2019 Annual Conference Proceedings

Transforming Decision Sciences Through Emergent Technologies

New Orleans, LA

50th Annual Conference

November 23 – 25, 2019

ISBN: 978-0-578-62648-2
ISSN: 2471-884X

CELEBRATING 50th ANNIVERSARY

DSI
DECISION SCIENCES INSTITUTE
Scholars · Educators · Problem Solvers

Table of Contents

DECISION SCIENCES INSTITUTE

On The Effects of Oil Price Fluctuation, Geopolitics, And the Global Financial Crisis On the dependence Structure Between Stock Markets: New Evidence From a Time-varying Copula Model

Noureddine Benlagha
Department of Finance and Economics,
College of Business and Economics,
Qatar University,
Email: nbenlagha@qu.edu.qa

ABSTRACT

This paper examines the impact of economic and geopolitical events on stock market dependence using copula statistical theory. We focus in particular on the time periods before and after oil price shocks and the 2017 political crisis among the Gulf Cooperation Council members (i.e. the Qatari blockade), to examine possible changes in the dependence structure between the Qatari stock market and other stock markets in a group of emerging and advanced countries. Our findings show that the time varying dependence behaviour is associated strongly with the oil price crash and the geopolitical crisis in the region. The findings also show that the 2008 GFC has a stronger impact than the price shocks and political crisis.

KEYWORDS: Stock markets, Oil prices, financial crisis, Geopolitics, Qatari blockade, Market dependence structure, Copula, GCC

JEL: C10; G01; G15

ACKNOWLEDGMENT
This work was supported by Research University Grant from Qatar University under the grant number QUUG-CBE-DMM-17/18-5.

INTRODUCTION

Over the past three decades, the member countries of the Gulf Cooperation Council (GCC) have made a significant effort to improve the profitability, the efficiency, and the size of their capital markets. A sustainable privatization program and independent capital market regulators have improved the integration among the GCC stock markets, strengthening each financial market and attracting local capital investment in the GCC region (Basher, Nechi and Zhu, 2014; Alotaibi and Mishra, 2017). However, at the same time, such rapid financial integration among the GCC countries may be exposing them to financial, economic, and political crises that could result in a greater restriction on capital mobility (Karolyi and Stulz, 2003; Balcilar et *al*., 2015 and Belkhir et al., 2017).

In the GCC area, two major events have occurred since the 2008 global financial crisis. Since June 2014, oil prices have dropped sharply after a period of stability at around $105 per barrel. The prices have remained low, dropping to $26.19 per barrel on 11 February 2016. The average oil price in the period between June 2014 and the end of 2017 was about $50 per barrel. Multiple causes as well as the consequences of such oil price fluctuation have been identified and discussed in the recent literature (Baffes *et al.*, 2015; Fantazzini, 2016; Prest, 2018).

Further, on 5 June 2017, the country quartet of Saudi Arabia (SA), the United Arab Emirates (UAE), Egypt, and Bahrain proclaimed they were breaking diplomatic and economic ties with Qatar, citing many political reasons and accusations. Consequently, the transcontinental imports of goods to Qatar have been handicapped since the suspension of all air and sea travel links between Qatar and the quartet. The sanctions have had diverse impacts in the short run, particularly on trade, capital, and human flow.

While the economic aspects of the oil price crash and the Qatari blockade have been examined in existing literature, the association between these two particular events and the stock market integration in the region has not yet been investigated. However, several arguments advocate that such a relationship or linkage exists. First, some previous studies have documented that stock market dependence is affected by financial and economic crises. Rizvi *et al. (2017)* claimed that throughout the Asian financial crisis (1997), the integration of the Japan stock market with global stock markets declined. (Yang et al., (2015) document that the dependence between international stock markets has been affected by the global financial crisis of 2008. Second, political crises, such as wars, terrorist attacks, and other catastrophic events, may affect the economic and financial integration between countries. Enders and Sandler, (1996) and Enders et al. (2002) find that acts of terrorism negatively affect the foreign direct investment (FDI) flows of the countries involved. Frijns et al. (2012) conclude that political crises with certain characteristics generally reduce the level of stock market integration in these regions.

This paper falls under the recent strand of literature that analyses the impacts of economic and political crises on cross-market linkages. Specifically, we examine the stock market dependence structure between the Qatari stock market and other country stock markets, focusing on the period of the pre- and post-oil price shock and the pre- and post-2017 GCC political crisis period.

The study of the Qatari stock market is critical for several key reasons. First, GCC countries are among the major oil producers in the word and Qatar is one of the largest suppliers of natural gas reserves to the global economy, following Iran and Russia. Therefore, its stock market may be affected by oil price fluctuations. Second, since May 2014, Qatar has been reclassified from a frontier market to an emerging market and become a more attractive destination for investors and for portfolio managers. Third, despite its segmentation from international markets, the Qatari financial market is very sensitive to regional political events. From this perspective, the Qatar financial market offers an interesting case study for analysing the effects of a political crisis on regional and international financial integration through an investigation of the impact of the blockade, commonly known as the Qatar-Gulf crisis (Q-GC), which begin 5 June 2017. Fourth, few empirical studies have examined the dependence between Qatar and other financial markets by employing the copula approach, which is used in describing the dependence between random variables**.** Finally, to the best of our knowledge, no empirical research regarding the dependence between the Qatar stock market and other stock markets can be found in the literature. The findings of the paper have important implications for cross-market risk managers and international asset pricing.

This paper fills the void in the extant literature by investigating the dependence structure between Qatar and the GCC emerging stock markets and developed stock markets and by identifying the different regimes that resulted from the decrease/fluctuation in oil prices since 2014 and the Q-GC since June 2017. In this paper, we discuss the impact of these two major economic and political events that can explain the presence of structural breaks when examining the dependence structure between the Qatari stock market and the set of emerging and developed stock markets.

To depict a significant dependence structure among the stock markets, we use the Student's-*t* copula construct (see (Demarta and McNeil, 2007 for more details). This construct is a well-known measure that can generate joint symmetric tail dependence and capture the positive and negative extreme dependencies. Furthermore, we employ a time-varying copula model to describe the effects of exogenous events on the dependence structure of the studied series.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the models and estimation method. Section 4 presents the data and some statistics. Section 5 provides the empirical results. Finally, Section 6 concludes.

LITERATURE REVIEW

During the last few decades, substantial empirical research has investigated the dependence structure among financial markets because of its implications for investors, portfolio managers, hedging strategy design, and risk management (e.g. Mokni and Mansouri, 2017; Al Nasser and Hajilee, 2016; Hussain and Li, 2018; Yang *et al.*, 2015); Abul, Nechi and Zhu, 2014; Hao and He, 2018).

Specifically, several studies investigate the bi-directional relationship between stocks and financial markets (Coelho et al., 2007; Lucey and Muckley, 2011; De Groot et al., 2012; Claus and Lucey, 2012; Karanasos et al., 2016; Yang and Hamoi, 2013; Aloui and Hkiri, 2014). However, most of these examine the dependence structure between developed and emerging stock markets (Mokni and Mansouri, 2017); Yang and Hamoi 2013; Hussain and Li, 2018), emerging and frontier stock markets (De Groot et al., 2012), and developed, emerging, and frontier stock markets (Yang et al., 2015; Mimouni et al., 2016). Overall, the findings are mixed and seem to depend on the periods and the stocks markets investigated as well as the econometric methods used.

For instance, Hatemi (2012) examines the financial linkages between the UAE and the US stock markets between 7 June 2005 and 5 April 2011. He finds evidence of the integration between the markets and that the degree of integration is higher during periods when the two markets are falling rather than rising. Claus and Lucey (2012) conclude that segmentation levels vary among the Asia Pacific stock markets, indicating a limited degree of integration among them.

Al Nasser and Hajilee (2016) apply the bounds testing technique to explore the existence of long- and short-run linkages between emerging stock markets (Brazil, China, Mexico, Russia, and Turkey) and advanced financial markets (the US, UK, and Germany). They find evidence of the existence of short-run integration between emerging and advanced stock markets. In their long-run exploration, they find that Germany's stock market is the only determinant of integration with emerging economies. Using Archimedean copulas, Yang et al. (2015) inspect the dynamic dependence among developed, emerging, and frontier stock markets. Their findings indicate evidence of strong dependence between the investigated European stock markets and a weak dependence between frontier markets and others. The authors also find that dependence is asymmetric among these markets.

In a more recent contribution, Mokni and Mansouri (2017) analyse the conditional dependence between international stock markets in three different time intervals: pre-crisis, subprime and debt crisis, and post the two crises. Using a long memory GARCH-copula model approach, they find a significantly strong correlation between European markets.

Taking China as a case of study, Hussain and Li (2018) examine the dependence pattern of China on six developed stock markets (the US, Canada, Japan, Australia, the UK, and Germany). Using dynamic copula constructs and extreme value theory, the authors find that the dependence varies over time. The findings show a substantial dependence between China and each of the Asian and European countries, and, conversely, a weak dependence between China and the US.

Other empirical studies examine stock market dependency in other regions. For example, Mensah and Alagidede (2017) apply copula models to explore the dependence structure of African stock markets (South Africa, Egypt, Kenya, and Nigeria) with the UK and US markets. The results indicate that the dependence is dynamic; it is weak in Egypt, Kenya, and Nigeria and strong in South Africa. In addition, the results support that the dependence between the studied stock markets is asymmetric, suggesting that stock return co-movement diverges in bear and bull markets.

However, there is no study that examines the dependence structure in the GCC countries. This means that few empirical studies have considered the relationships between the GCC member stock markets and other international stock markets.

In the context of the GCC countries, most studies focus on crude oil and its relationship with stock markets (e.g. (Ma et al., 2014; Nusair, 2016; Jouini and Harrathi, 2014; Mohanty et al., 2011; Arouri, Lahiani and Nguyen, 2011). The conclusions of these studies are mixed. For example, Naifar and Al Dohaiman (2013) find that oil prices have a significantly positive impact on GCC stock markets. Arouri, Lahiani, and Nguyen (2011) and Akoum et al. (2012) reveal similar findings. Nusair (2016) investigates the dynamic relationship between oil price shocks and GCC countries. He finds that positive oil price changes are significant in all the GCC countries with the expected positive sign. In addition, he finds that negative oil price changes are significant only in Kuwait and Qatar. In contrast, Al Janabi, Hatemi-J, and Irandoust (2010) find that the relationship between oil price and stock markets is insignificant.

However, only a few past studies examine the dependence between the GCC members' stock markets and other international stock markets. For instance, Abul, Nechi, and Zhu (2014) employ copula theory to analyse the dependence structure of seven GCC stock markets (Bahrain, Kuwait, Oman, Qatar, SA, Abu Dhabi, and Dubai) in terms of daily returns between 2004 and 2013. They find significant evidence indicating that GCC stock markets are characterized by persistent and strong volatility. They also find that dependence is asymmetric. Another study (Aloui and Hkiri, 2014) investigated empirically the short- and long-term relationships among six GCC stock market returns (Bahrain, Kuwait, Oman, Qatar, SA, and UAE) over the 2005-2010 period. The authors found that co-movement depends on both frequency and time and is strongly affected by a financial crisis.

Some other studies have documented that stock market dependence is influenced by financial and economic crises. For example, Rizvi and Arshad (2017) argue that throughout the Asian financial crisis, the integration of the Japan market with global stock markets declined. Yang et al. (2015) document how the global financial crisis of 2008 affected the dependence between international stock markets. These same results are supported in the works by Mensi et al. (2014) and De Groot et al. (2012). Recently, Karanasos et al. (2016) focused on two financial crises: the 1997 Asian financial crisis and the 2008 global financial crisis. They observe how the interdependence among eight stock markets intensifies during crises periods due to contagion effects. In addition, Mokni and Mansouri (2017) provide evidence that the dependence among European stock markets is strongly affected not only by the global financial crisis but also by the European crisis of 2010-2011. The dependence increases during crisis periods. In contrast, Alotaibi and Mishra (2017) show that the degree of integration in the GCC stock markets declines after the 2008 global financial crisis. In addition, Fayyad and Daly (2011) show that Qatar, UAE, and the UK are more responsive to oil shocks than the US and the other GCC financial markets. Extending this existing literature, our study seeks to identify the different regimes resulting from the decrease and fluctuations in oil prices since 2014.

While the influence of financial economic crises on the dependence between stock markets has been investigated, the impact of political crises has not. In existing literature, acts of terrorism are considered as the main proxy for political crises (see, for instance, Kollias et al. 2011; Kollias et al., 2013; Mnasri and Nechi, 2016; Kolaric and Schiereck, 2016). In these studies, the impact of the political crisis on the financial market is investigated based on the marginal distribution of the stock prices or returns. None of these studies has investigated the political crisis effects on the dependence or on the conditional correlations between two distributions. Moreover, among the existing research, this empirical paper is the first to study the impact of the political crisis on the dependence between stock markets by considering the Qatari blockade as a proxy for political crises.

Motivated by examining the dependence structure between Qatar and other financial markets, this paper advocates the use of copula functions, which have become popular as flexible models in this field. To detect an interpretable dependence structure among financial data, we use the Student's t copula model (see Demarta et al., 2005 for more details). This model is a well-known measure that can generate joint symmetric tail dependence and capture the positive and negative extreme dependencies. We use the time-varying copula construct that allows us to model extreme events such as the oil shocks prevalent since mid-2014 and the Qatar blockade since June 2017.

METHODOLOGY

The empirical objective of this paper is to demonstrate that it is economically and statistically relevant to jointly model the behaviours of two different stock returns using a flexible copula approach that accommodates the presence of tail dependence and the time variation due to various economic and political events. In a first stage, we estimated a battery of static copula models and their corresponding time varying constructions (Gaussian, Student'*s*, and *SJC*). Then we selected the suitable copula by using the log likelihood and AIC information criteria. As shown in the appendix A, the static student's-*t* copula and its dynamic version perform better in almost all pairs of return series.

In this section, the $ARMA(p,q) - GARCH(1,1) - skT$ model is first described for the marginal process of each stock return. Then, the static student's-*t* copula and the used estimation method are introduced. The third subsection describes the time varying dependence structure modeling using copula.

Marginal processes

In the empirical financial studies, asset returns $r_{i,t}$ display common features called stylized facts, as for example autocorrelation and heteroscedasticity. The $GARCH$ specifications are known to be the suitable models for these financial return series (Bensaïda, 2015). Likewise, the optimal selection of an appropriate conditional distribution function is fundamental for the robustness of the copula construction (Joe, 2005; Ning, 2010; So and Yeung, 2014). Consequently, among various $GARCH$ specifications we select an $ARMA(p,q) - GARCH(1,1) - skT$ which is one of the suitable models allowing to capture the conditional heteroscedasticity of financial asset returns (Fei et al. 2017).
The marginals of the daily index returns, denoted r_t, consist of estimating the following model:

$$r_t = a_0 + \sum_{i=1}^{p} a_i r_{t-i} + \sum_{i=1}^{p} b_j \varepsilon_{t-j} + \varepsilon_t, \quad (1)$$

$$\sigma_t^2 = c_0 + c_1 \sigma_{t-1}^2 + d_1 \varepsilon_{t-1}^2, \quad (2)$$

Where the filtered returns $x_t = \varepsilon_t / \sigma_t$, $t = 1, \dots, T$, are assumed to be iid and $x_t \sim skT(0,1,\nu)$, with $\nu > 2$ corresponds to the number of degree of freedom distribution as proposed by (Hansen, 1992).

The estimation of the parameters of this $ARMA(p,q) - GARCH(1,1) - skT$ model are performed by the maximum likelihood (ML) method. Following the prior empirical studies (see for instance Jondeau and Rockinger, 2006; Fei et al. 2017) we fix the lag orders in the conditional variance (equation 2) at one since this specification is employed to capture the conditional heteroscedasticity of financial asset returns. We also used the Akaike Information criterion (AIC) to determine the optimal AR and MA lag order combination $(p\ ,q)$.

The static student's-*t* Copula

Any d-dimensional function, defined on $[0,1]^d$, with marginal distributed as standard uniform is commonly known as a d-dimensional copula C. Sklar's theorem states that all density funaction F with margins $F_1, \dots, F_d$can be expressed as

$$F(x_1, \dots, x_d) = C\big(F_1(x_1), \dots, F_d(x_d)\big) \quad (3)$$

Any copula C can be employed join various univariate density functions $F_1, \dots, F_d$ to create a multivariate density function F with margins $F_1, \dots, F_d$.

$$C(u) := C(u_1, \dots, u_d) = F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right) \quad (4)$$

In the equation (4), F_1^{-1} correspond to the quantile functions of the margins. The copula C may be considered as the density function of the component wise probability transformed random vector$(F_1^{-1}(\mathrm{X}_1), \dots, F_d^{-1}(\mathrm{X}_d))'$.

The standardization of the marginal distributions guarantees that the copula remains invariant. Particularly, the student'*s-t* copula construction noted by (υ, μ, Σ) is identical to that of $t_d(\upsilon, 0, \Sigma)$ distribution where ρ is the correlation matrix implied by the dispersion matrix Σ. The unique copula is thus given by

$$C_{\nu,\rho}^{t}(u) = \int_{-\infty}^{t_\upsilon^{-1}(u_1)} \dots \int_{-\infty}^{t_\upsilon^{-1}(u_d)} \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^{\mathrm{d}}|\rho|}} \left(1 + \frac{x'\rho^{-1}x}{\nu}\right)^{-\frac{\nu+d}{2}} dx, \tag{5}$$

In this formula υ corresponds to the number of degree of freedom and t_υ^{-1} is the quantile function of a standard univariate student (t_υ) distribution. The notation can be simplified when considering the bivariate case and the notation becomes $C_{\nu,\rho}^{t}$.

Practically, the simulation of the student's-t copula is possible and can be performed by generatating a multivariate t-distribution random vector $\mathrm{X} \sim t_d(\upsilon, 0, \rho)$ by empolying the normal mixture construction (3) and then return a vector $U = (t_\upsilon(\mathrm{X}_1), \dots, t_\upsilon(\mathrm{X}_d))'$. It is also noted that the density of the t copula may be achieved using the equation (4) and implemented for the estimation resolutions has the form. This Student density distribution can be represented as follows:

$$c_{\nu,\rho}^{t}(u) = \frac{f_{\upsilon,\rho}(t_\upsilon^{-1}(\mathrm{u}_1), \dots, t_\upsilon^{-1}(\mathrm{u}_d))}{\prod_{i=1}^{d} f_\upsilon(t_\upsilon^{-1}(\mathrm{u}_i))}, u \in (0,1)^d, \tag{6}$$

where $f_{\upsilon,P}$ corresponds to the joint density of a $t_d(\upsilon, 0, \mathrm{P})$-distribution random vector and $f_\upsilon$is the density of the univariate standard t-distribution with υ degrees of freedom.

Properties of the student's-*t* Copula

To describe the dependence structure between two distributions by using any copula construction, it is important to evaluate Kendall's rank correlation allowing the depiction of the correlation between two entire marginal distributions. It is also important to assess the tail dependence coefficients to describe the pattern of dependence at the right and let tails of the distributions.

- Kendall's Rank Correlation

$$\rho_\tau(X_1, X_2) = E\left(sign\left(X_1 - \widetilde{X_1}\right)\left(X_2 - \widetilde{X_2}\right)\right), \tag{7}$$

where $\left(\widetilde{X_1}, \widetilde{X_2}\right)$ is a second independent couple with the identical distribution as (X_1, X_2).

It is important to note that kendall's tau rank correlation ρ_τ depends only on the copula but no (and not on the marginals (X_1 and X_2). The Kendall's tau rank correlation is presented by:

$$\rho_\tau(X_1, X_2) = 4\int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1. \tag{8}$$

- Tail Dependence Coefficients

The coefficient of upper tail dependence of X_1 and X_2 is defined by

$$\lim_{q \to 1^-} P\left(X_2 > (F_2^{-1}(q) \backslash X_1 > F_1^{-1}(q)\right) = \lambda_u, \tag{9}$$

provided the limit $\lambda_u \in [0,1]$ exists, and the coefficient of lower tail dependence, defined by

$$\lim_{q\to 0^+} P\left(X_2 > (F_2^{-1}(q)\backslash X_1 > F_1^{-1}(q)\right) = \lambda_l, \tag{10}$$

provided the limit $\lambda_l \in [0,1]$ exists. Thus these coefficients are limiting conditional probabilities that both margins exceed a certain quantile level given that one margin does.

The copula-based forms are $\lambda_u = \lim_{q\to 1^-} \frac{\bar{C}(q,q)}{1-q}$, $\lambda_l = \lim_{q\to 0^+} \frac{C(q,q)}{q}$, (11)

where $\bar{C}(u,u) = 1 - 2u + C(u,u)$ is the survivor function of the copula.

Estimation of the Student's- t copula

Writing $X_i = \left(X_{i,1}, \ldots, X_{i,d}\right)'$ for the i^{th} data vector, the method involves estimating the j^{th} marginal df F_j by

$$\hat{F}_j(x) = \frac{1}{n+1}\sum_{i=1}^{n} 1_{\{X_{i,j}\le x\}}. \tag{12}$$

The pseudo-sample from the copula is then constructed by forming vectors $\hat{U}_1, \ldots, \hat{U}_n$

$$\text{where } \hat{U}_i = \left(U_{i,1}, \ldots, U_{i,d}\right)' = (\hat{F}_1(X_{i,1}), \ldots, \hat{F}_d(X_{i,d}))'. \tag{13}$$

We notice that, even though the marginal vectors $X_1, \ldots, X_n$ are independent and identically distributed, the pseudo-sample data are dependent, because the marginal estimates $\hat{F}_j$ are issued from the entire original data vectors through the univariate samples $X_{1,j}, \ldots, X_{n,j}$.

Under some mathematical assumptions we can use the maximum likelihood method to estimate the parameters ν and ρ of $\text{student's} - t$ copula. The estimates are found by maximizing

$$\log L(\nu, \rho; \hat{U}_1, \ldots, \hat{U}_n) = \sum_{i=1}^{n} log c_{\nu,\rho}\left(\hat{U}_n\right) \tag{14}$$

Where $c_{\nu,P}^{t}$ denotes the density of the t copula in (6).

The time varying t-copula

There are evidence that dependence among financial assets does not remain constant over time (see Bouyé and Salmon, 2009) and Longin and Solnik, 2001). The dependence dynamics have important implications on the portfolio diversification reallocation strategies and help to identify how two assets behave jointly in various economic and financial conditions. Time-varying copulas are commonly considered as the dynamic generalizations of a Pearson correlation or Kendall's tau. In practice, time-varying copulas are frequently supposed to follow an autoregressive moving average (ARMA) (p, q) process, Patton (2006). Especially for the student's-t copula, the time varying copula construction is expressed as follows:

$$\rho_t = \Lambda\left(\Psi_0 + \Psi_1\rho_1 + \Psi_2\frac{1}{q}\sum_{j=1}^{q} \Phi^{-1}\left(u_{t-j}\right).\sum_{j=1}^{q} \Phi^{-1}\left(v_{t-j}\right)\right) \tag{15}$$

Where $\Lambda(x) = (1 - e^{-x})(1 + e^{-x})^{-1}$ is the modified logistic transformation. This transformation insure that the time varying coefficient of correlation be comprised in the interval $[-1,1]$.

DATA AND DESCRIPTIVE STATISTICS

Our empirical investigation relies on data from *Thomson Reuters Eikon* enhanced for the period spanning from August 1998 to June 2018. There are 14 countries in the sample divided into three groups: The first consists of the leading indices in GCC countries (Qatar, Saudi Arabia, United Arab Emirates, Oman, Kuwait and Bahrain. The second group is composed of four developed countries (USA, France, Germany and UK). The last group consists of four emerging financial markets (China, Turkey, Russia and India). The choice of these countries is based on the strength of the economic, financial and political relationship between Qatar and the others. The period of the empirical investigation is chosen to include the periods of pre and post oil price shocks and the pre and post 2017th GCC political crisis. This particular period covers also the global financial crisis (2007-2008) which significantly affected the volatility and the dependence between the stock markets of different nations.

The returns are calculated from the composite index of the major stock exchange in each country. Table 1 reports the summary statistics for the data series. The table shows that the GCC countries have the highest average returns among the three groups of countries studied (0.017 for GCC countries, 0.006 for the developed countries, and 0.0112 for the emerging ones). The highest volatility, measured by the standard deviation, is detected for the return series of the emerging countries (2.05); however, the GCC countries have the lowest variability with a standard deviation of 1.22. The comparison among countries shows that the Qatari stock market offers the highest average returns (0.033) and the highest volatility (2.507) among all the markets studied. Table 1 also shows that most return series exhibit skewness and excess kurtosis.

The Jarque-Bera test confirms this departure from the normality. The Ljung-Box test results suggests the presence of serial correlation in almost all return series. Moreover, the autocorrelation of the squared return series (Q^2 (10)) reveals that the dependence at long horizon is high enough to suggest a presence of ARCH effects. In addition, to test the stationarity, the ADF and PP unit root tests are performed and they fail to reject the hypothesis of stationarity for all return series indicating that the log first differencing is satisfactory to induce stationarity in prices.

The simplest way to describe the interdependence between financial markets is the use of the unconditional correlation coefficient (i.e., Pearson’s coefficient of correlation r) matrix. Graphically a scatter plot is also useful in describing the patterns of unconditional correlations.

The objective is to examine the direction and magnitude of market linkage or interrelationship without necessity to explore the driving forces, such as causalities, of the studied linear relationship. From a financial viewpoint, (Cheung and Ng, 1996)Cheung and Ng (1996) suggest that statistically significant correlation coefficients between inter-regional stock markets on the equal calendar day is considered as indication of spillover effects from the market to another.

Figure 1 displays the scatter plot and reports the correlation coefficients of daily returns and their respective significance levels.

The scatter plots clearly indicates the presence of extreme correlation and, thus, a presence of tail dependence for almost the return pairs. The tail dependence occurs in both the upper right and the lower left tail of the returns distribution pairs.

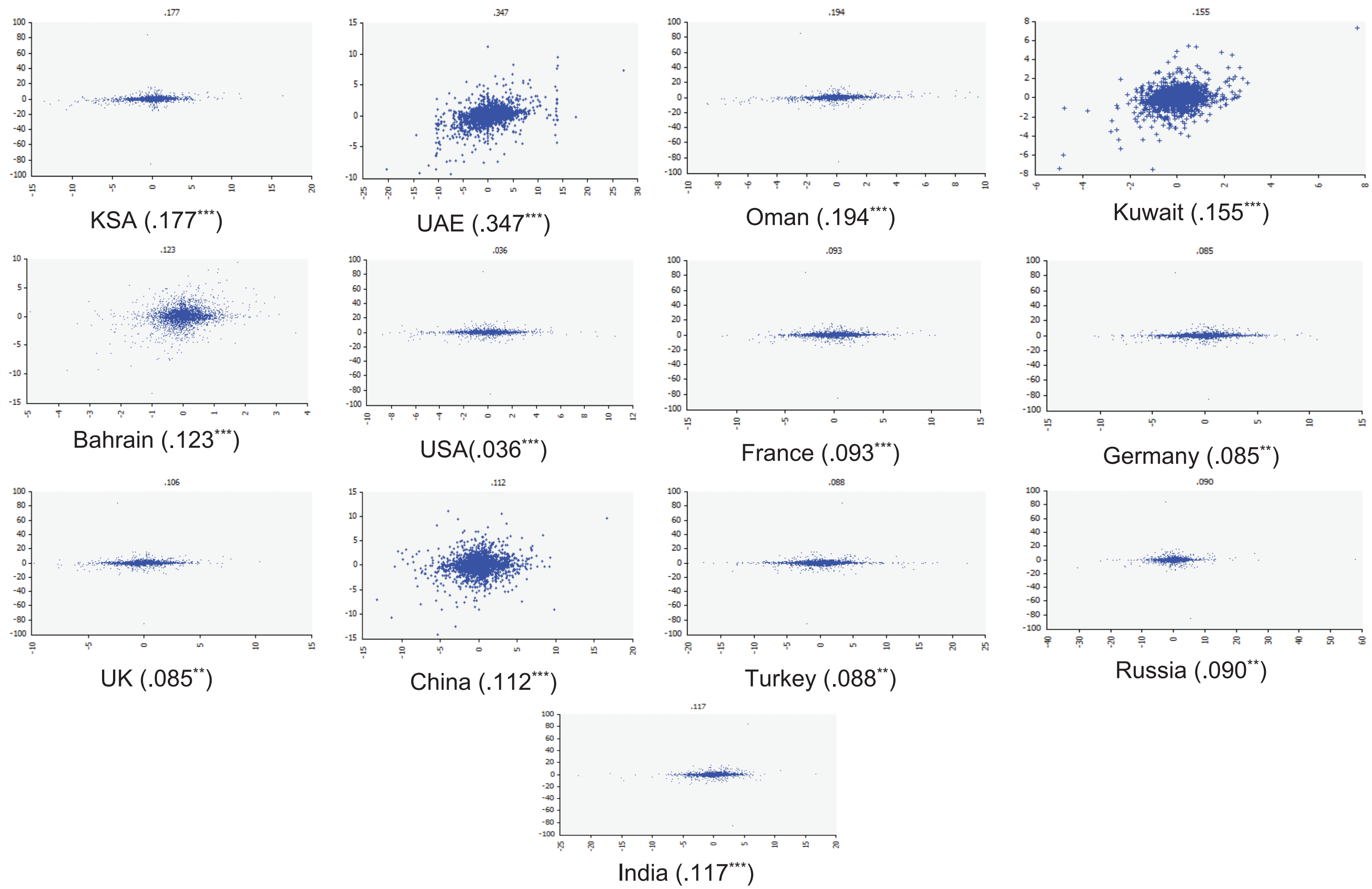
Moreover, the correlation patterns between returns seems to be non-linear. Accordingly, the Pearson’s correlations may be over- or under-estimated.The reported values show that the correlations between Qatari daily returns and the other GCC countries are positive and significant. On average, the correlation of the Qatari market equals 0.19; this value is higher than that obtained by Bley and Chen (2006) and lower than that obtained by Basher et al. (2014). The difference is mainly due to the periods studied and the frequency of the data in the previous studies. The Qatari market is more closely linked to the other GCC member markets than to the other emerging and developed markets studied. This indicates that the market integration is stronger locally (GCC area) than globally.

Table1: Summary statistics and ADF and PP unit root tests.

	Mean	Median	Max	Min	Std. dev.	Skuness	Kurtosis	. J–B	Q (10)	$Q^2(10)$	ARCH (10)	ADF	PP
Panel 1. GCC countries													
Qatar	0.033	0.042	84.423	-85.80	2.507	-0.543	624.90	0.000	16.197	2536.2	597.59	-47.64	-47.45
KSA	0.028	0.099	16.399	-13.490	1.52	-1.007	16.990	0.000	1041.2	2001.7	34.4	-22.77	-61.497
UAE	0.0315	0.0532	10.219	-12.157	1.769	-0.127	8.426	0.000	1149.3	1169.6	311.670	-37.879	-57.10
Oman	0.008	0.016	9.482	-8.699	0.991	-0.398	20.701	0.000	20.463	3083.5	171.20	-52.555	-52.342
Kuwait	1.6E-05	0.00017	0.050	-0.049	0.007	-0.048	8.031	0.000	10.884	273.06	53.24	-36.648	-36.79
Bahrain	0.005	0.000	3.613	-4.920	0.566	-0.388	9.325	0.000	30.991	328.42	119.99	-51.631	-54.096
Panel 2. Developed countries													
USA	0.025	0.047	10.089	-8.695	1.165	-0.215	10.680	0.000	6.401	3409.9	264.80	-65.77	-65.901
France	-0.009	-0.043	8.384	-8.868	1.439	0.162	6.459	0.000	17.54	1506.1	66.54	-62.355	-62.40
Germany	0.016	0.077	10.640	-7.552	1.428	-0.109	7.045	0.000	25.290	2023.7	96.144	-62.694	-62.796
UK	-0.008	0.022	9.384	-8.178	1.223	-0.116	8.069	0.000	34.265	2237.2	91.823	-47.236	-64.678
Panel 3. Emerging countries													
China	0.021	0.022	14.364	-18.369	1.947	-0.311	10.723	0.000	20.986	1416.0	62.675	-58.900	-58.924
Turkey	0.029	0.062	17.774	-19.979	2.387	-0.129	10.044	0.000	48.504	1136.4	61.746	-60.458	-60.526
Russia	-0.055	-0.073	19.178	-26.962	2.382	-0.162	19.880	0.000	40.621	1337.8	75.682	-59.547	-59.631
India	0.050	0.099	16.334	-13.054	1.504	-0.093	11.120	0.000	32.719	355.32	20.789	-44.905	-59.698

Notes: This table reports the main summary statistics for the return series for the 14 studied countries. The gives the p-value of the Jarque–Bera statistic test of normality and Q (10), Q^2 (10) which correspond to the Ljung–Box tests for 10th-order serial correlation in the returns and squared returns. The ADF and PP designate the Augmented Dickey and Fuller (1981) and Phillips and Perron (1988) unit root tests. *, **, and*** stand for 10%, 5% and 1% significance.

Figure 1: Scatter plots and correlation coefficients of market returns.



EMPIRICAL RESULTS

Marginal distribution estimations

Tables 2 and 3 report the maximum likelihood (ML) estimates for the conditional marginal distributions, ARMA-GARCH-*sk*t, of the daily returns for the GCC, emerging, and developed countries. The degrees of freedom (dof), ν, of the skt density is significantly small for almost all time series ranging between 3 and 6 for the GCC countries and between 6 and 9 for the emerging and developed countries. The low values of the dof suggest the index returns cannot be approximated by a normal distribution and that the developed countries' indices have fatter tails than the emerging and GCC countries' indices.

For the diagnostic checks of the marginal distributions, Panel B in Tables 2 and 3 show the result of the serial independence *Ljung–Box* (10) test on the first two moments of the estimated probability integral transformations, respectively: $(\hat{u}_t - \overline{u})^j$, $j \in \{1,2\}$, where $\hat{u}_t = F(\hat{x}_t)$, and $t = 1, \ldots, T$. Both the Q and the Q^2 statistics fail to reject the hypothesis of no serial correlation in the estimated residuals at conventional levels. Panel B in Tables 2 and 3 also reports the goodness-of-fit tests of Anderson-Darling and Cramer-von Mises for the null hypothesis that the transforms are $Uniform\ (0,1)$.

Tables 2 and 3 also show that the marginal models are well specified by the ARMA-GARCH-*sk*t specification and that the transforms are $Uniform\ (0,1)$. From the perspective of financial theory, the GCC, emerging, and developed countries' index returns display similar stylized facts as any other stock return series. Specifically, the investigated series show significant volatility clustering. Moreover, there is strong evidence of asymmetric behaviour among these returns. Finally, we see that the distribution of the stock returns fluctuates broadly according to the data series. Therefore, the standard GARCH specification under the highly flexible ARMA-GARCH-*sk*t model is appropriate to identify the true patterns of the studied index returns. Having well-specified marginal models is central to robust copula construction (Joe, 2005; Ning, 2010; So and Yeung, 2014).

Table 2: Estimation results for ARMA-GARCH-skt marginal models (GCC countries)

Countries	Qatar	KSA	UAE	Oman	Kuwait	Bahrain
Panel A. Parameter estimates (standard errors)						
Conditional mean						
Intercept	0.039*** (0.011)	-	0.057*** (0.020)	0.021** (0.009)	4.65E-05 (0.000)	0.007 (0.007)
AR1	0.221*** (0.016)	0.111*** (0.015)	0.050*** (0.018)	0.305*** (0.144)	-0.095*** (0.029)	0.079*** (0.016)
Conditional Variance						
Intercept	0.044*** (0.006)	0.031*** (0.006)	0.054*** (0.011)	0.021*** (0.003)	4.17E-06*** (1.49E-06)	0.013*** (0.003)
ARCH 1	0.599*** (0.055)	0.204*** (0.021)	0.159*** (0.017)	0.275*** (0.025)	0.101*** (0.022)	0.135*** (0.019)
GARCH 1	0.616*** (0.016)	0.815*** (0.013)	0.837*** (0.014)	0.739*** (0.016)	0.822*** (0.044)	0.847*** (0.017)
ν	3.231*** (0.145)	3.812*** (0.244)	5.113*** (0.466)	3.978*** (0.251)	6.228*** (1.167)	3.486*** (0.249)
Panel B. Goodness-of-fit tests						
Ljung–Box(10) test:						

1st moment	43.105	32.665	73.169	44.127	26.677	62.611
2nd moment	13.654	6.416	8.320	4.338	17.247	10.469
CvM test	3.346	10.542	2.365	6.549	0.501	8.317
AD test	20.710	57.067	14.433	39.280	3.447	46.422

Notes: Panel **A**. reports the parameters of the conditional mean and variance estimated using daily logarithmic return data, ν is the degree of freedom parameter for the student's *t* distribution. Standard errors are reported in parentheses. Panel **B**. reports the results of diagnostic tests. The Ljung-Box test is used to assess the null hypothesis of no autocorrelation up to a lag order of 10. CvM test and AD test are the Cramér–von Mises and Anderson–Darling goodness-of-fit tests. Here, the null hypotheses of the correct specification of the respective distribution function cannot be rejected at the 5% level of significance. *** and ** indicate significance at the 1% and 5% levels, respectively.

Table 3: Estimation results for ARMA-GARCH-skt marginal models (Emerging and Developed countries)

	China	Turkey	Russia	India	USA	France	Germany	UK
Panel A. Parameter estimates (standard errors)								
Conditional mean								
Intercept	0.083*** (0.023)	0.107*** (0.026)	-0.072*** (0.023)	0.096*** (0.018)	0.060*** (0.012)	-0.060*** (0.016)	0.076*** (0.017)	0.029** (0.013)
AR1	0.018 (0.017)	0.014 (0016)	0.004 (0.017)	0.023 (0.017)	-0.034** (0.017)	-0.036** (0.017)	-0.022 (0.017)	-0.046*** (0.017)
Conditional Variance								
Intercept	0.028*** (0.008)	0.036*** (0.009)	0.037*** (0.009)	0.027*** (0.007)	0.017*** (0.004)	0.016*** (0.005)	0.023*** (0.006)	0.017*** (0.004)
ARCH 1	0.069*** (0.008)	0.075*** (0.009)	0.096*** (0.010)	0.093*** (0.011)	0.097*** (0.011)	0.081*** (0.009)	0.097*** (0.010)	0.109*** (0.011)
GARCH 1	0.923*** (0.008)	0.919*** (0.008)	0.898*** (0.009)	0.898*** (0.010)	0.889*** (0.011)	0.913*** (0.009)	0.892*** (0.011)	0.882*** (0.011)
ν	6.936*** (0.808)	6.256*** (0.591)	6.536*** (0.706)	5.967*** (0.511)	7.945*** (0.904)	9.523*** (1.236)	9.154*** (1.256)	8.446*** (1.209)
Panel B. Goodness-of-fit tests								
Ljung–Box(10) test:								
1st moment	16.324	16.567	38.141	12.47	11.451	4.314	8.516	8.042
2nd moment	11.183	12.877	27.259	7.824	1.647	9.690	10.972	11.172
CvM test	1.783	1.539	4.722	3.311	1.213	0.853	1.601	0.998
AD test	10.270	9.965	27.156	17.537	7.957	5.796	11.508	7.925

Notes: See table 2

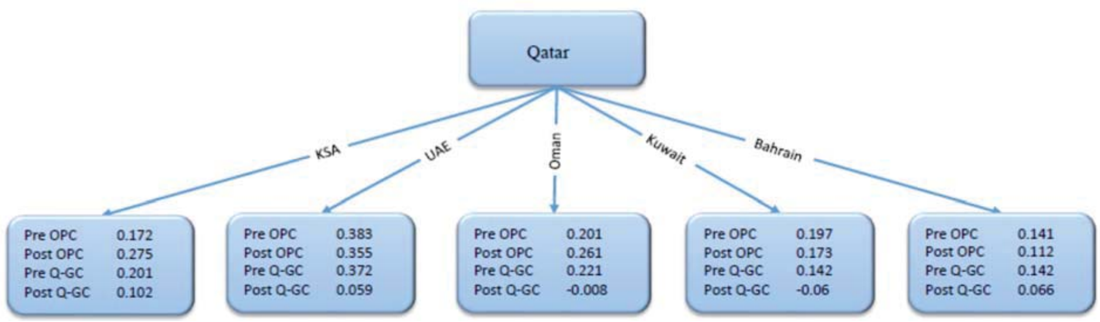
Static Student's-*t* copula results

Table 4 reports the results of the static Student's-*t* copula estimated parameters. Our analysis focuses on the discrepancy between the regional and international dependence structure.

Regional dependence

The results show that the order of the level of dependence for almost all the studied pairs is identical to that found by the linear coefficient of correlation examined previously. Specifically, for the GCC countries, the highest dependence is detected for the pair Qatar and UAE, followed by Oman and SA. However, a lower dependence is observed for the pair Qatar and Bahrain. These results reveal a strong integration of the Qatari stock market with the UAE and Oman stock markets and weaker integration of the former with the SA stock market. This low integration between the Qatari and Saudi markets is not specific to these two countries. Indeed, the Saudi market has the lowest levels of integration with each of the other GCC financial markets. Despite the fact that the Saudi stock market is the largest among the GCC members in terms of market capitalization, it remains inaccessible to foreign investors and risk managers (see, for instance, Alotaibi and Mishra, 2017).

Figure 2: The copula estimated dependence between Qatar and the other GCC stock markets



The results also show that the average of dependence significantly increases during the post-oil price shock period for the studied pairs in terms of the daily returns. In particular, in the GCC area, the increase in dependence is observed mainly for the following pairs: Qatar-SA, Qatar-UAE, and Qatar-Oman. However, the dependence with Kuwait and Bahrain, respectively, significantly drops after the oil shock.

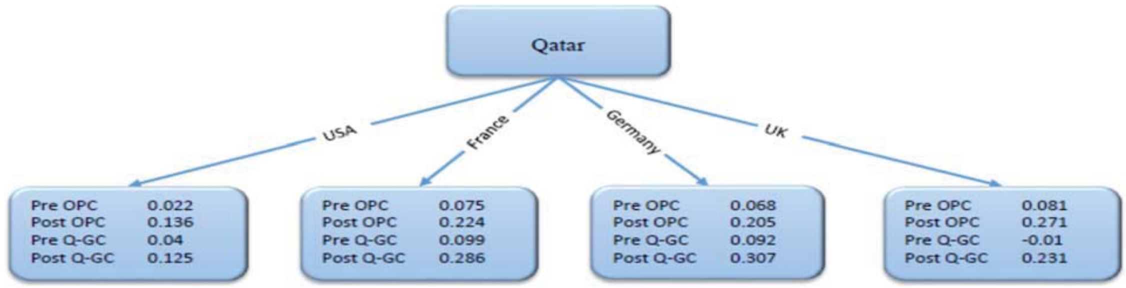
In the post-blockade period, the dependence between Qatari and each of the other GCC financial markets significantly diminishes, suggesting a decline in the level of regional integration. Interestingly, the dependence of the pairs of stock markets, Qatar-Oman and Qatar-Kuwait, becomes negative. Notably, these two countries are not among the blockade countries.

International dependence

Regarding the dependence between Qatari and each of the developed stock markets studied, the results indicate that the dependence between Qatar and US daily stock returns is the highest among the studied developed markets followed by the UK. The lowest level of dependence is between Qatar and Germany. In general, the dependence between the Qatar market and the developed financial markets is significantly weak. This implies that the Qatari financial market is isolated from most developed financial markets.

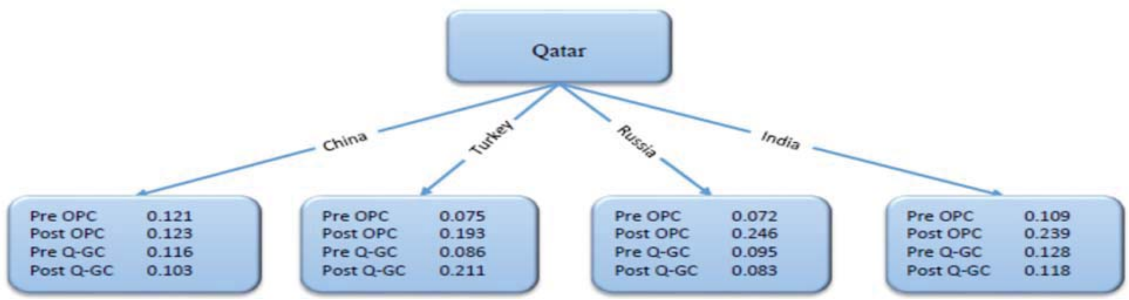
For the emerging markets, the highest level of dependence is between Qatar and China, followed by Qatar and India and Qatar and Turkey. Notably, there is a negative dependence between Qatar and Russia. According to the results, the Qatari financial market is integrated in the GCC area and is linked more closely to emerging markets than to developed ones.

Figure 3: The copula estimated dependence between Qatar and the developed stock markets



In the pre-oil price crash period, the increase in dependence is more significant with each of the developed countries (e.g. the dependence between Qatar and France increased from 0.075 before the oil price shock to 0.224 during the pre-crisis). Regarding the order, during the post-oil price shock, the Qatar financial market is linked more closely to the European countries (France, the UK, and Germany) than to the US market. The increase in dependence is also observed with all of the examined emerging countries.

Figure 4: The copula estimated dependence between Qatar and the emerging stock markets



As in the case of the developed markets, the order of dependence between Qatar and each of the emerging countries significantly changes after the oil price shock. After the shock, the highest level of dependence is between Qatar and Russia followed by Qatar and India and Qatar and Turkey. In contrast, the dependence between Qatari and China is the lowest among all the emerging markets studied.

Table 4: Static student's-*t* copula parameters estimates

	Entire period		**Pre oil price shock**		**Post oil price shock**		**Pre Blockade**		**Post Blockade**	
	ρ	ν	ρ	ν	ρ	ν	ρ	ν	ρ	ν
Panel 1. GCC countries										
KSA	0.198	3.767	0.172	3.819	0.275	4.70	0.201	3.857	0.102	99.96
UAE	0.360	3.301	0.383	2.855	0.355	5.164	0.372	3.383	0.059	14.05
Oman	0.219	4.682	0.201	4.96	0.261	4.367	0.221	4.742	-0.008	3.79
Kuwait	0.191	4.772	0.197	8.13	0.173	5.24	0.142	7.570	-0.060	99.9
Bahrain	0.049	27.975	0.141	7.034	0.112	14.46	0.142	7.570	0.066	11.22
Panel 2. Developed countries										
USA	0.137	5.358	0.022	8.306	0.136	4.774	0.040	8.203	0.125	4.328
France	0.004	23.97	0.075	18.19	0.224	6.18	0.099	13.75	0.286	6.963
Germany	0.063	20.863	0.068	25.39	0.205	6.01	0.092	17.33	0.307	8.756
UK	0.080	15.365	0.081	12.50	0.271	5.04	-0.01	99.99	0.231	7.308
Panel 3. Emerging countries										
China	0.161	9.241	0.121	4.455	0.123	4.76	0.116	4.483	0.1039	13.77
Turkey	0.065	93.919	0.075	16.85	0.193	20.53	0.086	16.87	0.2117	99.93
Russia	-0.025	11.440	0.072	6.81	0.246	4.66	0.095	6.681	0.083	12.35
India	0.128	5.278	0.109	4.97	0.239	4.37	0.128	5.212	0.118	3.356

Tail dependence

Tables 5a and 5b report the estimated tail dependence coefficients of the Student's-*t* distribution. The results indicate that for almost all the studied pairs of returns, the tail dependence parameters, $\tau_L and\ \tau_U$, of the Student's-*t* copulas are statistically significant, suggesting that the dependence at the lower and upper tails are symmetric.

The results also show that the pair Qatar and UAE has the highest tail dependence among the GCC countries followed by Qatar and SA. The lowest tail dependence is detected for the pair Qatar and Kuwait, and the tail dependence for the pair Qatar and Bahrain is statistically non-significant. The oil price shock does not affect the tail dependence parameter for all of the studied pairs. However, the blockade significantly affects the dependence structure in the GCC area. For all the pairs, except for Oman, the tail dependence parameter becomes non-significant, indicating that during the post blockade daily returns, the Gaussian copula is the most suitable to describe the dependence structure in the studied financial markets.

Table 5b shows that, except for Turkey, the tail dependence parameters between Qatar and the emerging countries' financial markets estimated for the full sample are statistically significant. The results also indicate that the tail dependence parameter between Qatar and the emerging financial markets significantly increases after the oil price shock. In contrast, the blockade significantly diminishes the dependence. More precisely, the tail dependence for the pairs Qatar and Turkey and Qatar and Russia becomes statistically insignificant.

Table 5b also indicates that for the developed countries, the tail dependence parameter estimated using the full sample is significant only for the pairs Qatar and the US and Qatar and the UK. It shows clearly that the oil price shock significantly increases the dependence at the tails between the daily returns of the Qatar market and the developed financial markets studied. Finally, the blockade event reduces remarkably the tail dependence between the Qatar market and the developed financial markets.

Table 5a: Estimation results for the tail dependence parameter (Qatar Vs GCC countries)

Countries	KSA	UAE	Oman	Kuwait	Bahrain
Panel 1. Full Sample					
$\tau_L = \tau_U$	0.136***	0.223***	0.107***	0.096**	1.819e^{-05}
Panel 2. Pre oil price shock					
$\tau_L = \tau_U$	0.126***	0.650***	0.0938**	0.035**	0.039**
Panel 3. Post oil price shock					
$\tau_L = \tau_U$	0.124***	0.136***	0.132***	0.079**	0.003**
Panel 4. Pre Blockade					
$\tau_L = \tau_U$	0.134***	0.2239***	0.106***	0.033*	0.033*
Panel 5. Post blockade					
$\tau_L = \tau_U$	1.0214e-^{14}	0.00233	0.080**	0.000	0.0065

Table 5b: Estimation results for the tail dependence parameter (Qatar Vs Emerging and Developed countries)

	China	Turkey	Russia	India	USA	France	Germany	UK
Panel 1. Full Sample								
$\tau_L = \tau_U$	0.0211**	1.243e-14	0.003**	0.028**	0.006**	3.97e-05	2.36e-04	0.0017**
Panel 2. Pre oil price shock								
$\tau_L = \tau_U$	0.088**	0.001	0.032**	0.071**	0.014**	6.51e-04	5.54e-05	0.004*
Panel 3. Post oil price shock								
$\tau_L = \tau_U$	0.080**	9.73e-04	0.116***	0.124***	0.082**	0.069**	0.068**	0.111***
Panel 4. Pre Blockade								
$\tau_L = \tau_U$	0.086	0.0011	0.037**	0.069**	0.016**	0.0034*	0.001*	0.000
Panel 5. Post Blockade								
$\tau_L = \tau_U$	0.0035**	1.29e-12	0.005**	0.001	0.094***	0.068**	0.047***	0.051*

Notes. ***, **, * denote significance at the 0.05, 0.10, 0.15 level respectively.

Overall, the static copula analysis provides two fundamental results. First, the oil price shock has a moderate impact on the dependence between Qatar and other GCC countries. Indeed, the strong integration between Qatar and the other GCC counties is not specific to Qatar. Actually, all the GCC member stock markets show the same integration patterns. This can be explained largely by the common GCC market status declared in 2008. The GCC common market status is aimed at forming a distinct location where the populations in the member countries have equal rights, including their engagement in several economic activities and services. It also calls for the unrestricted rights of property ownership and equity, movement of capital, and similar tax treatments. Hence, all these priorities, defined by the common GCC market status, are in favour of increasing the level of dependence and integration among the GCC members. Moreover, the different members of the GCC have taken similar actions to absorb the negative effects of the oil price crash.

The dependence of Qatar on both developed and emerging financial markets increased considerably after the blockade, indicating a growing level of international financial integration. This result is in contrast to that obtained previously on the integration between Qatar market and the other GCC stock market returns. The increase in the level of integration between Qatar and the emerging and developed financial markets after the blockade could relate to the government's support for the Qatari economy. Indeed, to deal with the undesirable effects of the GCC political crisis, the Qatari government has deposited billions of dollars into the local financial system, in particular, to prevent liquidity from drying up when Gulf residents started to withdraw their deposits. According to the central bank data, in August 2017, Qatar pumped 29.1 billion riyals into the financial market, growing public deposits from 273.5 billion riyals in July to 302.6 billion riyals. In August 2017, the central bank's international reserves and foreign currency liquidity increased to $38.3 billion, up from a low of $35.4 billion in July. In addition, the Qatar Investment Authority (QIA) has $300 billion in reserves that it could liquidate.

Further action has been to restructure supply chains in cooperation with new regional and international actors. In particular, Turkey's exports to Qatar grew by 42.3% in June and by 170.2% on an annual basis in July. Moreover, Qatar container lines have identified new shipping services via Oman, Kuwait, and the Indian subcontinent due to the blockade of Dubai as a trans-shipment port. Qatar is also looking to accelerate the effectiveness of its new Port with the objective of intensifying shipping routes to India, Oman, Turkey, and Pakistan. This would allow Qatar to import more food, construction materials, and other products. Securing food supplies is also one of the government's objective, as the country is dependent on imports for over 90% of its supplies.

These measures and the government's support have restored the trust of foreign investors in the Qatari financial system. New investors from several emerging and developed countries (e.g. Turkey, China, India) have replaced the GCC members. Consequently, the level of financial integration between Qatar and the international financial markets has increased. At the same time, the local integration (Qatar and the other GCC countries) has decreased significantly.

5.3 The time varying Student's-*t* copula results

In the static copula analysis, we compare the different regimes based on the fluctuation of oil prices since 2014 and the Q-GC crisis since June 2017. This first regime is based on the comparison between the pre- and post-events. The main technical finding in this first part sheds light on the fact that dependence between different markets is time varying. Moreover, the dependence between stock returns may be affected in various directions by economic and political events. Actually, the time varying copula analysis provides further information on the correlation patterns between the GCC members and the other stock markets during the period. The plots of the time varying parameters can provide interesting insights into the behaviour over time of the dependence between the Qatar stock market and the GCC, developed, and emerging stock market indices. The results of estimated time varying dependence parameters are reported in Tables 6a and 6b. These tables also report the estimated static Student's-t copula parameters for comparison.

The level of persistence is measured by the parameter Ψ_1, and the variation in the dependence process for the Student's-*t* copula is commonly captured by the parameter Ψ_2. As shown in Tables 6a and 6b, the persistence parameter Ψ_1 is significant at the 5% level in the time-varying student-*t* copula for almost all the stock return pairs. Moreover, this parameter is positive for all the pairs, except for Russia and the UK, in the post-blockade period. This result suggests a high level of persistence over time in the dependence behaviour between the Qatar market and the other international financial markets. This result is consistent with findings from previous studies (Chang and Hsueh, 2013; Wu and Liang, 2011; Wu and Lin, 2014).

The parameter Ψ_1 is also significant at the 5% level for practically all the pairs, showing considerable variations over time in the dependence between Qatar market and the other international stock markets. The positive values of this parameter for all countries suggest that the latest information on stock returns is a significant indicator for modelling the dynamic dependence between international stock markets.

Table 6.a: Estimation Results of constant and time varying dependence (Qatar Vs GCC, developed and emerging countries)

	Entire period					Before oil price crash 2014					After oil price crash 2014				
	Static STC		Dynamic STC			Static STC		Dynamic STC			Static STC		Dynamic STC		
	ρ	ν	Ψ_1	Ψ_2	Ψ_3	ρ	ν	Ψ_1	Ψ_2	Ψ_3	ρ	ν	Ψ_1	Ψ_2	Ψ_3
Panel 1. GCC countries															
KSA	0.198	3.767	0.527	0.232	-1.238	0.172	3.819	0.490	0.231	-1.522	0.275	4.70	-0.005	0.025	2.057
UAE	0.360	3.301	1.358	0.164	-1.985	0.383	2.855	0.663	0.148	-0.996	0.355	5.164	0.195	0.100	1.402
Oman	0.219	4.682	0.681	0.291	-1.808	0.201	4.96	0.470	0.272	-1.066	0.261	4.367	0.945	0.231	-2.004
Kuwait	0.191	4.772	0.674	0.257	-1.766	0.197	8.13	0.643	0.286	-1.974	0.173	5.24	0.622	0.326	-2.000
Bahrain	0.049	27.975	0.107	-0.061	0.093	0.141	7.034	0.345	0.326	-1.455	0.112	14.46	0.308	0.315	-2.014
Panel 2. Developed countries															
USA	0.137	5.358	0.682	-0.030	-2.057	0.022	8.306	0.0126	-0.013	1.489	0.136	4.774	0.694	0.140	-2.043
France	0.004	23.97	0.004	-0.069	-1.409	0.075	18.19	0.22	0.235	-1.607	0.224	6.18	1.132	-0.132	-2.09
Germany	0.063	20.863	0.164	0.107	-0.841	0.068	25.39	0.211	0.163	-1.514	0.205	6.01	0.996	-0.104	-2.086
UK	0.080	15.365	0.188	0.102	-0.591	0.081	12.50	0.034	0.034	1.483	0.271	5.04	1.352	-0.147	-2.103
Panel 3. Emerging countries															
China	0.161	9.241	0.566	0.230	-2.045	0.121	4.455	0.114	0.060	0.878	0.123	4.76	0.593	0.117	-2.003
Turkey	0.065	93.919	0.019	0.018	1.698	0.075	16.85	0.258	0.110	-1.628	0.193	20.53	0.803	-0.054	-2.076
Russia	-0.025	11.44	-0.114	-0.075	-1.915	0.072	6.81	0.032	0.045	1.432	0.246	4.66	0.923	0.148	-1.882
India	0.128	5.278	0.266	0.041	-0.272	0.109	4.97	0.496	-0.117	-1.990	0.239	4.37	0.986	0.177	-2.127

Notes. This table reports the results of constant and time varying dependence parameters. STC denotes the student's-t copula.

Table 6.b: Estimation Results of constant and time varying dependence (Qatar Vs GCC, developed and emerging countries)

	Entire period					**Before Blockade**					**After Blockade**				
	Static STC		Dynamic STC			Static STC		Dynamic STC			Static STC		Dynamic STC		
	ρ	ν	α	β	ω	ρ	ν	α	β	ω	ρ	ν	α	β	ω
Panel 1. GCC countries															
KSA	0.198	3.767	0.527	0.232	-1.238	0.201	3.857	0.593	0.225	-1.504	0.102	99.96	0.392	-0.114	-1.791
UAE	0.360	3.301	1.358	0.164	-1.985	0.372	3.383	1.163	0.048	-0.991	0.059	14.05	0.231	-0.439	-1.638
Oman	0.219	4.682	0.681	0.291	-1.808	0.221	4.742	0.601	0.276	-1.417	-0.008	3.79	0.023	-1.009	1.830
Kuwait	0.191	4.772	0.674	0.257	-1.766	0.142	7.570	0.380	0.303	-1.596	-0.060	99.9	-0.038	-0.130	1.357
Bahrain	0.049	27.975	0.107	-0.061	0.093	0.142	7.570	0.380	0.303	-1.596	0.066	11.22	0.140	0.120	-2.039
Panel 2. Developed countries															
USA	0.137	5.358	0.682	-0.030	-2.057	0.040	8.203	0.168	-0.078	-1.794	0.125	4.328	0.080	-0.138	1.566
France	0.004	23.97	0.004	-0.069	-1.409	0.099	13.75	0.299	0.236	-1.594	0.286	6.963	0.018	-0.045	2.066
Germany	0.063	20.863	0.164	0.107	-0.841	0.092	17.33	0.284	0.178	-1.494	0.307	8.756	1.524	-0.820	-1.719
UK	0.080	15.365	0.188	0.102	-0.591	-0.01	99.99	-0.043	0.123	-1.665	0.231	7.308	1.104	-0.373	-2.106
Panel 3. Emerging countries															
China	0.161	9.241	0.566	0.230	-2.045	0.116	4.483	0.137	0.073	0.690	0.103	13.77	0.506	-1.621	-0.913
Turkey	0.065	93.919	0.019	0.018	1.698	0.086	16.87	0.293	0.127	-1.612	0.211	99.93	1.038	-0.502	-2.061
Russia	-0.025	11.440	-0.114	-0.075	-1.915	0.095	6.681	0.038	0.038	1.517	0.083	12.35	0.029	-0.194	1.731
India	0.110	7.577	0.266	0.041	-0.272	0.128	5.212	$-6.8e^{-04}$	0.004	2.015	0.118	3.356	0.241	0.034	-0.281

Notes. See notes in Table. 6.a

Figure 5 illustrates the time path and average of the dependence structure between the Qatar market and the GCC, developed, and emerging stock markets, respectively. The area presented in grey corresponds to the post-oil price crash and the post-blockade periods. A third grey area corresponds to a period of high dependence between the studied stock returns. The average dependence is mostly weak (values near to zero) with uncommon exceptions. In the full dataset, the highest mean dependence (in absolute value) is found with the UAE with a value close to 0.4, followed by Oman, SA, and the other GCC stock markets. The lowest values are for Turkey and China. After the oil price crash, the average dependence increases for all the studied pairs, except for Bahrain. The increase of the average of dependence with the Qatari index is particularly substantial for the developed and emerging stock markets. For example, for the pair Qatar and the UK, the average dependence increases from 0.1 to 0.31 and for the pair Qatar and Russia, it increases from 0.089 to 0.247.

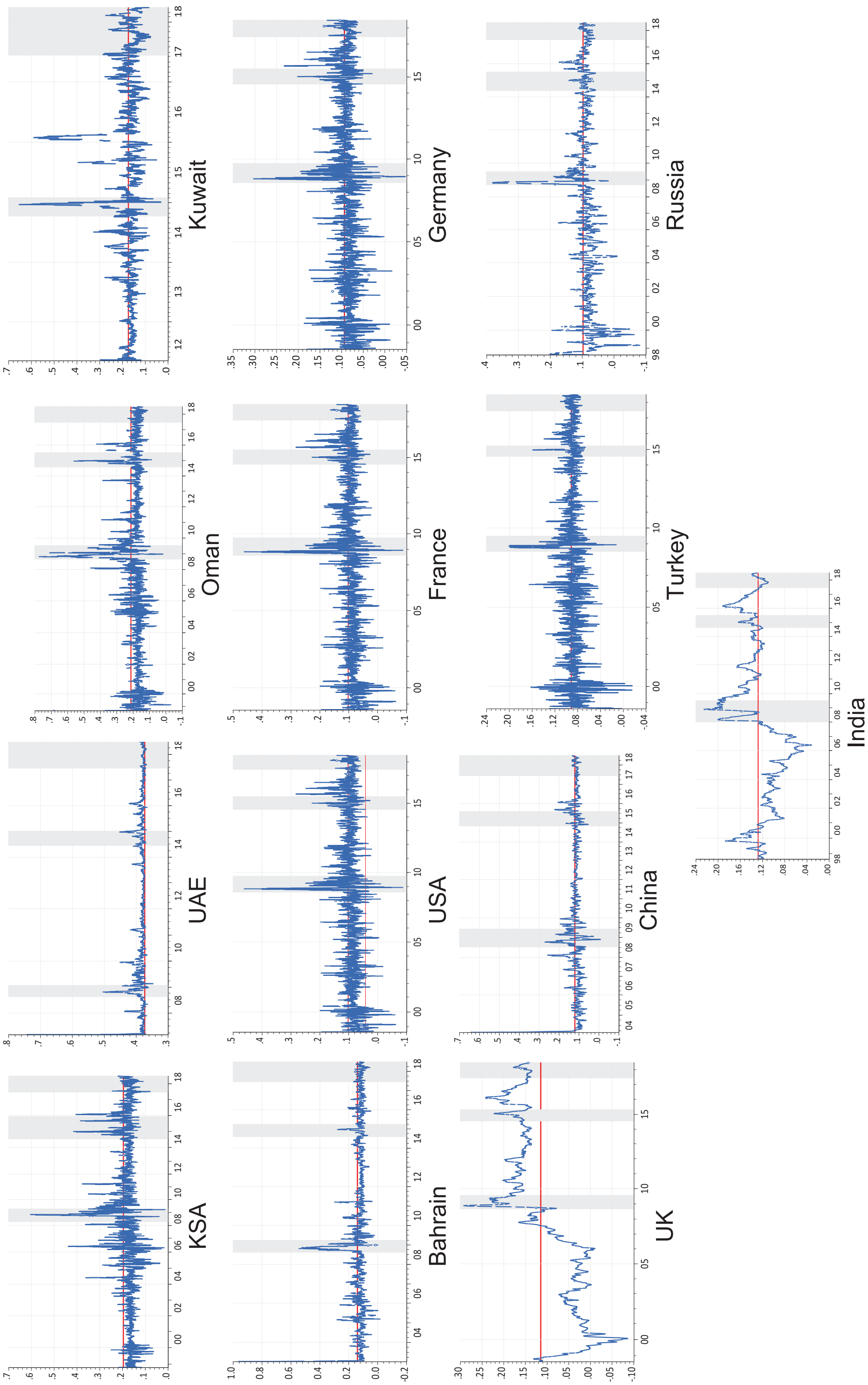
In the post blockade period, the average dependence decreases significantly for all the studied pairs. The mean of the dependence values in this period equals the values obtained for the full data.

Notably, Figure 5 shows a high dependence in the period 2007-2009 for all the studied pairs. This period corresponds to the recent global financial crisis. In this specific period, the dependence between the Qatar market and other stock markets peaks. In particular, the average dependence exceeds 0.6 between Qatar and the other GCC countries; reaches 0.3 between Qatar and the developed countries; exceeds 0.25 between Qatar and the emerging markets; and reaches 0.634 between Qatar and China. Thus, we provide evidence of an increasing level of market integration for all the markets in our sample due to the global financial crisis. This result is consistent with the existing literature on the impact of the recent global financial crisis on market integration (see, for instance, Frijns et al., 2012; Kumar and Okimoto, 2011; Alotaibi and Mishra, 2017).

This last result highlights the greater effectiveness of time-varying copula modelling. This implies that the static copula approach does not support the detection of hidden information related to the presence of regimes in the dependence between time series due to the impacts of several bad or good events.

Overall, our results show that the dependence between the Qatar market and the other international financial markets is time varying and significantly sensitive to political, economic, and financial crises.

Figure 5: Constant and time varying Student's-*t* copula parameter



24

CONCLUSION

This paper sets out to examine to what extent economic and political events affect stock market dependence. Empirically, we employ static and time varying copula approaches, which prove to be a convenient way to express multivariate distributions and take into consideration movement dependency, especially tail dependency between variables in the case of infrequent risk events. Broadly, the empirical findings report a time varying dependence between the Qatar stock market and the group of selected emerging and developed country stock markets. While previous studies have tried to investigate the impact of bad news on the dependence between financial markets by considering only one break (economic or political bad news), our study provides a more realistic exploration by considering more than one unique event. Indeed, the results of our empirical study show that the time varying behaviour of dependence is due to three major events.

Specifically, the oil price shock significantly affects the dependence structure between the Qatar stock market and the other stock markets. An increase in the international dependence does not significantly impact the local dependence (intra-GCC). The results also show that the oil price shock does not affect the tail dependence parameter for all the studied daily return pairs. In addition, our findings show that the blockade significantly affects the dependence structure in the GCC area. The blockade also affects the levels of dependence between Qatar and the rest of the developed and emerging countries. Finally, the empirical results show that the impact of the global financial crisis on the dependence structure is significantly stronger than the impact of either the oil price crash or the blockade.

In general, our results show that the dependence between the Qatar stock market and the other international financial markets is time varying and significantly sensitive to political, economic, and financial crises. Our results align in part with those of Charles and Darné (2006) and Dungey and Martin (2007), who point to market liberalization and financial crises as the main determinants of instability in the patterns of stock market links.

From a financial perspective, the increasing dependence between the Qatar stock market and other market returns during the post-oil price shock, the post-blockade, and the global financial crisis, corroborates the ‘contagion hypothesis’ during these periods. In recent financial literature (Forbes and Rigobon, 2002), ‘contagion effects’ are perceived as the significant increase in cross-market linkages after a shock to an individual country. Forbes and Rigobon (2002) explain the difference between ‘stock market interdependence’ and ‘the contagion effect’. The ‘contagion effect’ occurs if the stock market exhibits a significant increase in co-movement during the crisis compared to other periods of stability. However, ‘interdependence’ is viewed as a relatively constant correlation between stock markets and the absence of a significant interaction between stock markets and crises. Our results are similar to some previous studies including Garham and Kiviaho (2012), and Kiviaho et al. (2012), and Aloui and Hkiri (2014).

It is also important to note that the increasing level of dependence between stock market returns depends on the type of event considered. In fact, the results show that the global financial crisis has the largest impact on the increase in the dependence between the studied stock markets. However, the dependence between the Qatar stock market and each of the other GCC country markets is more sensitive to the oil price crash. This result suggests there is dissimilarity between the effects of the oil price shock and the blockade at the regional and international integration levels. This difference may be explained by the nature of the two events (economic versus geopolitical). These findings are important for portfolio diversification and asset management, risk management and international asset pricing. Future research should consider how financial dependence relates to financial and political crises, why these events have slowed dependence for some countries and increased it for others, and whether these effects are likely to be long lasting or temporary.

REFERENCES

Abul, S., Nechi, S. and Zhu, H. (2014) 'Journal of Multinational Financial Dependence patterns across Gulf Arab stock markets : A copula approach', *Journal of Multinational Financial Management*. Elsevier B.V., 25–26, pp. 30–50. doi: 10.1016/j.mulfin.2014.06.008.

Alotaibi, A. R. and Mishra, A. V (2017) 'Time varying international financial integration for GCC stock markets', *Quarterly Review of Economics and Finance*. Board of Trustees of the University of Illinois, 63, pp. 66–78. doi: 10.1016/j.qref.2016.03.001.

Aloui, C. and Hkiri, B. (2014) 'Co-movements of GCC emerging stock markets : New evidence from wavelet coherence analysis', *Economic Modelling*. Elsevier B.V., 36, pp. 421–431. doi: 10.1016/j.econmod.2013.09.043.

Arouri, M. E. H., Lahiani, A. and Nguyen, D. K. (2011) 'Return and volatility transmission between world oil prices and stock markets of the GCC countries', *Economic Modelling*. Elsevier B.V., 28(4), pp. 1815–1825. doi: 10.1016/j.econmod.2011.03.012.

Aun, S., Rizvi, R. and Arshad, S. (2017) 'Analysis of the efficiency – integration nexus of Japanese stock market', *Physica A*. Elsevier B.V., 470, pp. 296–308. doi: 10.1016/j.physa.2016.11.142.

Baffes, J. *et al.* (2015) 'The Great Plunge in Oil Prices: Causes, Consequences, and Policy Responses', *Ssrn*. doi: 10.2139/ssrn.2624398.

Balc, M. and Hammoudeh, S. (2015) 'Regional and global spillovers and diversi fi cation opportunities in the GCC equity sectors', 24, pp. 160–187. doi: 10.1016/j.ememar.2015.06.002.

Basher, S. A., Nechi, S. and Zhu, H. (2014) 'Dependence patterns across Gulf Arab stock markets: A copula approach', *Journal of Multinational Financial Management*. Elsevier B.V., 25–26, pp. 30–50. doi: 10.1016/j.mulfin.2014.06.008.

Belkhir, M., Boubakri, N. and Grira, J. (2017) 'Political risk and the cost of capital in the MENA region', *Emerging Markets Review*. Elsevier B.V., 33, pp. 155–172. doi: 10.1016/j.ememar.2017.08.002.

Bensaïda, A. (2015) 'The frequency of regime switching in fi nancial market volatility', *Journal of Empirical Finance*. Elsevier B.V., 32, pp. 63–79. doi: 10.1016/j.jempfin.2015.03.005.

Bouyé, E. and Salmon, M. (2009) *Dynamic copula quantile regressions and tail area dynamic dependence in forex markets*, *European Journal of Finance*. doi: 10.1080/13518470902853491.

Cheung, Y. W. and Ng, L. K. (1996) 'A causality-in-variance test and its application to financial market prices', *Journal of Econometrics*, 72(1–2), pp. 33–48. doi: 10.1016/0304-4076(94)01714-X.

Claus, E. and Lucey, B. M. (2012) 'Research in International Business and Finance Equity market integration in the Asia Pacific region : Evidence from discount factors ☆'. Elsevier B.V., 26, pp. 137–163. doi: 10.1016/j.ribaf.2011.09.002.

Coelho, R. *et al.* (2007) 'The evolution of interdependence in world equity markets — Evidence from minimum spanning trees', 376, pp. 455–466. doi: 10.1016/j.physa.2006.10.045.

Demarta, S. and McNeil, A. J. (2007) 'The t Copula and Related Copulas', *International Statistical Review*, 73(1), pp. 111–129. doi: 10.1111/j.1751-5823.2005.tb00254.x.

Enders, W., Sachsida, A. and Sandler, T. (2002) 'The Impact of Transnational Terrorism on U.S.

Foreign Direct Investment', (1996), pp. 517–531.

Enders, W. and Sandler, T. (1996) 'Terrorism and Foreign Direct Investment in Spain and Greece', 49, pp. 331–352.

Fantazzini, D. (2016) 'The oil price crash in 2014/15: Was there a (negative) financial bubble?', *Energy Policy*. Elsevier, 96(November 2014), pp. 383–396. doi: 10.1016/j.enpol.2016.06.020.

Fayyad, A. and Daly, K. (2011) 'The impact of oil price shocks on stock market returns: Comparing GCC countries with the UK and USA', *Emerging Markets Review*. Elsevier B.V., 12(1), pp. 61–78. doi: 10.1016/j.ememar.2010.12.001.

Fei, F., Fuertes, A. M. and Kalotychou, E. (2017) 'Dependence in credit default swap and equity markets: Dynamic copula with Markov-switching', *International Journal of Forecasting*. Elsevier B.V., 33(3), pp. 662–678. doi: 10.1016/j.ijforecast.2017.01.006.

Frijns, B., Tourani-Rad, A. and Indriawan, I. (2012) 'Political crises and the stock market integration of emerging markets', *Journal of Banking and Finance*. Elsevier B.V., 36(3), pp. 644–653. doi: 10.1016/j.jbankfin.2011.05.007.

Groot, W. De, Pang, J. and Swinkels, L. (2012) 'The cross-section of stock returns in frontier emerging markets ☆ , ☆☆'. Elsevier B.V., 19, pp. 796–818. doi: 10.1016/j.jempfin.2012.08.007.

Hansen, B. E. (1992) 'The likelihood ratio test under nonstandard conditions: Testing the markov switching model of gnp', *Journal of Applied Econometrics*, 7(S1), pp. S61–S82. doi: 10.1002/jae.3950070506.

Hao, J. and He, F. (2018) 'Univariate dependence among sectors in Chinese stock market and systemic risk implication ☆', *Physica A*. Elsevier B.V., 510, pp. 355–364. doi: 10.1016/j.physa.2018.05.142.

Hussain, S. I. and Li, S. (2018) 'The dependence structure between Chinese and other major stock markets using extreme values and copulas', *International Review of Economics and Finance*. Elsevier Inc. doi: 10.1016/j.iref.2017.12.002.

Al Janabi, M. A. M., Hatemi-J, A. and Irandoust, M. (2010) 'An empirical investigation of the informational efficiency of the GCC equity markets: Evidence from bootstrap simulation', *International Review of Financial Analysis*. Elsevier Inc., 19(1), pp. 47–54. doi: 10.1016/j.irfa.2009.11.002.

Joe, H. (2005) 'Asymptotic efficiency of the two-stage estimation method for copula-based models', 94, pp. 401–419. doi: 10.1016/j.jmva.2004.06.003.

Jouini, J. and Harrathi, N. (2014) 'Revisiting the shock and volatility transmissions among GCC stock and oil markets: A further investigation', *Economic Modelling*. Elsevier B.V., 38, pp. 486–494. doi: 10.1016/j.econmod.2014.02.001.

Karanasos, M., Yfanti, S. and Karoglou, M. (2016) 'International Review of Financial Analysis Multivariate FIAPARCH modelling of fi nancial markets with dynamic correlations in times of crisis', *International Review of Financial Analysis*. The Authors, 45, pp. 332–349. doi: 10.1016/j.irfa.2014.09.002.

Karolyi, G. A. and Stulz, R. M. (2003) 'Chapter 16 Are financial assets priced locally or globally?', in *Financial Markets and Asset Pricing*. Elsevier (Handbook of the Economics of Finance), pp. 975–1020. doi: https://doi.org/10.1016/S1574-0102(03)01025-2.

Kolaric, S. and Schiereck, D. (2016) ‘Are stock markets efficient in the face of fear? Evidence from the terrorist attacks in Paris and Brussels’, *Finance Research Letters*. Elsevier Inc., 18, pp. 306–310. doi: 10.1016/j.frl.2016.05.003.

Kollias, C., Kyrtsou, C. and Papadamou, S. (2013) ‘The effects of terrorism and war on the oil price-stock index relationship’, *Energy Economics*. Elsevier B.V., 40, pp. 743–752. doi: 10.1016/j.eneco.2013.09.006.

Kollias, C., Papadamou, S. and Stagiannis, A. (2011) ‘Terrorism and capital markets: The effects of the Madrid and London bomb attacks’, *International Review of Economics and Finance*. Elsevier Inc., 20(4), pp. 532–541. doi: 10.1016/j.iref.2010.09.004.

Kumar, M. S. and Okimoto, T. (2011) ‘Dynamics of international integration of government securities’ markets’, *Journal of Banking and Finance*. Elsevier B.V., 35(1), pp. 142–154. doi: 10.1016/j.jbankfin.2010.07.019.

Longin, F. and Solnik, B. (2001) ‘Extreme Correlation of International’, *The Journal of Finance*, LVI(2), pp. 649–676. doi: 10.1039/F19736900444.

Lucey, B. M. and Muckley, C. (2011) ‘International Review of Financial Analysis Robust global stock market interdependencies’. Elsevier Inc., 20, pp. 215–224. doi: 10.1016/j.irfa.2011.02.001.

Ma, F. *et al.* (2014) ‘Multifractal detrended cross-correlation analysis of the oil-dependent economies: Evidence from the West Texas intermediate crude oil and the GCC stock markets’, *Physica A: Statistical Mechanics and its Applications*. Elsevier B.V., 410, pp. 154–166. doi: 10.1016/j.physa.2014.05.023.

Mensah, J. O. and Alagidede, P. (2017) ‘How are Africa’s emerging stock markets related to advanced markets? Evidence from copulas’, *Economic Modelling*. Elsevier, 60, pp. 1–10. doi: 10.1016/j.econmod.2016.08.022.

Mensi, W., Hammoudeh, S. and Yoon, S. M. (2014) ‘How do OPEC news and structural breaks impact returns and volatility in crude oil markets? Further evidence from a long memory process’, *Energy Economics*. Elsevier B.V., 42, pp. 343–354. doi: 10.1016/j.eneco.2013.11.005.

Mnasri, A. and Nechi, S. (2016) ‘Impact of terrorist attacks on stock market volatility in emerging markets’, *Emerging Markets Review*. Elsevier B.V., 28, pp. 184–202. doi: 10.1016/j.ememar.2016.08.002.

Mohanty, S. K. *et al.* (2011) ‘Oil price movements and stock market returns: Evidence from Gulf Cooperation Council (GCC) countries’, *Global Finance Journal*. Elsevier Inc., 22(1), pp. 42–55. doi: 10.1016/j.gfj.2011.05.004.

Mokni, K. and Mansouri, F. (2017) ‘Conditional dependence between international stock markets: A long memory GARCH-copula model approach’, *Journal of Multinational Financial Management*. Elsevier B.V., 42–43, pp. 116–131. doi: 10.1016/j.mulfin.2017.10.006.

Naifar, N. and Al Dohaiman, M. S. (2013) ‘Nonlinear analysis among crude oil prices, stock markets’ return and macroeconomic variables’, *International Review of Economics and Finance*. Elsevier Inc., 27, pp. 416–431. doi: 10.1016/j.iref.2013.01.001.

Al Nasser, O. M. and Hajilee, M. (2016) ‘Integration of emerging stock markets with global stock markets’, *Research in International Business and Finance*. Elsevier B.V., 36, pp. 1–12.

Ning, C. (2010) ‘Journal of International Money Dependence structure between the equity market and the foreign exchange market – A copula approach’, *Journal of International Money and*

Finance. Elsevier Ltd, 29(5), pp. 743–759. doi: 10.1016/j.jimonfin.2009.12.002.

Nusair, S. A. (2016) 'The effects of oil price shocks on the economies of the Gulf Co-operation Council countries: Nonlinear analysis', *Energy Policy*. Elsevier, 91, pp. 256–267. doi: 10.1016/j.enpol.2016.01.013.

Patton, A. (2006) 'Modeling asymmetric exchange rate dependence', *International Economic Review*, 47(2), pp. 527–556. doi: 10.1111/j.1468-2354.2006.00387.x.

Prest, B. C. (2018) 'Explanations for the 2014 oil price decline: Supply or demand?', *Energy Economics*. Elsevier B.V., 74, pp. 63–75. doi: 10.1016/j.eneco.2018.05.029.

So, M. K. P. and Yeung, C. Y. T. (2014) 'Vine-copula GARCH model with dynamic conditional dependence', *Computational Statistics and Data Analysis*. Elsevier B.V., 76, pp. 655–671. doi: 10.1016/j.csda.2013.08.008.

Yang, L. *et al.* (2015) 'Modeling dependence structures among international stock markets: Evidence from hierarchical Archimedean copulas', *Economic Modelling*. Elsevier B.V., 51, pp. 308–314. doi: 10.1016/j.econmod.2015.08.017.

DECISION SCIENCES INSTITUTE

Maharatna PSUs and Financial Soundness - An Application of Altman's Z Score Model

Shyamsunder Chitta
Symbiosis Institute of Business Management, Hyderabad
Symbiosis International (Deemed University)
Email: shyam.chitta@sibmhyd.edu.in

Ravi Kumar Jain
Symbiosis Institute of Business Management, Hyderabad
Symbiosis International (Deemed University)
Email: director@sibmhyd.edu.in

ABSTRACT

Maharatna companies are top most PSUs contributing greatly to the development of our nation. These companies have existed for a long time and became self-sustained without depending on government for funds. In recent past some of the companies have started reporting losses continuously and are on the verge of bankruptcy. The government has been infusing funds to safeguard them, but is not able to improve their performance to the desired level. This paper attempts to study the reasons for the current situation of these companies by applying the famous Altman Z-Score model to verify whether these companies really suffer financially. The sample for the study comprises of Maharatna companies, and the results reveal that not all the Maharatna companies are performing as expected.

KEYWORDS: Altman Z-Score, Maharatna, Navaratna, PSUs, Bankruptcy

INTRODUCTION

In the present world of globalization, the business environment is very much turbulent, and the manner in which the business operates is changing over time. With increasing competition and rising costs, many companies are on the verge of bankruptcy or becoming bankrupt. In light of this, it is very much important to measure the financial soundness of the enterprises, so as to safeguard the interest of the stakeholders associated with them. We can estimate the financial soundness of the companies by establishing the relationship between various financial variables. There are various studies, which attempted to study the nature of bankruptcies in various industries, and one of the most prominent among them was done by E.I. Altman, who developed the model named after him as the Altman’s Z score model.

The Altman Z-score is the output of a credit-strength test that gauges a publicly traded manufacturing company's likelihood of bankruptcy. The Altman’s Z-score is based on five financial ratios calculated from data found in a company's annual report. It uses profitability, leverage, liquidity, solvency and activity to predict whether a company has high probability of becoming Insolvent.

Maharatna companies are the top most PSUs contributing greatly to the development of our nation. These companies have existed for a long time and became self-sustained without depending on government for funds. In recent past some of them started reporting losses continuously and are on the verge of bankruptcy. The government has been infusing funds to safeguard them, but is not able to improve their performance to the desired level.

OBJECTIVES OF THE PAPER

1. To investigate the financial performance of Maharatna Companies with respect to selected ratios.
2. To study the relevance of ratios used in Altman's Z score model to predict the probability of a company becoming bankrupt.
3. To predict the possibility of bankruptcy of Maharatna Companies using Altman's Z score model.

METHODOLOGY

Collection of data: The study is based on secondary data analysis, and the data was collected from Bloomberg database. The supporting data was collected from various other sources.

Tools used: The Altman's Z score model was used to predict the bankruptcy of the companies considered for the present study.

Altman's Z score model: The model is one of the most important models used for predicting the possibility of bankruptcy of the companies. The Z-score is a set of financial ratios in a multivariate context, based on multiple discriminant model, which is a statistical technique used to classify and/or make predictions in problems where the dependent variable appears in qualitative form (Altman, 1968).

The formula for the standard Z-Score in case of manufacturing firms is:

Z-Score = $1.2T_1 + 1.4T_2 + 3.3T_3 + 0.6T_4 + 0.999T_5$

T_1 = Working Capital / Total Assets
T_2 = Retained Earnings / Total Assets
T_3 = Operating Earnings / Total Assets
T_4 = Market Capitalization / Total Liabilities
T_5 = Sales / Total Assets

A score below 1.8 means it's likely the company is headed for bankruptcy, while companies with scores above 2.99 are not likely to go bankrupt. Investors can use Altman Z-scores to determine whether they should buy or sell a stock if they're concerned about the company's underlying financial strength. Altman's Z-Score determines how likely a company is to fail, by evaluating seven simple pieces of data, all of which should be available in the company's public disclosure.

Altman originally developed the Z-Score for manufacturing firms primarily because those were the companies in his original sample. However, the emergence of large, public service companies prompted him to develop a second Z-Score model for non-manufacturing companies. The formula is essentially the same as before; it just excludes the last component (sales / total assets) because Altman wanted to minimize the effects of manufacturing-intensive asset turnover.

The formula for the Z-Score in case of service firms is:
Z-Score = $1.2T_1 + 1.4T_2 + 3.3T_3 + 0.6T_4$

Interpretation of the variables

1. The first ratio (working capital / total assets) is an indicator of a firm's ability to make good on what it owes in the next few months.

2. The second ratio is a good indicator of how in debt the company is and whether it has a history of profitability.
3. The third ratio is a measure of efficiency in that it indicates how much the company generates in earnings by efficiently using its assets.
4. The fourth ratio is a fluid measure of the market's confidence in the company.
5. The fifth ratio is similar to the third ratio in that it measures the company's efficiency in generating sales from its assets.

Altman Z Score			
Weight	**Factor**	**Ratio**	**Significance**
1.2	A	Working Capital / Total Assets	It measures the net liquid asset of a company relative to the total assets.
1.4	B	Retained Earnings / Total Assets	It measures the financial leverage level of a company.
3.3	C	EBIT / Total Assets	It measures productivity of a company’s total assets.
0.6	D	Equity / Total Liabilities	It measures what portion of a company’s assets can decline in value before the liabilities exceed the assets.
0.999	E	Sales / Total Assets	It measures revenue-generating ability of a company’s assets.

Altman’s Z Score	Zone	Interpretation
Z>2.99	Safe Zone	There is low probability of bankruptcy and the company is healthy
1.81<Z<2.99	Grey Zone	These firms are considered as cases which should be watched with attention
Z<1.81	Distress Zone	There is high probability of bankruptcy

The Z-Score is a commonly used metric with wide appeal, though it is just one of many credit scoring models in use today, that essentially combine quantifiable financial indicators with a small number of variables in an attempt to predict whether a firm will fail. Over time, however, the Z-Score has proved to be one of the most reliable predictors of bankruptcy, so much so that analysts often equate certain Z-Scores with corresponding bond ratings. When Altman re-evaluated his methods by examining 86 distressed companies from 1969 to 1975 and then 110 bankrupt companies from 1976 to 1995 and later 120 bankrupt companies from 1996 to 1999, the Z-Score gave accurate results between 82% and 94%. However, if the company’s financials are misleading or incorrect, the Z-Score will be, too.

MAHARATNA COMPANIES

A state-owned enterprise in India is called a public sector undertaking (PSU) or a public sector enterprise, owned by the union government of India, or one of the many state or territorial governments, or both. PSUs are further classified as central public sector enterprises (CPSEs) and state level public enterprises (SLPEs). In 1951, there were just five enterprises in the public sector in India, which increased to 246 by March 1991. CPSEs are companies in which the direct holding of the Central Government or other CPSEs is 51% or more. The Ministry of Heavy Industries and Public Enterprises administer them. Certain public sector undertakings have been awarded additional financial autonomy, which have comparative advantages, giving them greater autonomy to compete in the global market to support them in their drive to become global giants.

Financial autonomy was initially awarded to nine PSUs and given Navaratna status in 1997. In 2010, the government established the higher Maharatna category, which raises a company's investment ceiling from Rs. 1,000 crore to Rs. 5,000 crore. The Maharatna firms can decide on investments of up to 15 per cent of their net worth in a project while the Navaratna companies could invest up to Rs 1,000 crore without explicit government approval.

The country has witnessed the introduction of a new industrial policy in the 1990s that emphasized delicensing, greater independence for profitable PSUs and restructuring of loss-making firms through the Bureau of Industrial Financing and Restructuring (BIFR). Other elements of the liberalization involved free entry to private sector firms in industries reserved exclusively for PSUs, and disinvestment of a small part of the government's shareholding (while still holding majority stocks) and listing PSUs on the stock exchanges.

Eligibility Criteria for grant of Maharatna status
CPSEs fulfilling the following criteria are eligible to be considered for grant of Maharatna status:

1. Having Navratna status
2. Listed on the Indian stock exchange, with a minimum prescribed public shareholding under SEBI regulations
3. An average annual turnover of more than Rs. 20,000 crore during the last three years
4. An average annual net worth of more than Rs.10,000 crore during the last three years
5. An average annual net profit of more than Rs. 2,500 crore during the last 3 years
6. Significant global presence or international operations

<u>The list of Maharatna CPSEs is given below</u>
1. Bharat Heavy Electricals Limited (BHEL)
2. Coal India Limited (CIL)
3. Gas authority of India Limited (GAIL)
4. Indian Oil Corporation Limited (IOCL)
5. National Thermal Power Corporation Limited (NTPC Limited)
6. Oil & Natural Gas Corporation Limited (ONGC)
7. Steel Authority of India Limited (SAIL)
8. Bharat Petroleum Corporation Limited (BPCL)

REVIEW OF LITERATURE

Altman (1968) tried to assess the quality of ratio analysis as an analytical technique. He took a sample of sixty-six firms in two mutually exclusive groups, namely, thirty-three bankrupt firms and thirty-three non-bankrupt firms, and developed the multiple discriminant analysis for manufacturing corporations. The author explained that since ratio analysis is susceptible to faulty interpretation and can be analyzed subjectively, there is a need to develop a model that suitably removes the possible ambiguities observed in traditional studies. He developed a model known as the Altman's Z score model, which has been used extensively in various research studies.

Altman, Haldeman, and Narayanan (1977) developed a new model to identify bankruptcy risks of the corporations, named as the Zeta model, which considered the developments related to recent business failures. The new model was effective in classifying bankrupt companies up to 5 years prior to failure on a sample of corporations consisting of manufacturers and retailers.

Altman (2002) reviewed two important credit scoring techniques, the Z-score (developed by the author himself in 1968) and KMV's Expected Default Frequency (EDF) model with respect to default probabilities in general. He explained that the relevance of these credit scoring models are in accordance to the changes brought about by the Basel II norms and the increasing bankruptcy filings by large corporates.

Samarakoon and Hasan (2003) investigated the ability of three versions of Altman's Z score model (Z, Z', and Z”) to predict bankruptcy in listed companies in Sri Lanka, and the results revealed that these models have a very high degree of accuracy in predicting corporate distress.

Jayadev (2006) provided empirical evidence on the acceptability of the Z-score model. The author used three forms of the Z score model and estimated the coefficients in all the 3 equations by using a sample of 112 companies. The results of the study revealed that Altman's model was capable of predicting default possibilities in most of the sample companies.

Gerantonis, Vergos, and Christopoulos (2009) took a sample of all listed companies on the Athens exchange for a period of 2002-2008 and concluded that Altman's Z score model can predict failures.

Hayes, Hodge, and Hughes (2010) attempted to construct and interpret Z score and apply it to the retail industry in the study period during 2007 and 2008. Their sample constituted the public retail companies which had assets greater than $1,000,000 and had declared bankruptcy during 2007-2008. The results of the period under study accurately classified eight out of nine firms under investigation as bankrupt.

Ramaratnam and Jayaraman (2011) used Altman's Z score model to predict, analyze, and compare the financial health of the major steel companies in India during 2006-2010. The Z score was further tested using the ANOVA to analyze the consistency and stability on the different ratios used in Altman's Z score.

Raiyani and Bhatasna (2011) studied the financial health of the textile industry in India by considering four major players during 2002-2009 and applied the Altman’s Z Score model to these companies to predict, analyze, and compare the financial health of these companies.

Ray (2011) made an attempt to apply Altman's Z Score model on the Indian automobile industry and tested whether the model can accurately predict the cases of bankruptcy in the Indian automobile industry for the study period from 2003-04 to 2009-10. The author considered a sample of 62 publicly traded companies listed on the Bombay Stock Exchange.

Reddy (2012) made an attempt to study the association between liquidity, profitability, and risk factors by employing Altman's Z Score model on Dr. Reddy's Laboratories Ltd. during the time period 2005 -2011, and the study revealed that the company was not suffering from financial distress.

Ahuja and Singhal (2014) have applied Altman's Z score model to investigate the financial position of the companies operating I the Indian textile industry, with a sample of fifteen companies and concluded that the model can be used by banks to predict the signs of bankruptcy of firms and devise various strategies to overcome it.

ANALYSIS AND RESULTS OF ALTMAN'S Z SCORE

Table-1(a): Altman's Z score of Maharatna Companies

Name of the Company	2018	2017	2016	2015	2014	Mean
CIL	2.423	3.656	3.612	3.439	3.246	3.275
BPCL	2.757	2.918	3.65	3.468	3.383	3.235
GAIL	2.119	2.161	2.156	1.947	2.053	2.087
IOCL	2.214	2.324	2.038	1.858	1.88	2.063
ONGC	0.614	0.876	0.91	0.941	1.071	0.882
SAIL	0.791	0.533	0.572	1.118	1.285	0.86
NTPC	0.652	0.676	0.733	0.859	1.05	0.794
BHEL	0.493	0.537	0.48	0.466	0.619	0.519
Mean	1.508	1.71	1.769	1.762	1.823	1.714

Table-1(b): Altman's Z score Model of the companies

Z Score	Name of the Company
Z>2.99	Coal India Limited, Bharat Petroleum Corporation Limited
1.81<Z<2.99	Gas Authority of India Limited, Indian Oil Corporation Limited
Z<1.81	Oil and Natural Gas Corporation, Steel Authority of India Limited, National Thermal Power Corporation, Bharat Heavy Electricals Limited

Analysis and Results
The data of eight Maharatna companies are analyzed and the Z score was obtained for the period 2014-2018 using the Altman's Z score model. Table-1(b) shows the consolidated results of the Z score using Altman's Z score model. Table-1(a) shows the mean Z score value of the companies during the last five years. As per the mean score, only two companies, namely Coal India Limited, Bharat Petroleum Corporation Limited had a score of Z>2.99, which suggests that these companies are financially safe. However, the companies, namely Oil and Natural Gas Corporation, Steel Authority of India Limited, National Thermal Power Corporation, Bharat Heavy Electricals Limited had an average Z score of less than 1.81, which indicates that these companies are not performing well and are in the bankruptcy zone.

Detailed analysis of various variables of Altman's Z score

From Table-1(a), we can observe that Coal India Limited has the highest average Z score of 3.275, which is due to the market capitalization, which was the highest during 2017, but has decreased during 2018. Also, its EBIT is higher compared to other companies, but is showing a declining trend during 2018. Bharat Petroleum Corporation Limited has the second highest Z score of 3.235, which is due to its EBIT and Sales.

Tables 2 to 6 provides the detailed analysis of the variables used for Altman's Z score model. Table-2 shows the working capital to total assets ratio, which reflects the operational efficiency of the company. The ratio of all the companies have shown a declining trend, most of them being negative, which indicates the company's inability to meet its current obligations.

Table-3 provides the ratio of retained earnings to total assets, which is an indicator of the extent to which the assets of the company has been financed by retained earnings. The results indicate that for most of the companies, the ratio is showing an increasing trend, which is good for the companies. However, for BHEL, the ratio is negative, which indicates that the company has borrowed debt in excess of retained earnings to finance its assets.

Table-4 shows the analysis of EBIT to total assets, which indicates how efficiently a company is using its assets to generate earnings before interest and taxes. The results indicate that the ratio for all the companies is showing a decreasing trend, which is an alarming situation, and needs to be taken care of.

Table-5 indicates the detailed analysis of equity to total liabilities, which indicates whether the company has sufficient net worth to meet the total debt obligations. This ratio is directly related to the market capitalization of the companies, which in turn depend upon the trends in the stock market, which are affected by the macro economic factors. The results reveal that the ratio for most of the companies shows an increasing trend, which is a good sign, but for few of them, it has decreased, which needs to be analyzed in light of the macro economic factors.

Table-6 shows the analysis of sales to total assets, which measures the effectiveness with which the firm uses its total assets to generate sales revenue. Except for SAIL and CIL, the ratio for other companies has shown a decreasing trend. It indicates that these companies are not able to utilize their assets to generate sales revenue.

Table-2: Working Capital / Total Assets

Name of the Company	2018	2017	2016	2015	2014	Mean
BHEL	0.328	0.372	0.374	0.385	0.364	0.365
CIL	0.017	0.123	0.206	0.384	0.372	0.220
ONGC	-0.096	0.043	0.058	0.053	0.054	0.022
GAIL	0.004	0.014	-0.002	0.011	0.036	0.012

NTPC	-0.028	-0.042	-0.020	0.035	0.081	0.005
BPCL	-0.079	-0.100	-0.048	-0.034	0.015	-0.049
IOCL	-0.117	-0.138	-0.043	-0.004	-0.003	-0.061
SAIL	-0.121	-0.195	-0.145	-0.060	-0.016	-0.107
Mean	-0.011	0.010	0.048	0.096	0.113	0.051

Table-3: Retained Earnings / Total Assets						
Name of the Company	2018	2017	2016	2015	2014	Mean
GAIL	0.502	0.489	0.472	0.452	0.449	0.473
SAIL	0.213	0.236	0.287	0.339	0.359	0.287
IOCL	0.306	0.302	0.038	0.017	0.028	0.138
CIL	0.083	0.165	0.056	0.136	0.200	0.128
BPCL	0.078	0.041	0.075	0.057	0.046	0.059
ONGC	0.050	0.044	0.050	0.046	0.070	0.052
NTPC	0.033	0.038	0.050	0.045	0.038	0.041
BHEL	-0.002	-0.007	-0.009	0.033	0.052	0.013
Mean	0.158	0.163	0.128	0.140	0.155	0.149

Table-4: EBIT / Total Assets						
Name of the Company	2018	2017	2016	2015	2014	Mean
CIL	0.490	0.761	0.778	0.627	0.626	0.657
BPCL	0.112	0.122	0.147	0.158	0.152	0.138
GAIL	0.120	0.126	0.131	0.102	0.109	0.118
ONGC	0.099	0.117	0.114	0.113	0.135	0.116
IOCL	0.116	0.126	0.119	0.077	0.032	0.094
NTPC	0.063	0.065	0.064	0.066	0.091	0.070
BHEL	0.025	0.026	0.010	-0.017	0.030	0.015

SAIL	0.020	-0.025	-0.048	0.039	0.044	0.006
Mean	0.131	0.165	0.164	0.146	0.152	0.152

Table-5: Equity / Total Liabilities

Name of the Company	2018	2017	2016	2015	2014	Mean
CIL	1.076	1.249	1.184	1.166	0.735	1.082
NTPC	0.185	0.213	0.242	0.270	0.326	0.247
SAIL	0.095	0.089	0.107	0.120	0.146	0.111
GAIL	0.127	0.098	0.057	0.053	0.056	0.078
ONGC	0.066	0.104	0.076	0.067	0.068	0.076
IOCL	0.056	0.030	0.018	0.016	0.013	0.026
BPCL	0.030	0.021	0.015	0.015	0.014	0.019
BHEL	0.024	0.017	0.015	0.014	0.012	0.016
Mean	0.207	0.228	0.214	0.215	0.171	0.207

Table-6: Sales / Total Assets

Name of the Company	2018	2017	2016	2015	2014	Mean
BPCL	2.358	2.569	3.110	2.900	2.792	2.746
IOCL	1.510	1.636	1.632	1.579	1.733	1.618
GAIL	0.941	0.988	1.031	0.933	0.990	0.976
SAIL	0.516	0.467	0.437	0.515	0.570	0.501
ONGC	0.292	0.314	0.349	0.398	0.421	0.355
NTPC	0.321	0.331	0.329	0.375	0.405	0.352
CIL	0.022	0.016	0.008	0.019	0.013	0.016
BHEL	0.006	0.005	0.003	0.006	0.004	0.005
Mean	0.746	0.791	0.863	0.841	0.866	0.821

CONCLUSION

Sound financial health is an important requirement for the survival and growth of a business. Since crucial business decisions are taken keeping in mind the financial capability of the firm, it is important to select and use optimal tools to analyze and predict financial strength of the firms. Altman's Z Score is one of the effective models which helps in judging the financial position of the firm and predicting bankruptcy.

A bolder roadmap for gradually getting the government out of the business must be prepared with a hard look at the real economic benefits from some of the profit-making state-owned firms as well. For now, India could leave the Mahartana's which hold about one-third of total assets of all PSUs in state hands, but with a plan to make them world class companies.

REFERENCES

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23* (4), 589-609.

Altman, E. I. (2002). Corporate distress prediction models in a turbulent economic and Basel II environment. Retrieved from http://pages.stern.nyu.edu/~ealtman/Corp-Distress.pdf

Altman, E.I., Haldeman R. G., & Narayanan, P. (1977). ZETATM* analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance, 1*, 29-54.

Bhavna Ranjan Ahuja, Nitish Singhal (2014). Assessing the Financial Soundness of Companies with Special Reference to the Indian Textile Sector: An Application of the Altman Z Score Model. *Indian Journal of Finance. 8(4), 38-48.*

Gerantonis, N., Vergos, K., & Christopoulos, A. G.(2009). Can Altman Z-score models predict business failures in Greece? *Research Journal of International Studies* , Issue 12, 21-28.

Hayes, S.K., Hodge, K.A., & Hughes, L.W. (2010). A study of the efficacy of Altman's Z to predict bankruptcy of specialty retail firms doing business in contemporary times. *Economics & Business Journal: Inquiries & Perspectives, 3*(1), 130-134.

http://bsepsu.com/maharatnas.asp

http://gailonline.com/home.html

http://www.bhel.com/

https://currentaffairs.gktoday.in/tags/maharatna-companies

https://investinganswers.com/financial-dictionary/financial-statement-analysis/altman-z-score-5188

https://strategiccfo.com/z-score-model

https://www.accountingtools.com/articles/the-altman-z-score-formula.html

https://www.bharatpetroleum.com/

https://www.coalindia.in/en-us/home.aspx

https://www.iocl.com/

https://www.jagranjosh.com/general-knowledge/list-of-maharatna-and-navratna-companies-in-india-1467721897-1

https://www.moneycontrol.com/

https://www.ntpc.co.in/

https://www.ongcindia.com/wps/wcm/connect/en/home/

https://www.sail.co.in/

https://www.wikipedia.org/

Jayadev, M. (2006). Predictive power of financial risk factors: An empirical analysis of default companies, *Vikalpa, 31* (3), 45-56.

Raiyani, J.R., & Bhatsana, R.B. (2011). A study on financial health of textile industry in India : A 'Z'- score approach. *Indian Journal of Finance, 5* (1), 9-16, 25.

Ramaratnam, M. S., & Jayaraman, R. (2011). A study on measuring the financial soundness of select firms with special reference to Indian steel industry An empirical view with Z score. *Asian Journal of Management Research, 1* (2), 724- 735.

Ray, S. (2011). Assessing corporate financial distress in automobile industry of India: An application of Altman's model. *Research Journal of Finance and Accounting, 2* (3), 155 -168.

Reddy, V. C. (2012). Analysis of liquidity, profitability, risk and financial distress: A case study of Dr. Reddy's Laboratories Ltd., *Indian Journal of Finance, 6* (12), 5 - 17.

Samarakoon, L. P., & Hasan, T. (2003). Altman's Z-Score models of predicting corporate distress: Evidence from the emerging Sri Lankan stock market. *Journal of the Academy of Finance, 1*, 119-125.

DECISION SCIENCES INSTITUTE

A Numerical Taxonomy of Stroke Patient Medical History: Implications for Operations Management

Brandon Lee
University of Dayton
Email: blee03@udayton.edu

Aleda Roth
Clemson University
Email: aroth@clemson.edu

Lawrence Fredendall
Clemson University
Email: flawren@clemson.edu

ABSTRACT

This study describes a numerical taxonomy of stroke patients, which was developed with *k-means* cluster analysis methods, using 22 medical history variables. The patient sample comes from a stroke dataset of a tertiary referral hospital that houses an award-winning comprehensive stroke center. The data contain the patient-level data of 8,531 patients for a time period from January 2009 to March 2017. We provide insights for the operations management issues of stroke patient care processes and discuss the necessity to develop specific operating procedures and resource management policies for different patient groups.

KEYWORDS: Taxonomy, Patient Classification, Healthcare Operations, Stroke

INTRODUCTION

Organizations attempt to reduce process variations by standardizing their processes, which in turn improves the quality of products or services provided to their customers. However, standardizing hospital processes has proven to be challenging. One reason for this is due to inherently heterogeneous patient populations (Tucker et al., 2007). This study aims to provide an in-depth understanding of a patient population from a set of medical history variables involving a stroke patient sample by developing a numerical taxonomy. A numerical taxonomy is an effective method to determine distinct groups within which common characteristics are presented (Miller and Roth, 1994). As such, a numerical taxonomy using cluster analysis methods is useful in classifying patient populations.

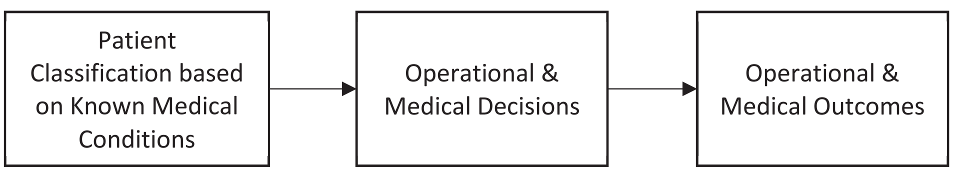
We reviewed two streams of literature on patient classification: i) the medical literature and ii) the operations management literature. First, in the medical research community, cluster analysis methods have been widely used to classify patients and compare patient clusters on medical outcomes. For example, Vanfleteren et al. (2013) use a self-organizing map cluster analysis method to classify chronic obstructive pulmonary disease (COPD) patients. They find 5 distinct patient groups based on the 13 comorbidities and conduct a further comparison analysis among patient groups on the patients' medical conditions. Also, Vu et al. (2011) use an agglomerative hierarchical clustering method to classify fall injury patients. They find 5 distinct patient groups based on the 22 comorbidities. Each study addresses a specific morbidity, and its related comorbidities are different from those of other studies based on different morbidities.

As such, each patient classification study generates a unique set of patient clusters. Studies on patient classification in the medical literature do not address how patient classification triggers medical or operational decisions, which in turn affects patient care outcomes. Rather, they heavily focus on medical outcomes or physician conditions among the different patient groups.

Second, patient classification in the operations management literature has been primarily applied to appointment scheduling (AS) methods. These studies utilize patient classification for sequencing and determining appointment intervals to examine the efficiency of the scheduling system based on outcome variables such as patient wait time, physician idle time, and overtime. Patient classification in this stream of literature could be as simple as new-return patients (Cayirli et al., 2008), outpatient-inpatient-emergency patients (Kolisch & Sickinger, 2008), or based on more patient characteristics such as self-reported health conditions, medication, etc. (Salzarulo et al., 2016). Although these studies for AS show how different patient classes or groups lead to different operational decisions, the use of medical information (e.g., comorbidities) is substantially lacking. Research that addresses both the link between patient classification and operational or medical decisions, and patient classification based on patients' medical conditions are scant. Hu et al. (2018) present a similar notion to what we are trying to achieve. Instead of patient classification, they use patient risk scores based on vital signs, lab results, diagnoses, and patient demographics to make informed decisions on the transfer of patients to the intensive care unit (ICU).

We attempt to answer the following research question with this study: Does a numerical taxonomy for patient classification trigger more informed medical and operational decisions for stroke patient care? Figure 1 describes the conceptual framework of how patient classification could help improve operational and medical outcomes. We first present how we classify patient groups using a numerical taxonomy and then discuss how such a classification could lead to informed medical and operational decisions for better outcomes.

Figure 1: Conceptual Framework



MEDICAL CONTEXT OF STROKE

Stroke is one of the leading causes of deaths, accounting for 5.2% of the total number of deaths in the United States, according to the 2015 National Center for Health Statistics (2017). Stroke care is highly time dependent because a typical stroke patient loses 1.9 million neurons every minute (Saver, 2006). Due to its severity and prevalence, stroke care has received national attention. The American Heart Association initiated "*Get With The Guidelines-Stroke (GWTG-Stroke)*" in order to provide hospitals with standardized patient data collection methods so that participating hospitals could measure their performance and compare their own performance with that of other hospitals. The ultimate goal of *GWTG-Stroke* is to improve stroke care processes, which essentially means faster treatment delivery to patients. Of the key performance measures from *GWTG-Stroke*, Door-to-Needle (DTN) time is considered one of the most important stroke care performance measures. The DTN time refers to the time elapsed between the arrival of a stroke patient at the hospital and the infusion of a thrombolytic medication called TPA (Tissue Plasminogen Activator) to the patient. The national target for the

DTN time recommended by the *Target:Stroke* initiative (which is part of *GWTG-Stroke*) is 60 minutes. Rewards (e.g., *Target:Stroke* Honor Roll) are given to participating hospitals who meet the target. *Target:Stroke* also provides 10 best practices for hospitals to follow in order to achieve or exceed the national target for the DTN time. Thus far, however, the *GWTG-Stroke* and *Target:Stroke* initiatives have not addressed the necessity of prompt gathering or access to patient medical history information, nor have they addressed information transfer from emergency medical service (EMS) providers to emergency department (ED) clinicians. Also, the national target and rewards system of *Target:Stroke* are not based on different patient medical conditions that hospitals receive.

METHODS

Research Setting and Data Collection

This study is based on a patient-level stroke dataset from a tertiary referral hospital that houses a comprehensive stroke center (CSC), the highest level of certification based on patient care capability. The hospital has over 800 certified beds, was certified as a primary stroke center (PSC) in 2009, and has been certified as a CSC since November 2017 by the *Joint Commission*. This hospital has actively participated in *Get With the Guidelines-Stroke (GWTG-Stroke)* since 2006 and the *Target:Stroke* initiative since 2011. Both programs are governed nationally by the American Heart Association. The hospital has won performance-based awards in stroke care (i.e., *GWTG-Stroke Gold-Plus* and *Target:Stroke Honor Roll Elite Plus*), and has been recognized by the American Heart Association.
The stroke dataset includes 8,531 patient visit episodes from January 2009 to March 2017. The dataset includes patient-level variables, such as patient demographic information, patient medical history, stroke type, arrival mode, time stamps of process milestones, insurance information, doctor information, etc. Patient demographics are summarized in Table 1. Approximately 45% of the patients in this sample were 70 years of age or older. There were about 3% more female patients than male patients. 78% of the patients were white, 21% were black, and less than 2% of the patients were Hispanic.

Table 1: Descriptive Statistics of Patient Demographics

Characteristics	**Sample (Total=8531)**	**Proportion**
Age Group		
< 50 years old (1)	1127	13.2%
50-59 years old (2)	1603	18.8%
60-69 years old (3)	1993	23.4%
70-79 years old (4)	1907	22.4%
80+ years old (5)	1901	22.3%
Gender		
Male (0)	4134	48.5%
Female (1)	4397	51.5%
Race		
White (0)	6635	77.8%
Black (1)	1791	21.0%
Asian (2)	52	0.6%
Native (3)	5	0.1%
Not Documented (4)	48	0.6%
Ethnicity		
Non-Hispanic (0)	8387	98.3%
Hispanic (1)	144	1.7%

Taxon Variables

In order to identify different groups of patients based on their medical conditions, we chose medical history variables as taxon variables, which were available from the stroke dataset. Such variables are standard patient medical history variables from *GWTG-Stroke* by the American Heart Association and include 22 comorbidity and stroke risk factor variables (American Heart Association, 2018 - *Get With The Guidelines®-Stroke Overview*). All of them are binary variables (= 1 for a condition present, = 0 for a condition not present). Table 2 describes the medical history variables. Approximately 4% of the patients in the sample did not have any of the medical conditions.

Table 2: Descriptions of Taxon Variables

Medical History Variable (Taxon)	Percent of Total Patients
Atrial Fibrillation or Atrial Flutter	15.4%
Coronary Artery Disease / Prior Myocardial Infarction	29.3%
Carotid Stenosis	5.5%
Current pregnancy	0.1%
Depression	11.4%
Diabetes Mellitus	33.2%
Drugs/Alcohol Abuse	5.8%
Dyslipidemia	48.4%
Family History of Stroke	7.8%
Heart Failure	10.2%
Hormone Replacement Therapy	1.3%
Hypertension	77.1%
Migraine	2.3%
Obesity/Overweight	37.4%
Previous Stroke	24.9%
Previous TIA	8.9%
Prosthetic Heart Valve	1.1%
Peripheral Vascular Disease	7.1%
Renal Insufficiency - Chronic	7.2%
Sickle Cell	0.1%
Sleep Apnea	2.8%
Smoker	26.8%

Cluster Analysis Procedures

We employed a cluster analysis to identify patient groups based on the medical history variables. The SAS ACECLUS (Approximate Covariance Estimation for CLUStering) procedure and the FASTCLUS procedure were used (Miller and Roth, 1994; Rosenzweig et al., 2011; and Lee et al., 2015). The ACECLUS procedure enabled us to "obtain approximate estimates of the pooled within-cluster covariance matrix," using the 22 medical history variables (SAS/STAT® 13.2 User's Guide: The ACECLUS Procedure 2018). The FASTCLUS procedure was subsequently employed to conduct a disjoint cluster analysis. According to the SAS/STAT® 13.2 User's Guide: The FASTCLUS Procedure (2018), “the FASTCLUS procedure uses Euclidean distances, so the cluster centers are based on least squares estimation. This kind of clustering method is often called a *k-means* model, since the cluster centers are the means of the observations assigned to each cluster when the algorithm is run to complete convergence. Each iteration reduces the least squares criterion until convergence is achieved.” Since all of the 22 medical history variables are coded in 0/1 binary codes and the units are consistent across all variables, standardization of the variables is not necessary.

To ascertain a suitable number of clusters, we contemplated using the R^2 and pseudo-F statistic. These statistics show notable increases with respect to the homogeneous nature of the clusters and the practical interpretability of the patient clusters based on ANOVA, Chi-square test, and Tukey (for continuous variables) or Dunn (for binary variables) pairwise comparison tests involving cluster mean or proportion differences. The four-cluster model best fulfills these two criteria. According to Wilk’s Lambda criterion and the F-statistic, the null hypothesis (four clusters are equal across all taxon variables) may be rejected (F-statistic= 562.95, $p <.0001$).

RESULTS

In this section, we report on the stroke patient groups based on patients’ prior medical history before arrival at the hospital for the treatment of stroke symptoms. We also report the differences among the patient groups on demographic, time performance, and other process-related variables.

Patient Groups based on Medical History

The four resultant patient groups are described in Table 3 in cluster centroid (mean) scores. Since we use binary codes, the scores essentially represent the percentage of patients in a cluster who had a particular medical history prior to the occurrence of stroke symptoms. We labeled these four clusters of patients as “Metabolic,” “Cardiovascular,” “Behavior,” and “Low Risk.”

Cluster 1: Metabolic. We label cluster 1 as “metabolic” because the patients in this group had a relatively higher likelihood of having conditions related to metabolic syndrome: diabetes, dyslipidemia, hypertension, and obesity (American Heart Association, 2018 - What is Metabolic Syndrome?). The 2,148 patients of cluster 1 represent 25 percent of the entire stroke patients in our sample.

Cluster 2: Cardiovascular. We label cluster 2 as “cardiovascular” because the patients in this group had a relatively higher likelihood of having cardiovascular diseases, such as atrial fibrillation or flutter, coronary artery disease, carotid stenosis, heart failure, previous stroke, previous TIA, a prosthetic heart valve, peripheral vascular disease, and renal insufficiency (American Heart Association, 2018 - What is Cardiovascular Disease?). Although patients in this group did not have the highest cluster means for previous stroke and renal insufficiency, they had the highest cluster means on all other cardiovascular disease histories. The 1,702 patients of cluster 2 represent 18 percent of the entire stroke patients in our sample.

Cluster 3: Behavior. Because this group of patients had the highest cluster means on both of the behavioral risk factors - drug and alcohol abuse and smoking among the four clusters, this group is labeled as “behavior.” The 2,001 patients of cluster 3 represent 24 percent of the entire stroke patients in our sample.

Cluster 4: Low Risk. The patients in this group had the highest cluster means on only two out of the 22 medical history variables (i.e., pregnancy and hormone replacement therapy) among the four clusters. We label cluster 4 as “low risk” patients. For all other medical history variables, this group had relatively low cluster means among the four clusters. The 2,680 patients of cluster 4 represent 31 percent of the entire stroke patients in our sample.

Table 3: Cluster Results - Patient Groups by Medical History

Medical History Variable	**1. Metabolic (n=2148)**	**2. Cardio-vascular (n=1702)**	**3. Behavior (n=2001)**	**4. Low Risk (n=2680)**	**Chi-square Statistic (p-value)**
Atrial Fibrillation / Flutter					
Cluster Mean	0.122	**0.422**	0.037	0.098	1226.24
Different Groups (Pairwise)	(2, 3)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.327	0.494	0.189	0.298	
Coronary Artery Disease					
Cluster Mean	0.314	**0.849**	0.125	0.047	3600.84
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.464	0.358	0.331	0.212	
Carotid Stenosis					
Cluster Mean	0.062	**0.100**	0.066	0.014	159.11
Different Groups (Pairwise)	(2, 4)	(1, 3, 4)	(2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.242	0.300	0.248	0.117	
Current Pregnancy					
Cluster Mean	0.000	0.001	0.000	**0.004**	15.32
Different Groups (Pairwise)	(4)	(4)	(4)	(1, 2, 3)	(p=0.002)
Std. Dev.	0.000	0.024	0.022	0.061	
Depression					
Cluster Mean	**0.243**	0.058	0.078	0.075	473.01
Different Groups (Pairwise)	(2, 3, 4)	(1)	(1)	(1)	(p<0.000)
Std. Dev.	0.429	0.233	0.268	0.263	
Diabetes Mellitus					
Cluster Mean	**0.727**	0.433	0.128	0.103	2597.94
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2)	(1, 2)	(p<0.000)
Std. Dev.	0.446	0.496	0.334	0.304	
Drugs/Alcohol Abuse					
Cluster Mean	0.036	0.016	**0.113**	0.062	182.68
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.187	0.127	0.317	0.240	
Dyslipidemia					
Cluster Mean	**0.812**	0.738	0.507	0.043	3452.88
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.391	0.440	0.500	0.204	
Family History of Stroke					
Cluster Mean	**0.125**	0.045	0.068	0.068	99.46
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2)	(1, 2)	(p<0.000)

Medical History Variable	**1. Metabolic (n=2148)**	**2. Cardio-vascular (n=1702)**	**3. Behavior (n=2001)**	**4. Low Risk (n=2680)**	**Chi-square Statistic (p-value)**
Std. Dev.	0.331	0.207	0.252	0.252	
Heart Failure					
Cluster Mean	0.107	**0.278**	0.035	0.037	795.29
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2)	(1, 2)	(p<0.000)
Std. Dev.	0.309	0.448	0.185	0.190	
Hormone Replacement					
Cluster Mean	0.014	0.008	0.007	**0.019**	18.43
Different Groups (Pairwise)		(4)	(4)	(2, 3)	(p<0.000)
Std. Dev.	0.117	0.087	0.083	0.138	
Hypertension					
Cluster Mean	0.926	**0.932**	0.923	0.429	2569.98
Different Groups (Pairwise)	(4)	(4)	(4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.261	0.252	0.267	0.495	
Migraine					
Cluster Mean	**0.034**	0.006	0.017	0.029	38.13
Different Groups (Pairwise)	(2, 3)	(1, 4)	(1)	(2)	(p<0.000)
Std. Dev.	0.180	0.080	0.131	0.167	
Obesity/Overweight					
Cluster Mean	**0.839**	0.089	0.215	0.300	2853.28
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.368	0.285	0.411	0.458	
Previous Stroke					
Cluster Mean	0.308	0.343	**0.392**	0.035	994.35
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.462	0.475	0.488	0.184	
Previous TIA					
Cluster Mean	0.100	**0.127**	0.091	0.053	77.59
Different Groups (Pairwise)	(2, 4)	(1, 3, 4)	(2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.300	0.334	0.288	0.224	
Prosthetic Heart Valve					
Cluster Mean	0.014	**0.019**	0.007	0.007	18.10
Different Groups (Pairwise)		(3, 4)	(2)	(2)	(p<0.000)
Std. Dev.	0.119	0.136	0.083	0.084	
Peripheral Vascular Disease					
Cluster Mean	0.099	**0.113**	0.057	0.033	134.45
Different Groups (Pairwise)	(3, 4)	(3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.299	0.316	0.232	0.179	
Renal Insufficiency					
Cluster Mean	**0.165**	0.071	0.039	0.022	411.86

Medical History Variable	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	Chi-square Statistic (p-value)
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2)	(1, 2)	(p<0.000)
Std. Dev.	0.372	0.256	0.194	0.148	
Sickle Cell					
Cluster Mean	0.001	0.001	0.000	0.001	0.69
Different Groups (Pairwise)					(p=0.876)
Std. Dev.	0.031	0.024	0.022	0.033	
Sleep Apnea					
Cluster Mean	**0.079**	0.015	0.008	0.010	269.88
Different Groups (Pairwise)	(2, 3, 4)	(1)	(1)	(1)	(p<0.000)
Std. Dev.	0.269	0.123	0.092	0.102	
Smoker					
Cluster Mean	0.116	0.095	**0.684**	0.189	2358.96
Different Groups (Pairwise)	(3, 4)	(3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.321	0.294	0.465	0.391	

Notes. The cluster means indicate the percentage of patients in the patient group that had the medical condition. The numbers in parentheses represent the groups that are significantly different from the particular group at the 0.05 significance level, based on the pairwise post-hoc Dunn test with Bonferroni adjustments. The bolded numbers indicate the highest group centroid for that variable.

Underlying Dimensions

A multiple group discriminant analysis was conducted to validate our model. We performed the SAS canonical discriminant analysis (i.e., SAS CANDISC) as a component of the SAS FASTCLUS procedure. We used four clusters as the dependent variables and the 22 medical history variables as the independent variables. Table 4 shows the discriminant analysis findings. The canonical loadings signify the correlations between the 22 medical history variables and the three discriminant functions. This is the result of N-1 dimensions, where N indicates the number of clusters (4 clusters). The measure of the multivariate association (γ_6), which is the average of the squared canonical correlations, is .583. Thus, 58.3% of the variance in patient group membership is explained by the set of the 22 medical history taxon variables (Cramer & Nicewander, 1979).

Table 4: Results of Canonical Discriminant Analysis

Canonical Correlation Function	Eigenvalue	Canonical Correlation (R_c)	Squared Canonical Correlation	p-value
1	2.312	0.835	0.698	<.0001
2	1.214	0.741	0.548	<.0001
3	1.012	0.709	0.503	<.0001

	Canonical Loadings			Canonical Coefficients		
Variables	Function 1	Function 2	Function 3	Function 1	Function 2	Function 3
Atrial Fibrillation / Flutter	0.258	-0.295	-0.314	0.195	-0.321	-0.299

Variables	Canonical Loadings			Canonical Coefficients		
	Function 1	Function 2	Function 3	Function 1	Function 2	Function 3
Coronary Artery Disease	**0.604**	-0.455	-0.326	0.600	-0.693	-0.560
Carotid Stenosis	0.139	-0.090	0.039	0.012	-0.026	-0.023
Current Pregnancy	-0.044	0.002	-0.030	-0.020	-0.025	-0.008
Depression	0.151	0.267	0.031	0.040	0.161	-0.008
Diabetes Mellitus	**0.584**	0.331	-0.114	0.565	0.511	-0.243
Drugs/Alcohol Abuse	-0.102	-0.013	0.168	-0.024	-0.073	0.023
Dyslipidemia	**0.741**	0.004	0.209	0.819	0.061	0.382
Family History of Stroke	0.045	0.135	0.026	0.018	0.031	-0.023
Heart Failure	0.269	-0.208	-0.195	0.117	-0.171	-0.140
Hormone Replacement	-0.031	0.038	-0.037	-0.005	-0.014	-0.010
Hypertension	**0.558**	-0.128	0.385	0.456	-0.220	0.674
Migraine	-0.020	0.087	-0.005	-0.005	-0.037	-0.020
Obesity/Overweight	0.258	**0.724**	0.034	0.290	0.943	0.022
Previous Stroke	0.298	-0.132	0.299	0.209	-0.160	0.397
Previous TIA	0.104	-0.050	0.014	-0.005	-0.016	-0.002
Prosthetic Heart Valve	0.049	-0.011	-0.026	0.017	0.014	0.030
Peripheral Vascular Disease	0.148	-0.019	-0.019	0.008	0.010	-0.017
Renal Insufficiency	0.213	0.173	0.002	0.047	0.125	0.037
Sickle Cell	-0.005	0.008	-0.008	-0.008	0.001	0.008
Sleep Apnea	0.138	0.183	-0.005	0.002	0.072	0.027
Smoker	-0.206	-0.189	**0.672**	-0.109	-0.307	0.930

Notes. The high loadings in canonical functions ±|0.40| are bold numbers. Wilk's Lambda = 0.07 (F value = 562.95, df = 66, p <.0001) indicates a significant overall multivariate relationship for the 22 medical history taxon variables. The three canonical correlations (R_{c1} = 2.312, R_{c2} = 1.214, R_{c3} = 1.012) are statistically significant.

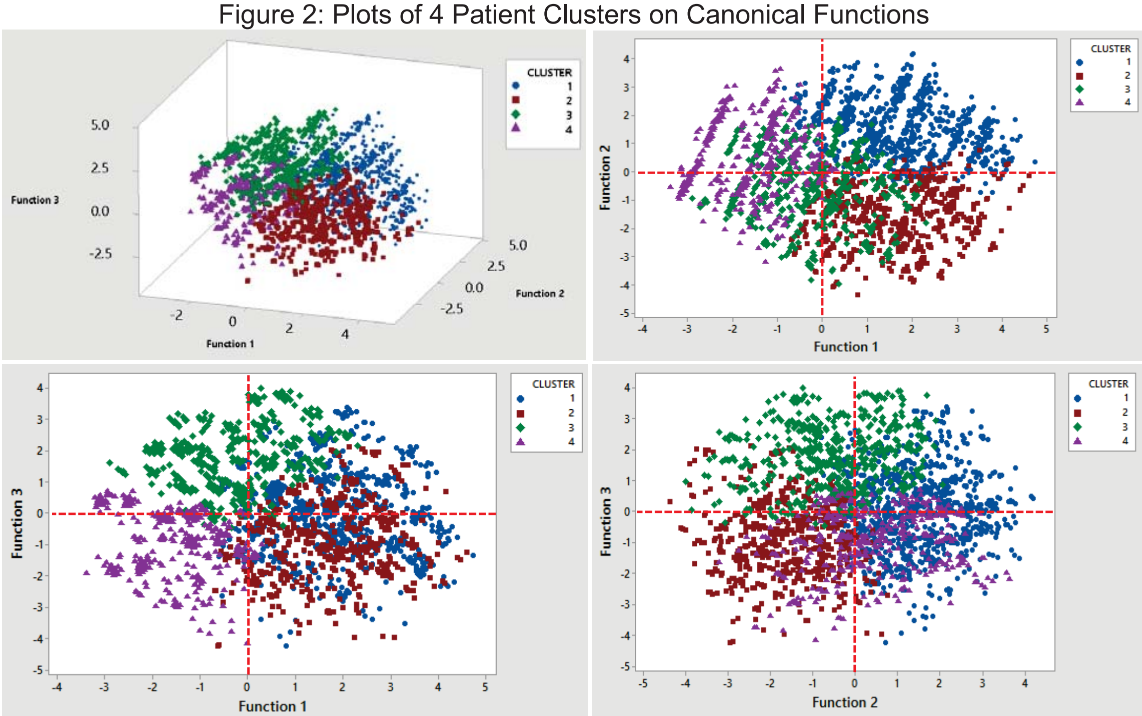
We interpret the underlying stroke patient dimensions based on the standardized canonical loadings. As indicated on the "function 1" column of canonical loadings in Table 4, the first canonical function has high loadings on some of the cardiovascular diseases and metabolic syndrome, such as coronary artery disease, diabetes mellitus, dyslipidemia, and hypertension. The second canonical function has a high loading on obesity. The third canonical function has a high loading on smoker.

However, labeling the dimensions using the canonical loadings is still not trivial. As an alternative approach, a graphical analysis is conducted in order to properly label the three dimensions. Derived from the canonical coefficients, the 4 clusters of the patients are depicted on a 3-dimensional coordinate, as shown in Figure 2. Patients in each cluster are located within distinct spaces in the 3-dimensional coordinate.

Based on the 3-D and 2-D plots of Figure 2 and the averages of the canonical coefficients for all patients, the locations of the patient clusters are summarized, as indicated in Table 5. Along with the first dimension (i.e., function 1), the metabolic and cardiovascular patient clusters have a high-level average canonical coefficient, whereas behavior and low risk patients have a low- level average canonical coefficient. This pattern is similar to the average ages of the patient clusters. Considering this similar pattern, we interpret the first dimension as being "age related," where a high canonical coefficient represents an older patient, and a low coefficient represents a younger patient.

Along with the second dimension (i.e., function 2), the average canonical coefficient is high for the metabolic patient cluster and low for the cardiovascular patient cluster. The average canonical coefficients for the behavior and low risk clusters are between high and low. With the exception of the low risk patient cluster, this pattern is similar to the proportion of female patients in each patient cluster. Accordingly, we interpret the second dimension as being "gender related," where a high canonical coefficient tends to be female, and a low coefficient tends to be male.

Figure 2: Plots of 4 Patient Clusters on Canonical Functions



Cluster Key: 1 = Metabolic; 2 = Cardiovascular; 3 = Behavior; 4 = Low Risk

Lastly, along with the third dimension (i.e., function 3), the average canonical coefficient is high for the behavior patient cluster and low for the cardiovascular and low risk patient clusters. The average canonical coefficient for the metabolic cluster is between high and low. This pattern is similar to the proportion of black patients in each patient cluster. The pattern is the opposite of the proportion of white patients. Accordingly, we interpret the third dimension as being "race related," where a high canonical coefficient tends to be black, and a low coefficient tends to be white. The black-white race dimension is plausible with our patient sample, as the two race categories account for 98.8% of all patients (i.e., 21.0% black and 77.8% white).

Table 5: Pattern Comparison between Canonical Functions and Demographics

Function or Demographic Characteristic	**1. Metabolic (n=2148)**	**2. Cardio-vascular (n=1702)**	**3. Behavior (n=2001)**	**4. Low Risk (n=2680)**
Function 1	High (1.58)	High (1.58)	Middle (-0.50)	Low (-1.90)
Function 2	High (1.51)	Low (-1.62)	Middle (-0.61)	Middle (0.27)

Function or Demographic Characteristic	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)
Function 3	Middle (0.06)	Low (-0.88)	High (1.70)	Low (-0.76)
Age (years)	67.6	74.2	62.9	63.6
Female (%)	53.1%	47.5%	48.8%	54.9%
Black (%)	23.6%	15.2%	26.5%	18.5%
White (%)	75.3%	83.9%	72.6%	79.7%

Notes. The numbers in parentheses for the canonical functions represent the averages of the canonical coefficients for each cluster.

Statistical Cross-validation

A cross-loading analysis was performed to ensure that the estimates were stable. Table 6 presents the results of the cross-validation using a Jackknifing method in the SAS DISCRIM procedure. The total error rate (i.e., the percentage of misclassified cases) from the statistical cross-validation is 0.029, indicating that the 22 medical history taxon variables aptly classified the stroke patients.

Table 6: Cross-validation of the Number of Observations and Percentages

	1	2	3	4	
From/To	**Metabolic**	**Cardiovascular**	**Behavior**	**Low Risk**	**Total**
1	**2052 (95.53%)**	67 (3.12%)	12 (0.56%)	17 (0.79%)	2148 (100%)
2	17 (1%)	**1662 (97.65%)**	20 (1.18%)	3 (0.18%)	1702 (100%)
3	6 (0.3%)	29 (1.45%)	**1937 (96.8%)**	29 (1.45%)	2001 (100%)
4	2 (0.07%)	39 (1.46%)	4 (0.15%)	**2635 (98.32%)**	2680 (100%)
Error Rates from Cross-validation	0.045	0.024	0.032	0.017	**0.029**

Cluster Comparisons based on Associated Variables

Based on the clusters, we systematically examined the variables associated with patient group membership. The following sub-sections present the differences among the four patient clusters in terms of demographic variables, time performance variables, and other system variables.

<u>Demographic Variables</u>

Table 7 presents the comparisons among the four patient groups, based on each of the demographic variables. The “metabolic” patient group has the highest BMI (Body Mass Index) among the four groups. The “cardiovascular” patient group is the oldest group, has the highest proportion of white patients, and has the highest percent of individuals with Medicare and private insurance. Such differences are statistically significant from each of the other three patient groups. The “behavior” patient group has the highest proportion of black patients. Finally, the “low risk” patient group has the highest proportion of female patients.

Table 7: Demographic Variables Associated with Patient Groups

Variable	**1. Metabolic (n=2148)**	**2. Cardio-vascular (n=1702)**	**3. Behavior (n=2001)**	**4. Low Risk (n=2680)**	**F or Chi-square Statistic (p-value)**
Age (Years)					
Cluster Mean*	67.63	**74.23**	62.95	63.56	247.04
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2)	(1, 2)	(p<0.000)
Std. Dev.	12.72	11.96	13.62	17.24	
Female					
Cluster Mean**	0.53	0.48	0.49	**0.55**	30.75
Different Groups (Pairwise)	(2, 3)	(1, 4)	(1, 4)	(2, 3)	(p<0.000)
Std. Dev.	0.50	0.50	0.50	0.50	
White					
Cluster Mean**	0.75	**0.84**	0.73	0.80	81.83
Different Groups (Pairwise)	(2, 4)	(1, 3, 4)	(2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.43	0.37	0.45	0.40	
Black					
Cluster Mean**	0.24	0.15	**0.26**	0.19	89.54
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.42	0.36	0.44	0.39	
Asian					
Cluster Mean**	0.00	0.01	0.01	0.01	4.32
Different Groups (Pairwise)					(p=0.229)
Std. Dev.	0.06	0.07	0.07	0.09	
Hispanic					
Cluster Mean**	0.02	0.01	0.01	0.02	7.79
Different Groups (Pairwise)		(4)		(2)	(p=0.051)
Std. Dev.	0.13	0.11	0.12	0.15	
BMI (Index)					
Cluster Mean*	**31.99**	27.08	26.80	27.24	291.33
Different Groups (Pairwise)	(2, 3, 4)	(1)	(1, 4)	(1, 3)	(p<0.000)
Std. Dev.	6.86	6.45	6.60	6.64	
Medicare					
Cluster Mean**	0.73	**0.81**	0.55	0.50	556.33
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.45	0.39	0.50	0.50	
Medicaid					
Cluster Mean**	0.15	0.14	**0.20**	0.10	79.30
Different Groups (Pairwise)	(3, 4)	(3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)

Variable	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	F or Chi-square Statistic (p-value)
Std. Dev.	0.36	0.35	0.40	0.31	
Private Insurance					
Cluster Mean**	0.48	**0.57**	0.39	0.52	149.37
Different Groups (Pairwise)	(2, 3, 4)	(1, 3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.50	0.49	0.49	0.50	

* The cluster mean indicates the averages of the specified variable by each patient group. The numbers in parentheses represent the groups that are significantly different from the particular group at the 0.05 significance level, based on the Tukey pairwise tests.
** The cluster means indicate the percentage of patients in the patient group that met the specified variable condition. The numbers in parentheses represent the groups that are significantly different from the particular group at the 0.05 significance level, based on the pairwise post-hoc Dunn test with Bonferroni adjustments.
Note. The bolded numbers indicate the highest group centroid for that variable.

Time Performance Variables

Figure 3 visually describes the steps of the stroke patient care process and the time intervals that are used as performance measures for stroke patient care. Once a stroke patient arrives at the hospital, the patient receives a brain imaging in the CT (computer tomography) scan room in order to determine the type of stroke that the patient has. Then, the eligible patient receives a thrombolytic medication called TPA (Tissue Plasminogen Activator) (Davis et al., 2009). Once emergency care after TPA infusion is complete, the patient is admitted either to the regular floor or the ICU (Intensive Care Unit). Then, the patient is discharged from the hospital as his or her condition improves. The most critical in-hospital time interval is the door-to-needle (DTN) time. DTN is broken down into two sub-time intervals: Door-to-CT (DTC) and CT-to-Needle (CTN).

Figure 3: Stroke Patient Care Process Flow

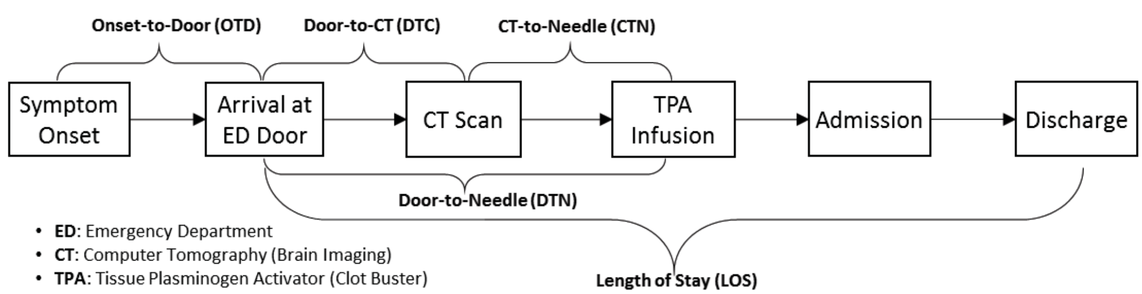


Table 8 presents comparisons among the four patient groups, based on each of the time performance variables. Interestingly, there is no significant difference in the DTN time among the four patient groups. However, among those who received TPA, the “metabolic” patient group took the longest DTC time (i.e., the front-end of the care process), while the “low risk” patient group took the shortest DTC time. Due to patients being overweight, it is possible that additional time may have been spent for a metabolic patient prior to the CT scan, e.g., transfer from the EMS stretcher to the CT bed. For the low risk patient group, the diagnosis time may have been shorter prior to the CT scan due to the smaller number of comorbidities that a patient had.

On the other hand, the “cardiovascular” patient group took the longest CTN time (i.e., the back-end of the care process), and the “metabolic” patient group took the shortest time. The TPA treatment decision by the neurologist may have been longer for the cardiovascular patient group because the patients in that group may have already been treated with anti-coagulant (blood-thinning) medications. Because using TPA treatment on such patients may cause certain complications, additional steps may have been taken to make a TPA treatment decision.

The differences in DTC and CTN among the different patient groups provide important information to stroke care teams regarding the prompt and proper use of the patient’s medical history. Obtaining the medical history information prior to the patient’s arrival or immediately after the patient’s arrival could be a critical factor in allocating the right resources, both in terms of clinicians and diagnostic/treatment equipment.

Table 8: Time Performance Variables Associated with Patient Groups

Variable	**1. Metabolic (n=2148)**	**2. Cardiovascular (n=1702)**	**3. Behavior (n=2001)**	**4. Low Risk (n=2680)**	**F-Statistic (p-value)**
Door-to-Needle (TPA Treated)					
Cluster Mean	55.70	58.36	55.94	56.89	0.31
Different Groups (Pairwise)					(p=0.818)
Std. Dev.	31.01	25.14	27.81	27.94	
Door-to-CT (TPA Treated)					
Cluster Mean	**18.92**	15.80	17.31	15.43	2.67
Different Groups (Pairwise)	(4)			(1)	(p=0.046)
Std. Dev.	16.26	10.16	17.01	12.60	
CT-to-Needle (TPA Treated)					
Cluster Mean	36.78	**42.55**	38.63	41.46	2.48
Different Groups (Pairwise)	(2)	(1)			(p=0.060)
Std. Dev.	24.60	22.33	21.29	25.11	
Door-to-CT (TPA Not Treated)					
Cluster Mean	109.90	157.73	168.59	123.68	1.16
Different Groups (Pairwise)					(p=0.323)
Std. Dev.	293.20	777.27	1458.32	589.30	
Door-to-CT (All Patients)					
Cluster Mean	95.18	139.61	144.24	106.50	1.24
Different Groups (Pairwise)					(p=0.293)
Std. Dev.	269.96	727.56	1336.78	541.73	
Onset-to-Door (TPA Treated)					
Cluster Mean	99.59	91.26	90.36	86.26	1.99
Different Groups (Pairwise)					(p=0.114)
Std. Dev.	65.05	56.45	57.88	59.16	
Onset-to-Door (TPA Not Treated)					
Cluster Mean	399.28	339.77	346.41	368.28	1.20
Different Groups (Pairwise)					(p=0.307)
Std. Dev.	500.93	397.40	441.47	685.05	

Variable	1. Metabolic (n=2148)	2. Cardiovascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	F-Statistic (p-value)
Onset-to-Door (All Patients)					
Cluster Mean	302.95	278.69	265.82	283.83	0.76
Different Groups (Pairwise)					(p=0.518)
Std. Dev.	436.89	362.44	385.84	588.36	
Length of Stay (Days)					
Cluster Mean	7.193	7.378	7.828	**8.631**	4.59
Different Groups (Pairwise)	(4)	(4)		(1, 2)	(p=0.003)
Std. Dev.	12.180	13.704	12.826	17.709	

Notes. The cluster mean indicates the average time spent for the specified time interval by each patient cluster. The numbers in parentheses represent the groups that are significantly different from the particular group at the 0.05 significance level, based on the Tukey pairwise test with Bonferroni adjustments. The bolded numbers indicate the highest group centroid for that variable.

Other System and Outcome Variables

Table 9 presents the comparisons among the four patient groups, based on other system and outcome variables. As for the type of stroke, the "metabolic" patient group had the highest proportion of ischemic and TIA (transient ischemic attack) patients among the four groups, whereas the "low risk" group had the highest proportion of hemorrhagic patients. Also, the "cardiovascular" patient group had the highest usage of EMS (emergency medical services). This could be partly because the cardiovascular patient group is the oldest group that has already experienced the benefit of using EMS over private transportation. Also, the cardiovascular patient group had strokes in healthcare settings, compared to the other patient groups. In line with the highest EMS usage, the cardiovascular group had the highest percentage of pre-notifying the hospital of the pending patient's arrival and the highest percentage of stroke alert activations. However, the cardiovascular group still had the highest percentage regarding the in-hospital mortality rate, which is due to the high severity of medical conditions prior to having a stroke.

Table 9: Other Variables Associated with Patient Groups

Variable	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	Chi-square Statistic (p-value)
Stroke Type - Ischemic					
Cluster Mean	0.730	**0.746**	0.705	0.643	68.63
Different Groups (Pairwise)	(4)	(3, 4)	(2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.44	0.44	0.46	0.48	
Stroke Type - TIA					
Cluster Mean	0.134	**0.135**	0.112	0.091	30.11
Different Groups (Pairwise)	(3, 4)	(3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.34	0.34	0.32	0.29	
Stroke Type - Hemorrhagic					

Variable	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	Chi-square Statistic (p-value)
Cluster Mean	0.124	0.116	0.177	**0.255**	196.13
Different Groups (Pairwise)	(3, 4)	(3, 4)	(1, 2, 4)	(1, 2, 3)	(p<0.000)
Std. Dev.	0.33	0.32	0.38	0.44	
Stroke Type - No Stroke					
Cluster Mean	**0.011**	0.002	0.005	**0.011**	15.01
Different Groups (Pairwise)	(2, 3)	(1, 4)	(1)	(2)	(p=0.002)
Std. Dev.	0.10	0.04	0.07	0.10	
Arrival Mode - EMS					
Cluster Mean	0.471	**0.557**	0.473	0.457	46.75
Different Groups (Pairwise)	(2)	(1, 3, 4)	(2)	(2)	(p<0.000)
Std. Dev.	0.50	0.50	0.50	0.50	
Arrival Mode - Transfer					
Cluster Mean	**0.205**	0.155	0.183	0.171	18.37
Different Groups (Pairwise)	(2, 4)	(1, 3)	(2)	(1)	(p<0.000)
Std. Dev.	0.40	0.36	0.39	0.38	
Arrival Mode - Private					
Cluster Mean	0.274	0.222	0.299	**0.310**	45.06
Different Groups (Pairwise)	(2, 4)	(1, 3, 4)	(2)	(1, 2)	(p<0.000)
Std. Dev.	0.45	0.42	0.46	0.46	
Arrival Mode - Inpatient					
Cluster Mean	0.050	**0.065**	0.044	0.062	10.71
Different Groups (Pairwise)		(3)	(2, 4)	(3)	(p=0.013)
Std. Dev.	0.22	0.25	0.21	0.24	
Occurrence - Non Healthcare					
Cluster Mean	0.875	0.818	**0.891**	0.879	50.47
Different Groups (Pairwise)	(2)	(1, 3, 4)	(2)	(2)	(p<0.000)
Std. Dev.	0.33	0.39	0.31	0.33	
Occurrence - Acute care					
Cluster Mean	0.009	**0.016**	0.013	0.011	4.93
Different Groups (Pairwise)	(2)	(1)			(p=0.177)
Std. Dev.	0.09	0.13	0.11	0.11	
Occurrence - Chronic care					
Cluster Mean	0.048	**0.090**	0.042	0.040	61.92
Different Groups (Pairwise)	(2)	(1, 3, 4)	(2)	(2)	(p<0.000)
Std. Dev.	0.21	0.29	0.20	0.19	
Occurrence - Outpatient					
Cluster Mean	**0.014**	0.007	0.005	0.006	13.26
Different Groups (Pairwise)	(2, 3, 4)	(1)	(1)	(1)	(p=0.004)
Std. Dev.	0.12	0.08	0.07	0.08	
Pre-notification					
Cluster Mean	0.368	**0.390**	0.346	0.346	11.45
Different Groups (Pairwise)		(3, 4)	(2)	(2)	(p=0.010)
Std. Dev.	0.48	0.49	0.48	0.48	

Variable	1. Metabolic (n=2148)	2. Cardio-vascular (n=1702)	3. Behavior (n=2001)	4. Low Risk (n=2680)	Chi-square Statistic (p-value)
Stroke Alert					
Cluster Mean	0.226	**0.254**	0.240	0.219	8.45
Different Groups (Pairwise)		(4)		(2)	(p=0.038)
Std. Dev.	0.42	0.44	0.43	0.41	
TPA Exclusion					
Cluster Mean	0.251	**0.293**	0.244	0.230	23.01
Different Groups (Pairwise)	(2)	(1, 3, 4)	(2)	(2)	(p<0.000)
Std. Dev.	0.43	0.46	0.43	0.42	
Delay due to Patient Condition					
Cluster Mean	**0.019**	0.004	0.010	0.013	19.43
Different Groups (Pairwise)	(2, 3)	(1, 4)	(1)	(2)	(p<0.000)
Std. Dev.	0.14	0.06	0.10	0.11	
Arrived during Night Shift					
Cluster Mean	0.422	0.461	0.468	**0.471**	13.89
Different Groups (Pairwise)	(2, 3, 4)	(1)	(1)	(1)	(p=0.003)
Std. Dev.	0.49	0.50	0.50	0.50	
Arrived on the Weekend					
Cluster Mean	0.243	0.262	0.259	0.264	3.11
Different Groups (Pairwise)					(p=0.376)
Std. Dev.	0.43	0.44	0.44	0.44	
In-hospital Mortality					
Cluster Mean	0.069	**0.116**	0.085	0.108	26.25
Different Groups (Pairwise)	(2, 4)	(1, 4)	(2, 4)	(1, 3)	(p<0.000)
Std. Dev.	0.25	0.32	0.28	0.31	
Discharge to Home					
Cluster Mean	**0.521**	0.386	**0.521**	0.485	62.89
Different Groups (Pairwise)	(2, 4)	(1, 3, 4)	(2)	(1, 2)	(p<0.000)
Std. Dev.	0.50	0.49	0.50	0.50	

Notes. The cluster means indicate the percentage of patients in the patient group that met the specified variable condition. The numbers in parentheses represent the groups that are significantly different from the particular group at the 0.05 significance level, based on the pairwise post-hoc Dunn test with Bonferroni adjustments. The bolded numbers indicate the highest group centroid for that variable.

DISCUSSION

Theoretical Contributions

We developed a numerical taxonomy of stroke patient medical histories by performing a cluster analysis. This taxonomy contributes to the medical literature by extending patient classification based on comorbidities and risk factors to the realm of stroke patient care and relating such patient classification with the operational variables. This study also contributes to the operations management literature by extensively using medical history variables to classify stroke patients and by relating such patient classification with the medically relevant operational issues. That is, we go beyond extant operations literature on patient classification, which is primarily focused on issues like patient appointment scheduling. Such literature only addresses time performances but does not address the operational issues that can influence medical outcomes. Results of our study suggest that patients in different patient groups may require different levels of resources, both human resources and physical resources such as diagnostic and treatment devices and medications. A hospital can improve time performances by allocating appropriate resources. For example, because the “metabolic” patient group takes a longer DTC time (i.e., the front-end process) potentially due to a longer time for patient bed change, more crew members may be assigned for the bed change. Also, because the “cardiovascular” patient group takes a longer CTN time (i.e., the back-end process), the key decision makers for TPA treatment such as radiologists and neurologists should start getting involved in the care process as early as possible. Moreover, because patients in the “metabolic” and “cardiovascular” groups are more likely to have an ischemic stroke or TIA (Transient Ischemic Attack), stroke care clinicians can begin preparing for ischemic stroke treatments including TPA treatment as soon as they identify the patient classification group. Our study can be an important platform for future research on how such patient classification affects operational outcomes that are associated with medical outcomes.

Our study contributes to the service operations literature, as it provides insights on how understanding customer (i.e., patients) groups might improve operational performance and customer satisfaction in a service setting where there is an inherently heterogeneous customer population.

Managerial Contributions

Our analysis results indicate that understanding patient groups based on medical history information can be critical in treating stroke patients in a timely manner. Hospital organizations must ensure the fast compilation of the patient’s medical history or timely access to the patient’s medical records. Two issues should be considered. First, the role of EMS providers is important for obtaining the patient’s medical history. If the EMS provider can make available the important medical history of a patient when he/she gives advanced notification of the pending stroke patient to the destination hospital, the hospital’s stroke team can be more appropriately organized, and the relevant equipment or medication can be prepared in advance. Our taxonomy suggests that a hospital may benefit from developing specific care procedures (i.e., resource allocations) for each of the stroke patient groups. That way, the in-hospital diagnosis and treatment time can be minimized. Currently, the *GWTG-Stroke* and *Target:Stroke* initiatives by the American Heart Association do not address ways of obtaining a patient’s medical history. Coordination between EMS and ED is certainly important, and developing a standard operating procedure on obtaining a patient’s medical history can improve such coordination. Second, the electronic medical records (EMRs) of patients should be readily available to the stroke team and shared between hospitals and between EMS providers and hospitals. That way, the stroke team does not have to rely on verbal information from the patient or the patient’s family members about his or her medical history. Moreover, from the physician’s point-of-view, many aspects of the patients need to be considered for the treatment plan in a short time period. Having 4 patient groups instead of 22 medical history variables would certainly reduce the information overload for the physician.

Currently, our research site hospital uses a universal target and reward criterion, regardless of the patient’s condition. As our analysis shows, there might be inherent differences in the realistically required time to treat stroke patients among the different patient groups. The hospital might have to consider a differentiated reward system for different patient groups in order to ensure fairness.

CONCLUSIONS

For a rich understanding of patient classification, we developed a numerical taxonomy using 22 medical history variables, including comorbidities and risk factors associated with stroke. We performed a *k-means* cluster analysis method, which resulted in 4 distinct patient groups. We labeled the 4 patient groups as "metabolic," "cardiovascular," "behavior," and "low risk," based on the holistic judgments from the cluster means. These 4 patient groups were located within distinct spaces in a 3-dimensional space, where the dimensions were developed based on canonical functions. The patient groups are different in terms of several demographic variables, such as age, gender, race, and the type of insurance that the patient has. More interestingly, they are different in terms of the operational aspects of stroke care, such as in-hospital time performance, mode of hospital arrival, stroke occurrence location, EMS pre-notification, stroke alert, etc. This suggests that the hospital could benefit from standardizing the patient care procedures for each of the patient groups and standardizing the way that the stroke team obtains the patient's medical history. Future research should investigate the casual relationship between patient classification and operational/medical outcomes.

REFERENCES

American Heart Association (2018). *Get With The Guidelines®-Stroke Overview* http://www.heart.org/HEARTORG/Professional/GetWithTheGuidelines/GetWithTheGuidelines-Stroke/Get-With-The-Guidelines-Stroke-Overview_UCM_308021_Article.jsp#.WbiVGLL5i9K

American Heart Association (2018). *What is Cardiovascular Disease?* http://www.heart.org/HEARTORG/Conditions/What-is-Cardiovascular-Disease_UCM_301852_Article.jsp#.WxyzKO7RX4Y

American Heart Association (2018). *What is Metabolic Syndrome?* http://www.heart.org/idc/groups/heart-public/@wcm/@hcm/documents/downloadable/ucm_300322.pdf

Berta, W. B., & Baker, R. (2004). Factors that impact the transfer and retention of best practices for reducing error in hospitals. *Health Care Management Review, 29*(2), 90-97.

Cayirli, T., Veral, E., & Rosen, H. (2008). Assessment of patient classification in appointment system design. *Production and Operations Management, 17*(3), 338-353.

Cramer, E. M., & Nicewander, W. A. (1979). Some symmetric, invariant measures of multivariate association. *Psychometrika, 44*(1), 43-54.

Davis, S. M., & Donnan, G. A. (2009). 4.5 Hours. *Stroke, 40*(6), 2266-2267.

Hu, W., Chan, C. W., Zubizarreta, J. R., & Escobar, G. J. (2018). An examination of early transfers to the ICU based on a physiologic risk score. *Manufacturing & Service Operations Management, 20*(3), 531-549.

Kolisch, R., & Sickinger, S. (2008). Providing radiology health care services to stochastic demand of different customer classes. *OR spectrum, 30*(2), 375-395.

Miller, J. G., & Roth, A. V. (1994). A taxonomy of manufacturing strategies. *Management Science, 40*(3), 285-304.

National Center for Health Statistics (US. (2017). *Health, United States, 2016: With Chartbook on Long-term Trends in Health*.

Salzarulo, P. A., Bretthauer, K. M., Côté, M. J., & Schultz, K. L. (2011). The impact of variability and patient information on health care system performance. *Production and Operations Management, 20*(6), 848-859.

Salzarulo, P. A., Mahar, S., & Modi, S. (2016). Beyond patient classification: Using individual patient characteristics in appointment scheduling. *Production and Operations Management, 25*(6), 1056-1072.

SAS/STAT® 13.2 User's Guide (2018) *The ACECLUS Procedure* http://support.sas.com/documentation/cdl/en/statug/67523/HTML/default/viewer.htm#statug_aceclus_overview.htm

SAS/STAT® 13.2 User's Guide (2018) *The FASTCLUS Procedure* http://support.sas.com/documentation/cdl/en/statug/67523/HTML/default/viewer.htm#statug_fastclus_overview.htm

Saver, J. L., Smith, E. E., Fonarow, G. C., Reeves, M. J., Zhao, X., Olson, D. M., & Schwamm, L. H. (2010). The "golden hour" and acute brain ischemia: Presenting features and lytic therapy in> 30 000 patients arriving within 60 minutes of stroke onset. *Stroke, 41*(7), 1431-1439

Tucker, A. L., Nembhard, I. M., & Edmondson, A. C. (2007). Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management Science, 53*(6), 894-907.

Vanfleteren, L. E., Spruit, M. A., Groenen, M., Gaffron, S., van Empel, V. P., Bruijnzeel, P. L., ... & Franssen, F. M. (2013). Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine, 187*(7), 728-735.

Vu, T., Finch, C. F., & Day, L. (2011). Patterns of comorbidity in community-dwelling older people hospitalised for fall-related injury: A cluster analysis. *BMC Geriatrics, 11*(1), 45.

DECISION SCIENCES INSTITUTE
Cloud ERP adoption in healthcare industry

Uzay Damali
University of Wisconsin La Crosse
udamali@uwlax.edu

Mehmet Kocakulah
University of Wisconsin La Crosse
mkocakulah@uwlax.edu

Ahmet Ozkul
University of New Haven
aozkul@newhaven.edu

ABSTRACT

In this study, we explored the factors that encourage or discourage healthcare organizations to adopt Cloud ERP, and whether organizational contingencies (e.g., organizational size) affect this adoption. Based on literature review, we proposed a research model of Cloud ERP adoption. Our model and findings can guide healthcare IT manager and support future research on this topic.

KEYWORDS: Cloud, ERP, Healthcare, Technology adoption

INTRODUCTION

Enterprise resource planning (ERP) systems support effective decision making in organizations by integrating various business functions (e.g., accounting, finance, human resources and operations) through extensive data processing (i.e., collecting, analyzing and sharing of data). Recently, there has been an accelerated transition from on-premise ERP systems to Cloud ERP systems in a variety of service and manufacturing industries. On-premise ERP systems (including data, software, and database servers) are internally hosted, and maintained by user organizations; while Cloud ERP systems are externally hosted, and related data are maintained by a third party vendor (Peng and Gala, 2014). To investigate this transition, a significant body of literature has recently been made available.

This literature primarily identified three issues: (1) the factors that encourage or discourage organizations adopt Cloud ERP, (2) whether the organizational contingencies (primarily organization size) influences Cloud ERP adoption, and (3) whether the industry type influences on Cloud ERP adoption. Regarding the first issue, Peng and Gala (2014) building on compressive literature review identified the factors that promote Cloud ERP adoption, these are: (1) reduced cost, (2) improved system speed and performance, (3) improved system upgrade, and (4) improved mobility and accessibility. They also identified factors that obstruct Cloud ERP adopting, these are: (1) lack of data privacy and control, (2) lack of data security, (3) vendor lock-in, (4) integration difficulties, and (5) organizational challenges.

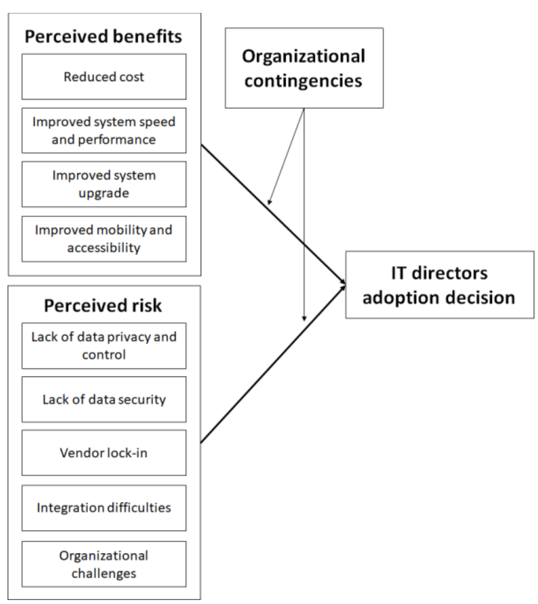
Regarding the second and third issue, this literature found evidence that the rate of the Cloud ERP adoption was dependent on organizational contingencies, specifically organizational size. This literature found that small and medium sized organizations were more likely to adopt Cloud ERP compared to large sized organizations. Also, they found out that Cloud ERP adoption has been successful in a wide variety of service and manufacturing industries. However, to the best of our knowledge, there are only a few number of studies on Cloud ERP adoption focused on the healthcare industry.

To address this gap, we investigate the following research question: *what factors influence Cloud ERP adoption in healthcare industry, and what organizational contingencies affects the adoption rate?* Healthcare is one of the major industries in the US (17 % of the GDP). There has been a significant effort in this industry to reduce costs and improve quality. One of these efforts that has become increasingly important is the adoption of Cloud technologies. However, the rate of adoption, particularly Cloud ERP, has been relatively slow compared to other industries. We believe our research findings can guide healthcare IT managers for the Cloud ERP adoption decisions.

LITERATURE REVIEW AND RESEARCH MODEL

The fundamental difference between on-premise ERP and Cloud ERP is where the ERP system is located and maintained. On-premise ERP is maintained by the organization internally, while a Cloud ERP system is maintained by the ERP vendor externally (Peng and Gala, 2014). This difference between on-premises ERP and Cloud ERP is associated with factors that enhance or obstruct Cloud ERP adoption. In this section, we summarize the literature on these factors, based on Peng and Gala's (2014) model, and proposed a research model for Cloud ERP adoption in healthcare (Figure 1). We built the model primarily considering Peng and Gala's (2014), because their model adequately covers the domain of research problem.

Figure 1. Cloud ERP adoption model



Factors that enhance cloud ERP adoption

Peng and Gala (2014) discusses the following four factors to enhance Cloud ERP adoption: (1) reduced cost, (2) improved system speed and performance, (3) improved system upgrade, and (4) improved mobility and accessibility.

Reduced costs

One of the benefits of Cloud ERP systems is its cost advantage over on-premises ERP systems. The literature discusses five different topics under cost advantage. Cloud ERP requires (1) less hardware investment; (2) less human resources for system operation and maintenance; (3) less start-up cost; (4) less total cost of ownership; and (5) have better cost transparency. Regarding the first topic, the research identified that on-premise ERP systems require a facility to run servers and hardware; while Cloud ERP systems do not require servers at the organization, significantly reducing these costs (Peng and Gala, 2014; Purohit et al., 2012; Castellina, 2011). Regarding the second topic, the research identified that organizations need to have costly IT professionals to operate and maintain an on-premise ERP system; while organizations reduces the need for these professionals with Cloud ERP systems (Peng and Gala, 2014; Johansson et al., 2015; Elmonem et al., 2016). Regarding the third topic, the research identified that any ERP system has certain upfront costs such as initial subscription fees and employee training. These costs are less for Cloud ERP systems (Peng and Gala, 2014; Johansson et al., 2015; Purohit et al., 2012; Elmonem et al., 2016; Castellina, 2011). Regarding the fourth topic, several research considered that the overall cost (i.e., hardware, human resources and upfront) associated with on-premise and Cloud ERP systems and found that Cloud ERP has a lower total cost of ownership (Castellina, 2011; Johansson et al., 2015). Regarding the fifth topic, the research found that the predictability of the cost is an important factor for adopting an ERP system. Relative to the on-premise ERP systems, Cloud ERP allows costs to be more transparent and predictable making it more desirable for organizations (Johansson et al., 2015; Elmonem et al., 2016). However there is research that suggest the opposite, such as there are many hidden costs of Cloud ERP (Elmonem et al., 2016; Gupta et al., 2017).

Improved system speed and performance

Another benefits of Cloud ERP systems is the improved system speed and performance. Three different topics are discussed under this factor. Cloud ERP (1) is faster to implement; (2) allows access to leading technologies and software; (3) enhances user satisfaction, efficiency, and multi-tasking. Regarding the first topic, the research provided extensive amount of evidence that on-premises ERP systems require significantly longer time to implement than Cloud ERP implementation, because Cloud ERP requires less hardware/software and employee at the focal organization (Johansson et al., 2015, Purohit et al., 2012; Elmonem et al., 2016; Elragal and El Kommos, 2012; Castellina, 2011). The faster, perhaps seamless, implementation allows organizations to achieve ERP benefits quickly. Regarding the second topic, the research reported that Cloud ERP systems allow access to the most advanced technology that fits the complex needs of customers and can be integrated with the other systems of the customers, supporting to achieve higher speed and performance (Johansson et al., 2015; Castellina, 2011). However, there are opposite claims, stating the integration can be problematic for Cloud ERP (Peng and Gala, 2014; Elmonem et al., 2016; Gupta et al., 2017). Regarding the third topic, the research explained that Cloud ERP results in higher level of satisfaction, efficiency and multi-tasking. Since the vendors use hundreds of servers and large data storage capacity, an ongoing operation (such as MRP calculation) would not slow down another ERP functionality, such as searching or printing specific products or customers. This high performance system would eventually lead to increased client satisfaction and efficiency (Peng and Gala, 2014).

Improved system upgrade

Another benefits of Cloud ERP is the improved system upgrade over on-premises ERP systems. Three different topics are discussed under this factor. Cloud ERP allows (1) centralized, automated and seamless upgrades by the vendor; (2) higher scalability and flexibility; and (3) access to the latest software and upgrades. Regarding the first topic, the research explained that Cloud ERP upgrades is centralized, meaning all the upgrades needed is managed by one party (i.e., ERP vendor), and upgrades are automated and seamless because this is vendors' core competency (Peng and Gala, 2014; Castellina, 2011). Regarding the second topic, the research found that Cloud EPR is more flexible and scalable (Peng and Gala, 2014; Johansson et al., 2015; Purohit et al., 2012; Elmonem et al., 2016; Elragal and El Kommos, 2012). Flexibility (i.e., ability to customize the ERP system) and scalability (i.e., ability to change capacity in a cost efficient manner) enables organizations to adapt rapidly and with minimal cost to the changing needs of the market and thus to compete more efficiently with other organizations (Johansson et al., 2015). Regarding the third topic, the research found that ERP vendors usually has the most advanced IT infrastructure and expertise, so they can offer the latest upgrades more frequently (Johansson et al., 2015; Elmonem et al., 2016). Further, this research explains that improved upgrading capability is particularly important for small and medium size organizations. These organizations need flexibility and scalability because of significant uncertainty in their operations. In addition, they do not have enough IT resources to constantly upgrade their systems and access to the most advanced technologies and software (Johansson et al., 2015).

Improved mobility and accessibility

Another benefit of Cloud ERP is the improved mobility and accessibility. Accessibility and mobility means that system can be accessed from anywhere through a web browser and a mobile device (Peng and Gala, 2014; Johansson et al., 2015; Purohit et al., 2012; Elmonem et al., 2016). Peng and Gala (2014) stated that accessibility and mobility is particularly important for sales personnel who may require access to the ERP system outside the organization. In addition, they stated that accessibility and mobility is important for organizations with multiple branches. Multi-site on-premise ERP installations may require additional hardware and on-site IT personnel, significantly increasing the cost. Cloud ERP systems are already mobile by design. In most cases, all it takes is a standard browser to reach company's Cloud ERP system. With Cloud ERP, the system is used by more people, more frequently in wider functional areas of business (Purohit et al., 2012; Johansson et al., 2015; Elmonem et al., 2016).

Factors that hinder Cloud ERP adoption

The literature discusses the following five factors frequently as hindering Cloud ERP adoption: (1) lack of data privacy and control, (2) lack of data security, (3) vendor lock-in, (4) integration difficulties, and (5) organizational challenges.

Lack of data privacy and control

One of the hindering factors of Cloud ERP adoption is the lack of data privacy and lack of control compared to on-premises ERP systems. Data privacy and control is about user's concerns over the vendor's not properly and cautiously handling of data and not being transparency of the data handling processes. Specifically, the literature discusses three different topics under this factor. These are: (1) lack of control in system and data storage; (2) losing technical capabilities; and (3) lack of trust in the vendor. Regarding the first issue, the research reported that lack of control can be a major issue because of inconsistent data protection laws and policies. In addition, some vendors may not be transparency of their data storage processes, such as the data being stored in different locations around the globe (Peng and Gala, 2014). Also, client concerns are reported that data in the Cloud can be exposed to the competition, get corrupt or lost, causing significant damages (Castellina, 2011; Gupta et., 2017). In addition, these concerns increased when the data is sensitive, such as patient information (Elmonem et al., 2016). Further, the research reported that the lack of control is not only about data, but also other processes such as upgrading process (Castellina, 2011), which is another concern for the organization. Regarding the second issue, Cloud ERP may result in organizations to lose technical competencies, since they become dependent on the technical capabilities of the vendor. This creates strategic risks in the long-term, since the organization may not properly evaluate the risks and take the mitigating actions related to information technology (Elmonem et al., 2016; Saa et al., 2017). Regarding the third issue, storing organization's private data in a cloud ERP system requires trust in the vendor. Lack of integrity, consistency, transparency and reliability in data processing and storage practices may lead to lack of trust by the client which may hinder Cloud ERP adoption (Gupta et al., 2017; Saa et al., 2017).

Lack of data security

Another hindering factor of Cloud ERP systems is lack of data security over on-premises ERP systems. Data security prevents unauthorized access to computers, databases and websites. The research frequently reported that lack of data security is a major concern for organizations to adopt Cloud ERP (Peng and Gala, 2014; Johansson et al., 2015; Elmonem et al., 2016; Gupta et al., 2017; Saa et al., 2017; Elragal and El Kommos, 2012; Castellina, 2011). Specifically, Peng and Gala (2014) reported that poor data protection practices may result in unauthorized data access. Further, Saa et al. (2017) stated that vendor organization can leak the information and Cloud user may not have control or security protocols and standards to prevent it. Also, smaller Cloud ERP vendors may have poor data protection measures, network monitoring and lower data management standards (Peng and Gala, 2014; Gupta et al., 2017). However, today many third party Cloud technology providers are large companies such as Amazon, Google, Microsoft and IBM, which provide extensive data security measures. Also it should be noted that the majority of data security issues are caused by human errors/actions rather than technologic weakness (Peng and Gala, 2014).

Vendor lock-in

Another risks of Cloud ERP systems is the vendor lock-in, which creates customer dependency on a vendor for the ERP services, and incurs high switching costs. There are five different topics discussed under this factor. These are: (1) high cost of ERP migration; (2) High long-term cost; (3) legal and contractual restrictions; (4) vendor's network dependence; and (5) variance and uncertainty in service quality of different vendors. Regarding the first topic, the research identified that organizations may not implement Cloud ERP because of high switching costs. Cloud ERP adoption requires a high level of commitment, time and effort for the client, leading to high switching cost. Although this is a common situation with any large software/hardware system, Cloud ERP may introduce additional difficulties to switch vendors due to distributed nature of servers, database systems, and software platforms used in the Cloud ERP systems (Peng and Gala, 2014; Johansson et al. 2015; Elmonem et al., 2016; Gupta et al., 2017). Regarding the second topic, the research identified that organizations may not implement Cloud ERP because of high long-term cost. Although start-up costs with Cloud ERP could be lower than on-premises ERP installation, annual subscription fees, and fee based upgrades could add up to significant cost figures. (Elmonem et al., 2016; Gupta et al., 2017). Regarding the third topic, the research identified that organizations may not implement Cloud ERP because of legal and contractual restrictions. The Service Level Agreement (SLA) may give the Cloud vendor additional rights to make the switch more difficult (Peng and Gala, 2014; Saa et al., 2017). Regarding the fourth topic, the research identified that organizations may not implement Cloud ERP because Cloud ERP can make the organization dependent on not only to the vendor but also vendor's network. Many Cloud ERP vendors partner with IT services providers (e.g. Amazon, IBM, Google, Microsoft) in terms of IT infrastructure and software platforms. This supply chain network may introduce additional dependencies on third parties for the client (Gupta et al., 2017; Saa et al., 2017). Regarding the fifth topic, the research identified that organizations may not implement Cloud ERP because of the variance and uncertainty in service quality of different cloud vendors. Cloud ERP is a relatively new technology compared to on-premise ERP systems, which may create uncertainties and concerns about the Cloud ERP service quality. Also, vendor-to-vendor service quality differences may be difficult to assess when selecting the right vendor, or hesitate to change the vendor (Peng and Gala, 2014).

Integration difficulties

Another risks of Cloud ERP systems is the integration difficulties. There are two topics discussed under this factor; these are: (1) customization and integration difficulties; and (2) functionality limitations. Regarding the first topic, the research identified that organizations may not implement Cloud ERP because of customization and integration difficulties (Peng and Gala, 2014; Johansson et al., 2015; Elmonem et al., 2016; Gupta et al., 2017; Castellina, 2011). Integration difficulties mostly stem from the fact that Cloud ERP systems could be incompatible with existing legacy systems at the client. Also, vendors may prohibit or limit third party software access needed for seamless integration and impose it in their service contracts. (Peng and Gala, 2014). On the other hand, in terms of functionality limitations, Cloud ERP modules offered by the vendor may not be as comprehensive or functional as the on-premises counterpart, which could lead to the use of more legacy and complementary software systems along with Cloud ERP, imposing more need for integration (Elmonem et al., 2016; Gupta et al., 2017).

Organizational challenges

Another risks of Cloud ERP systems is the organizational challenges. There are three main organizational challenges: (1) lack of top management support; (2) resistance to change; (3) inadequate understanding of Cloud ERP technology and business processes. Regarding the first topic, the literature emphasized the importance of upper level management support for Cloud ERP adoption. There are several reasons the top management may not support Cloud ERP, including financial risks, and strategic mismatch (Peng and Gala, 2014; Elmonem et al., 2016). Regarding the second topic, the research found that some organizations may be reluctant to change from a traditional system to a Cloud based system, primarily due to unwillingness for making process and cultural changes. This research mentioned that the change management could be difficult and organization may not know how to communicate the change and implications resulting from the Cloud ERP adoption (Peng and Gala, 2014; Johansson et al., 2015). Regarding the third topic, managerial lack of adequate technical knowledge on Cloud technology, its capabilities and limitations, inability to assess its benefits, costs and risks, unfamiliarity with legal, jurisdictional and compliance issues of Cloud ERP are cited by the researchers as hindrance for Cloud ERP adoption. The research also found that the organization may have difficulty understanding how to implement Cloud ERP in their internal business processes, resulting not to choose Cloud ERP (Peng and Gala, 2014; Elmonem et al., 2016; Gupta et al., 2017).

Organizational contingencies

Typical organizational contingencies, such as organizational size, industry and organizational structure (decentralized vs. centralized) could be significant factors for Cloud ERP adoption and success. The literature specifically reports that the *size* of the organization moderates the impact of motivating and hindering factors on the adoption decision (Johansson et al., 2015; Purohit et al., 2012; Elmonem et al., 2016; Gupta et al., 2017; Saa et al., 2017; Elragal and El Kommos, 2012). As shown earlier, the research explains that improved upgrading capability and scalability are particularly important for small and medium sized organizations. Since these organizations may grow fast and receive volatile demand, need for flexibility, scalability and upgradability are important features expected from an ERP system. As an additional challenge, they may not have enough IT resources to upgrade on a continuous basis, and access to the latest technologies and software. Therefore, Cloud ERP could be the most viable option for them to have the state of the art technology (Johansson et al., 2015). Also, while small and medium sized organizations could be more lenient on Cloud ERP adoption due to lower upfront costs and faster implementation times, a larger organization with a better financial standing could approach Cloud ERP with less enthusiasm (Castellina, 2011;Saa et al., 2017; Elragal and El Kommos, 2012).

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE
Open Source Software Communities (OSSPCs) as a Good Solution for the Productivity Dilemma: An Empirical Research Proposal

Zhengzhong Shi
University of Massachusetts at Dartmouth
Email: zshi@umassd.edu

ABSTRACT

Productivity dilemma is about the contradiction between pursuing efficiency and driving innovations and it is about the inconsistent requirements in terms of organizational systems, structures, behaviors by exploitation and exploration activities. Organizational ambidexterity as the balance between exploitation and exploration activities has been proposed as the solution to resolve this dilemma. Both structural and contextual factors impacting ambidexterity have been identified and ambidexterity's impacts on organizational performance has been tested. This rich stream of research on productivity dilemma and ambidexterity in organizations can be applied to explain why OSSPCs could resolve the productivity dilemma and consequently become very successful.

KEYWORDS: OSS Project Community, Productivity Dilemma, Ambidexterity, and Research Proposal

INTRODUCTION

Based on data in the automobile industry, Abernathy (1978) found that a firm's decline is highly related to its efficiency and productivity efforts and raised the issue of productivity dilemma. He argued that for firms to gain sustained competitive advantages overtime, it is imperative to be both efficient/productive and innovative/flexible. While to improve the efficiency of a production system is to exploit existing capabilities, to be innovative and flexible is to explore new capabilities. March (1991) proposed that there is a need for the balance between exploitation and exploration for firms to be successful both in the short term and in the long run. This argument is further developed in the research stream of dynamic capabilities by Teece, Pisano, and Shuen 1997). In addition, Benner and Tushman (2003) proposed that ambidextrous organizations can help with meeting this requirement for the balance between exploitation and exploration activities. In their reflection of the 2003 paper, Benner and Tushman (2015) further specified the need for the combination of both the structural ambidexterity and contextual ambidexterity to better manage this balance between exploitation and exploration. Structural ambidexterity can be produced by firms to have both organic and mechanic units that are integrated at the senior manager level. Contextual ambidexterity can be produced by firms through encouraging employees making their own decisions regarding time and efforts dedicated to efficiency or innovation activities (Gibson and Birkinshaw, 2004).

Open source software project communities (OSSPCs) have been playing a critical role in software development for several decades and good examples are such as the Linux community, open source enterprise systems communities, and many communities supporting Apache foundation's open source software projects (e.g., Spark, Kafka, Hadoop, etc.). In this paper, we propose that the structure of the OSSPC management and the flexibility of community members in terms of freely deciding the tasks they would like to work on and the tools they may use demonstrate an almost perfect integration of the structural ambidexterity and contextual ambidexterity. And this integration,, in addition to other factors such as community trust dynamics, may very well help explain the underlying reasons why OSSPCs can be very successful in software development, facilitating both incremental innovations such as bug fixes and discontinuous innovations such as new feature development and change of technology trajectories for software development. In this paper, we will propose a research plan to empirically test the open source software community ambidexterity hypothesis.

THE CONCEPT OF AMBIDEXTERITY

The concept of ambidexterity is initially developed in the context of organizations, which means that while firms exploit existing competences, knowledge, and assets for efficiency, refinement, cost leadership, and continuous improvement, they also explore new fields through discovery, global search, disruptive innovations, and new competence building. While the purpose of exploitation is to make full use of existing knowledge and assets for the short-term oriented profit maximization, the goal of exploration is to ensure their long-term viability and sustained performance, especially during the period of industry revolution when existing dominant product designs are overthrown and new emerging designs are competing to dominate in the industry.

In early studies of ambidexterity, scholars proposed that it is very challenging for organizations to simultaneously maintain effective exploration and efficient exploitation due to the competition for scarce firm resources and the embedded conflicts between exploration and exploitation in terms of their demands for different organizational structures, processes, systems, and cultures (March, 1991; Sidhu, Commandeur, and Volberda. 2007). While to exploit existing routines, assets, and knowledge requires discipline, tight coupling, and continuous process improvement, to explore new fields effectively requires flexibility, loose coupling, and process transformation. Consequently, without establishing appropriate organizational contexts and structural designs, exploitation and exploration may both suffer.

Later studies demonstrate that while simultaneous implementation of exploitation and exploration strategies is indeed very challenging and replete with barriers, it is possible for firms to be ambidextrous through balancing exploitation and exploration and some empirical evidences show that firms properly matching exploitation with exploration perform better in the long run and have a high probability to survive disruptive technological innovations (Gibson and Birkinshaw, 2004). Cao, Gedajlovic, and Zhang (2009) further developed measurements of ambidexterity (including both the balanced and combined dimensions) and provided empirical evidence proving the positive impact on organizational performance of the balanced and combined dimensions of ambidexterity and their interactions.

Further, Porter’s (1996) development of the concept of productivity frontier also provides the theoretical foundation for the feasibility of simultaneous exploitation and exploration. In his theory, he argued that before companies reach the productivity frontier, “they can often improve on multiple dimensions of performance at the same time. … What were once believed to be real trade-offs—between defects and costs, for example—turned out to be illusions created by poor operational effectiveness” (Page 62, Porter, 1996). Exploration is risky and costly, but it is the foundation for the long-term viability in case of industry discontinuous innovations. Exploitation is to generate continuous adjustments for incremental innovations with low risk and investments and it is good for benefiting from existing capabilities. Before companies reach their productivity frontiers, based on Porter’s theory, they can move forward along both the exploitative and explorative paths within the constraints of resource limitation. With today’s dramatic technology advancement (especially in Internet and cloud computing supported technologies such as AI and machine learning), the productivity frontier is pushed further outward. Thus, companies may be able to better simultaneously handle resource competition and structural challenges raised by exploitation and exploration activities with well nurtured ambidexterity.

THE AMBIDEXTERITY OF OSSPCS

The concept of ambidexterity could be especially useful in the context of open source software project communities (OSSPCs). For example, Temizkan and Kumar (2015) applied social network and organizational ambidexterity theories to the field of OSS development and found that while exploitation oriented teams (such as patch development teams) consist of a smaller number of densely connected subgroups and emphasize close collaborations, exploration oriented teams (such as feature request and development teams) consist of a larger number of sparsely connected subgroups and emphasize the need to access new ideas. They further proposed that the identification of ambidextrous developers in an open source software community helps develop the concept of OSSPC ambidexterity. While it is true that an OSSPC cannot be treated as an organization with strict employer-employee relationships, community members can be viewed as forming a virtual organization through sharing common values and beliefs (Stewart and Gosain, 2006), participating community activities by following joining scripts (von Krogh, Spaeth, and Lakhani, 2003), and establishing appropriate structural conditions (Lee and Cole, 2003). For example, in the study by Lee and Cole (2003) about the OSSPC for the Linux development, they identified two tiers of community members (core vs. periphery) and parallel code branches (development branch vs. stable branch) as community organizing structures resulting in structural ambidexterity.

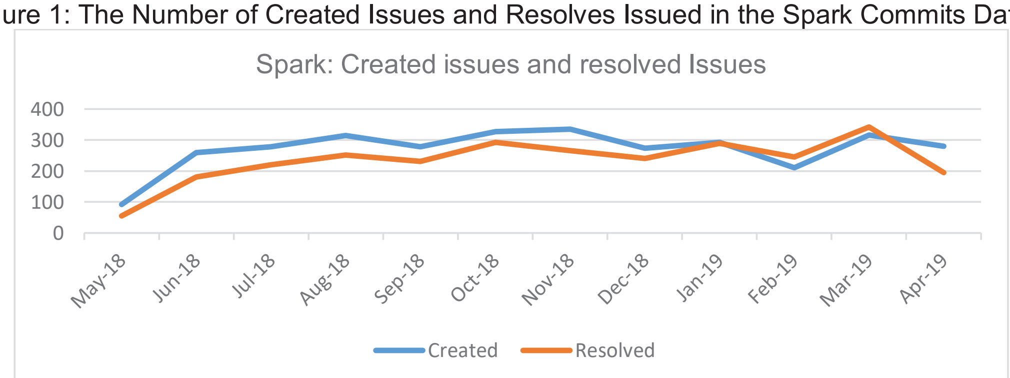
These community structures are not separate entities like mechanic and organic business units in a company, which are integrated at the senior managers' level. While employees in organizations can change their positions, they are not as flexible in switching their focus between exploitative and explorative activities beyond assignments from managers as community members who can switch between exploitation and exploration activities freely without consultation with any other members. In addition, members themselves decide the problems they want to solve, choose the tools to solve those problems, and control the pace of problem solving. Indeed, it is our belief that this behavior pattern of freely switching between exploitative and explorative activities is the key factor resulting in the contextual ambidexterity proposed by Gibson and Birkinshaw (2004) in the context of OSSPCs. Gibson and Birkinshaw believe that contextual ambidexterity differs from structural ambidexterity remarkably in that the former is achieved by building a set of systems and processes enabling individuals make decisions about their time and energy allocation for alignment and adaptability. Their empirical testing further demonstrated the critical role contextual ambidexterity plays in enhancing firm performance. For the four groups of firms in their data, they found that the highly ambidextrous firms perform the best and the moderately ambidextrous firms have the second-best performance. Following their conceptual development and empirical findings, it is our belief that the OSSPC contexts encourage simultaneous exploitation and exploration, which results in the OSSPC contextual ambidexterity.

As Benner and Tushman (2015, page 12) proposed, "that structural and contextual ambidexterity are complementary". The goal of this paper is to apply the rich findings from the research stream in organization ambidexterity to the field of OSSPC and empirically demonstrate that the OSSPCs are indeed a good solution for the productivity dilemma raised by Abernathy (1978) through the integration of structural and contextual ambidexterity.

A RESEARCH PROPOSAL FOR THE EMPIRICAL STUDY

In this section, we will sketch out an outline for the future empirical study on whether OSSPCs could be a good solution for the productivity dilemma from the community ambidexterity perspective. First, we plan to use the open source software project commits as the source of data for the empirical study. One example OSSPC is Apache Spark™ community. Spark "is a unified analytics engine for large-scale data processing. Apache Spark achieves high performance for both batch and streaming data, using a state-of-the-art DAG scheduler, a query optimizer, and a physical execution engine" [https://spark.apache.org/, accessed on May 22, 2019]. The Spark's commits data are hosted in the Jira site [https://issues.apache.org/jira/projects/SPARK/, accessed on May 22, 2019] and the Spark community is very active. As indicated in the following figure, for the past year from May 2018 to April 2019, the Spark community created 3256 issues and resolved 2807 issues (Figure 1).

Figure 1: The Number of Created Issues and Resolves Issued in the Spark Commits Data



In terms of the different types of tasks facing the Spark community, the following list shows the details.

1. Epic - Created by JIRA Software - do not edit or delete. Issue type for a big user story that needs to be broken down.
2. Bug - A problem which impairs or prevents the functions of the product.
3. Dependency upgrade - Upgrading a dependency to a newer version
4. Improvement - An improvement or enhancement to an existing feature or task.
5. New Feature - A new feature of the product, which has yet to be developed.
6. Story - Created by JIRA Software - do not edit or delete. Issue type for a user story
7. Task - A task that needs to be done.
8. Sub-task - The sub-task of the issue

For the purpose of investigating the OSSPC ambidexterity in our research, we are interested in the bug, improvement, and new feature tasks. Bugs and improvement tasks represent incremental innovations with the focus on exploitation of existing software capabilities. New features represent the efforts of exploring new capabilities for the software.

Structurally, for the Spark community, besides the master branch of the source codes, there are two active source code branches (2.3 and 2.4) [https://github.com/apache/spark/branches, accessed on May 22, 2019]. These different branches are to store the source codes for different versions, some of which are stable for the production use with focus on exploiting existing software capabilities and some of which are for the explorative purpose with experimented features and no or few production-based tests. Contextually, community members can freely make their own decisions in terms of which branch of source codes they would like to work on and which type of tasks they would like to complete. The aggregated member behaviors of the whole community represent the organic distribution of efforts and time dedicated to exploitative and explorative activities respectively, forming the OSSPC ambidexterity resulting from both the contextual and structural factors embedded in the Spark community.

Second, we would like to divide the available bug, improvement, and new feature commit data into different periods. Each period could be two or four or six months as the community may have its own developmental cycles. Then we will count the exploitative activities and the explorative activities in each period. Further, based on Cao, Gedajlovic and Zhang (2009), we will also calculate the balanced and combined dimensions of the ambidexterity for each period and examine the evolution of the Spark's community ambidexterity. In addition, we plan to collect the web exposure data of the Spark project (primarily using google search) and its downloads for each period to examine the impact of community ambidexterity on its popularity and downloads over time. We assume that the web popularity and downloads for each period can be used to approximate the market share of the Spark project in the big data analytics and AI market.

Third, we will collect more commits data from other OSSPCs and do a comparison study in terms of the evolution of the community ambidexterity and their impacts on web popularity and the number of downloads per period. Other potential OSSPCs we may include in the future empirical research are such as the Dolibarr ERP/CRM community and the Vtiger CRM community, both of which have very active members.

CONCLUSION

In this paper, we review the concept of productivity dilemma and the proposed ambidexterity as the solution to this dilemma in the context of organizations. We notice that over time, researchers agree that structural and contextual ambidexterity are complementary, and both are needed to better resolve the challenge of the productivity dilemma. We conjecture that the structural design and the contextual factors embedded in OSSPCs may result in community ambidexterity resolving the productivity dilemma. On the one hand, there are multiple source code branches with the stable branches for the production purpose and other unstable branches for the development purpose. These different branches automatically organize the community members into groups working on commits for exploitative and explorative activities and this structural design is the driver for the structural ambidexterity in the OSSPCs. On the other hand, the freedom of the community members in deciding the problems and tasks in different source code branches they would like to work on over time is the key contextual factor enabling contextual ambidexterity. Lastly, we sketch out an outline for the empirical study in terms of commits data collection and analysis. We believe that this paper, through drafting a research proposal on the OSSPCs as a good solution to the productivity dilemma from the community ambidexterity perspective, will lead to fruitful research studies.

REFERENCES

Abernathy, W. J. (1978). The productivity dilemma. Baltimore: Johns Hopkins University Press.

Benner, M. J., & Tushman, M. (2002). Process management and technological innovation: A longitudinal study of the photography and paint industries. Administrative Science Quarterly, 47(4), 676-707.

Benner, M. J., & Tushman, M. L. (2003). Exploitation, exploration, and process management: The productivity dilemma revisited. Academy of management review, 28(2), 238-256.

Benner, M. J., & Tushman, M. L. (2015). Reflections on the 2013 Decade Award—“Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited” Ten Years Later. Academy of Management Review, 40(4), 497-514.

Cao, Qing, Eric Gedajlovic, and Hongping Zhang. (2009) "Unpacking organizational ambidexterity: Dimensions, contingencies, and synergistic effects." Organization Science 20.4, 781-796.

Gibson, Cristina B., and Julian Birkinshaw. (2004) "The antecedents, consequences, and mediating role of organizational ambidexterity." Academy of management Journal 47, no. 2, 209-226.

Lee, G.K. and Cole, R.E., (2003). From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. Organization science, 14(6), pp.633-649.

March, James G. (1991) "Exploration and exploitation in organizational learning." Organization science 2.1, 71-87.

Porter, M. E. (1996). What is strategy? Harvard Business Review, 74(6): 61– 81.
Stewart, K.J. and Gosain, S., (2006). The impact of ideology on effectiveness in open source software development teams. Mis Quarterly, pp.291-314.

Temizkan, Orcun, and Ram L. Kumar. (2015) "Exploitation and Exploration Networks in Open Source Software Development: An Artifact-Level Analysis." Journal of Management Information Systems 32.1,116-150.

Von Krogh, G., Spaeth, S. and Lakhani, K.R., (2003). Community, joining, and specialization in open source software innovation: a case study. Research policy, 32(7), pp.1217-1241.

Sidhu, J. S., H. R. Commandeur, H. W. Volberda. (2007). The multifaceted nature of exploration and exploitation: Value of supply, demand, and spatial search for innovation. Organ. Sci. 18 20-40.

DECISION SCIENCES INSTITUTE

The Role of Working Relationships in SCM Performance

Liliana Rivera
School of Management Universidad del Rosario, MIT Center for Transportation and Logistics
Email: myriam.rivera@urosario.edu.co mlrivera@mit.edu

Andrea S. Patrucco
Penn State University, Department of Project and Supply Chain Management
Email: asp72@psu.edu

Yossi Sheffi
MIT Center for Transportation and Logistics
Email: sheffi@mit.edu

Christopher Mejia Argueta
MIT Center for Transportation and Logistics
Email: cmejia@mit.edu

ABSTRACT

This paper evaluates the relationships between support networks (social capital) and skills (human capital) on both employees' job satisfaction and companies' expansion of their supply chain (SC). A structural equation model (SEM) is tested based on four constructs: relationship orientation, job skills, job satisfaction, and supply chain expansion. To do this, we conducted a survey of Colombian supply chain professionals and managers. The results reaffirm the literature's previous findings of a strong relationship between employees' social capital and their job skills. We also found a significant relationship among SC members' skills and job satisfaction. SC expansion determinants were not conclusive, probably because other factors affected the expansion of companies; however, the research overall highlights the indirect importance of human resources management on supply chain performance.

KEYWORDS: Human Resource Management; Skills; Satisfaction; Latin America; Performance

INTRODUCTION

As the importance and influence of supply chain functions grow in all businesses, the importance of the supply chain's human resources is becoming evident (Barnes and Liao, 2013). Long-term, high-performance supply chain management (SCM) depends, in part, on human capital and data-driven decision-making processes (Barnes and Liao, 2012; Waller and Fawcett, 2013). However, scant research is available to characterize the influence of multiple factors on organizational performance and only a handful of papers link Human Resource management (HRM) to overall supply chain performance (Gunasekaran et al. 2004; Ou et al. 2010; Smith–Doerflein et al. 2011; Swart et al. 2012; Lengnick-Hall et al. 2013; Gómez-Cedeño et al. 2015).

The available HRM studies in SCM have either focused on deepening the understanding of diverse job skills or on mapping social relationships to see their effects on human and organizational performance (Barnes and Liao, 2012). Yet only a couple of these studies propose metrics related to this overall performance. Furthermore, little research has examined the impact of the job satisfaction of SCM professionals on SC performance (Antoncic and Antoncic, 2012).

To address this gap in the literature, this paper evaluates the (somewhat obvious) hypothesis that in order to deliver high-quality organizational SC performance, supply chain employees must possess a sufficient level of skills, an adequate network supporting them, and satisfaction with their jobs. Even though this research does not establish causality, its aim is to contribute to setting the foundations of how to develop future talent in SCM in order to respond to challenges in the discipline, such as rapidly changing environments, fragmented logistics operations, growing urbanization rates and dynamic data-driven decision-making processes (Rivera at al., 2019). These contextual features are present in emerging regions such as Latin America and countries such as Colombia. Thus, our contribution targets a better understanding of the role that high-skilled SCM professionals should have in transferring skills, to make SCM a more active member of the talent development process via social relationships and employees' job satisfaction, and to promote organizational and SC growth.

Therefore, some unique aspects that distinguish our study from the existing body of similar works in the literature involve the inclusion of employee's skills, satisfaction, and social networking in order to model, analyze, and derive conclusions about their performance in SCM. Hence, this paper contributes to the literature 1) empirically, by providing new insights about the aforementioned constructs of HRM in a Latin American country through primary data collected via surveys, and 2) methodologically, by explicitly modeling the relationships among these constructs and their effect on overall supply chain performance. Our model differs from (and builds onto) the existing body of traditional structural equation models (SEM) in that it enables us to first evaluate the relationships that have a significant effect on skills development and how these skills influence employee's satisfaction and supply chain performance.

Section 2 reviews the theoretical background, explains the relationship between HRM and SCM, and provides some context for the HRM in Latin America. Section 3 presents the research model and hypotheses. Section 4 depicts the methodology, questionnaire design and data collection. Section 5 presents the results, including Confirmatory Factor Analysis (CFA), SEM and cluster analysis. Finally, Section 6 discusses the results and Section 7 concludes with final observations.

LITERATURE REVIEW

Supply chain management and logistics are topics that have gained traction in the academic and practitioner communities in the last two to three decades. In fact, SCM has shifted from having a secondary role in organizations, to becoming a core strategic function that builds a competitive advantage among diverse stakeholders (i.e., suppliers, manufacturers, retailers, distributors and end customers) (Mentzer 2001; Gunasekaran et al. 2004; Hult et al. 2007). In consequence, the characteristics required of supply chain personnel have also evolved dramatically (Slone et al. 2007; Flöthmann and Hoberg, 2017). This evolution has also affected the way in which companies perform HRM for the purpose of recruiting, training, and retaining superior SCM professionals to improve the performance of organizations and grow their productivity and profitability (Swart et al. 2012).

Human resources have proven to be a key factor for the undertaking of successful supply chain activities, ensuring greater customer satisfaction and enhanced organizational performance (Menon, 2012). In addition, the role played by individuals is critical for the execution of SCM and its complex web of relationships. Hence, HR practices in firms must be aligned with SCM in order to promote the integration of the chain and safeguard better business outcomes. (Gómez-Cedaño et al. 2015). However, few studies in the literature describe the relationship between human resource activities or organization variables and supply chain success. This omission may be due to the emphasis on technical or “hard” determinants of supply chain performance (Shub and Stonebraker, 2009).

On the other hand, there is an ongoing shortage of SC managers who have the broad range of competencies and managerial skills needed to promote operational innovation (Jordan and Bak, 2016). Moreover, SCM personnel may not be sufficiently aligned with the objective of developing the “soft” side of SC managers. Perhaps the reason behind it all is that organizational development (OD) is often left to organizational leaders and managers who are tasked with such responsibilities but may not be trained to execute them. (Ellinger and Ellinger, 2014). Regarding the lack of research that examines the so-called “soft side” of SCM, more research in this area would help organizations to better manage strategically significant supply chain processes with the adequate connection between the needs of the chain, the SCM personnel and the responsibilities asked of them (Ellinger and Ellinger, 2014; Othman and Ghani, 2008).

HRM in Latin America

Despite the noted progress of Latin America in terms of access to education and healthcare, as well as increased prosperity and decreasing poverty rates during the last decade, the region still needs further improvement (UNDP, 2017). There is still a remarkable lack of high-quality human capital, which hinders regional economic growth and obliges firms to work with a large base of low-skilled workers in various productive sectors (Santos, 2016). Even worse, rapid technological pace threatens to broaden the existing gap between the skills required by regional labor markets and those available from regional human talent. This further polarizes the socioeconomic situation in developing countries, where most of the labor is trained for mechanical and manual tasks to develop a low-cost workforce, at the same time limiting the professionalization in technology, data-driven decision-making processes, critical thinking, and non-routine cognitive skills (Elvira and Davila, 2005; Perez-Arrau et al. 2012). Further trends, such as growing urbanization and poor policymaking, shape labor markets and in some cases are likely to be even more important than technology (Santos, 2016).

Elvira and Davila (2005) is, to our knowledge, the first study to describe the state-of-the-art of HRM in Latin America. Since that pioneering work, solid constructs related to HRM in SCM such as social relationships, job satisfaction, and personal contact were identified to be important manifestations of the Latin American working environment. Later, Perez-Arrau et al. (2012) showed that HRM depends on cultural and country-related characteristics that affect human and organizational performance. Despite the authors' focus on Chile, their study updates the state-of-the-art of HRM in Latin America, without analyzing particularities and implications of SCM professionals. Therefore, even with the growing efforts to address HRM in Latin America, there are still several research opportunities available.

HRM and SCM: which relationships?

It is clear that HRM has a deep impact in configuration of SCM professionals and workforce, which in turn has a significant consequence for organizational and company performance (Gómez-Cedeño et al. 2015). But there is scarce research connecting HR skills and SCM performance, especially in Latin America. By looking at recent literature (e.g., Mangan and Christopher, 2005; Murphy and Poist, 2007; Derwik and Hellström, 2017; Flöthmann et al. 2018), four important constructs that allow for mapping and building a detailed model of how SCM professionals and logistics managers might affect organizational and supply chain performance emerge. Thus, it is important to investigate the effects of multiple HR factors from a twofold perspective: 1) from a typical HRM perspective, by analyzing *relationship orientation* (i.e., social connections and networking), employees' satisfaction and their work skills; and, 2) from a SCM perspective, by linking employee features (e.g., experience, technical training) to supply chain expansion and organizational growth (Brauner et al. 2013).

Relationship orientation. Social relationships reflect the connective power of networking for the employee. Such relationships can be in the form of organizational collaboration, camaraderie, professional networking, and friendship that facilitate personal communications, knowledge co-creation, knowledge sharing, and empathy among organizational members (Elvira and Davila, 2005; Hitt and Ireland, 2007; McCallum and O'Connell, 2009, Wood et al. 2016). This idea relies on systems where the party and the state are intertwined and let individuals facilitate access to broader and better sources of information; develop influence, control and power to achieve goals; increase solidarity; and reduce the need for formal controls (Adler and Kwon, 2002). In particular, it measures the extent to which individuals have established connections to potentially influential people, such as leaders, policy makers, and decision makers (Barnes and Liao, 2012).

Employee-related features. Hohenstein et al.'s (2014) recent extensive literature review on human capital management in SCM and Flöthmann et al.'s (2018) study of the key competencies of SCM personnel can be considered as a benchmark when dealing with the level of job *skills*. In particular, their framework suggests the inclusion of four categories of skills: 1) technical expertise; 2) general management skills; 3) interpersonal skills; and 4) analytical and problem-solving abilities. These categories are defined by the employee's socioeconomic profile, education, experience, skills, values and abilities that increase their value in the workplace (Elvira and Davila, 2005; Forret, 2006; Murphy and Poist, 2007; Derwik and Hellström, 2017).

Employee's satisfaction. In line with several past (Wanous et al. 1997) and more recent (Dobrow and Ganzach, 2018) studies, employees will more effectively contribute to organizational (and SCM) success when they are satisfied with their work status, both in terms of its economic and non-economic aspects. The factors contributing to job satisfaction include working environment, flexible job descriptions, team work, and rewards that lead to the well-being and high performance of the personnel (Menon, 2012; Goffnett et al. 2012). Humphrey et al. (2007) used a large meta-analytical review to show that job satisfaction depends heavily on motivational-human capital characteristics. Moreover, Antoncic and Antoncic (2011) found that job satisfaction characteristics significantly influence a firm's growth.

Organizational/supply chain expansion. Finally, the growth potential expresses the traditional dimension to measure supply chain coverage (van Hoek, 2001), and it is linked to the characteristics of people (Fawcett and Waller, 2013). Since the diffusion of the SCM culture, authors have debated the need for SCM and whether supply chain evolution is strictly related to the people operating in it (Chapman et al., 2000); in this regard, Gunasekaran et al. (2004) proposed a set of performance metrics to measure SCM activities' growth. The authors linked SCM activities – plan, source, make/assemble and deliver – at different levels with the aim to create flexibility, effectiveness, accuracy and speed in the flow of goods/services, information, cash, and knowledge along the supply chain. Besides this focus, the current literature has little evidence on the connection between the SC expansion and the characteristics of SC professionals.

The next section proposes a research model that is grounded in the above theoretical considerations. Three constructs are considered and three related hypotheses are proposed for the relationships among them. This model may serve to increase the understanding of the relationships among HRM-related activities and SCM performance.

THEORETICAL DEVELOPMENT/MODEL

The need to select highly competent supply chain employees is grounded in the knowledge-based view (KBV) of the company (Grant, 1996). This theory proposes the idea that knowledge and capabilities may become a source of competitive advantage that supply chain management can exploit (Handfield et al. 2105). Strategic knowledge, in particular, can be rooted both in the individuals and in the organizations (McCallum and O'Connell, 2009; Wood et al. 2016).

Personal knowledge can be seen as the combination of individual skills and abilities possessed by an employee and applied in day-to-day job activities (Fossum and Arvey, 1986). As cited in the HRM literature (Wright et al. 2001; Kang and Snell, 2009), highly skilled employees can be a strategic resource that facilitates competitive advantage to a firm. Furthermore, these skills and the expertise accumulated by employees over time become a source of sustained differentiation, especially, if the skills are valuable, rare, inimitable, and non-substitutable (Barney, 1991).

This is particularly true in the context of SCM. SCM has become a profession that requires a mixed of technical and managerial skills. These skills need to be cultivated by adopting a multidisciplinary and cross-functional perspective, in order to link numerous functions within and across companies that manage an integrated value flow (Flöthman et al. 2018).

Therefore, this paper is grounded in the idea that in order to deliver high-quality performance and help companies expand and grow, supply chain employees need to possess a sufficient level of skills, a broad network supporting them, and be satisfied with their jobs. To analyze these linkages, the theoretical model illustrated in *Figure 1* has been tested.

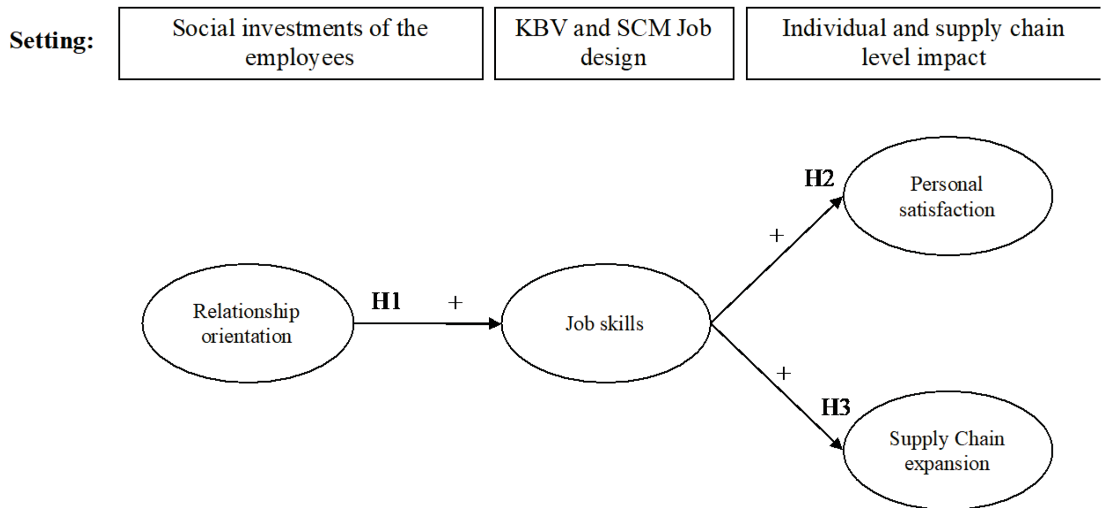


Figure 1. Research model

When we talk about competence and skills, we adopt the perspectives provided by McCormick (1976) as well as Fossum and Arvey (1986), who define skills as *developed abilities that facilitate learning or the more rapid acquisition of knowledge*. These skills influence performance and they are not subject to a big change.

An important property of skills is that they can be trained or acquired "on the job" and that several elements can contribute to enhancing them (Mumford et al., 1999). Under social exchange relationships theory, the establishment of an extensive network of relationships promotes interpersonal exchange and individual growth, which may affect skills and knowledge improvement (Möller and Halinen, 1999; Baron and Markman, 2000).

Among the different possible external relationships for an employee – those involving people who have a relevant role and/or qualifications for the employee's context (i.e., managers, executives, policy makers) – are recognized as being the most valuable in fostering individual skills development (Barnes and Liao, 2012; Dowell et al. 2015; Schermuly and Meyer, 2016). This is even more important in supply chain management, where the possibility to build relationships and grow networking (i.e., connection power) is amplified by the complexity of the unit of analysis. This puts employees in contact with several supply chain actors such as suppliers, distributors, and retailers at diverse levels (Wright and Kaine, 2015).

Based on these reasons, we can thus formulate the following hypothesis:

H1: In SCM, a greater orientation toward employees having external business relationships is associated with a greater level of job skills

By developing skilled workers and investing in human resources, companies can obtain payback on two levels: employee satisfaction and organizational performance. If we examine the HRM literature, there is a broad set of scholars (Messersmith et al. 2011; Kehoe and Wright, 2013; Kianto et al. 2016) who discussed that, in companies pushing for development of employee competencies – in line with job requirements and the organizational environment – the employee's personal job satisfaction is higher.

SCM jobs require multifaceted skills to manage flexibility, change, risk, complex decision-making processes, creative thinking, and technical knowledge. SCM positions give employees the

opportunity to express their capabilities in a more extensive way, as well as to improve and demonstrate their abilities. Companies also reward workers for the complex tasks and challenges they have to face by providing them higher benefits in terms of salary, non-economic privileges, work-life balance, and better work environment (Lai Wan, 2007). Altogether, this results in higher satisfaction of the employees in their jobs, maximizing the company's ability to retain the best performers and develop high-potential employees to become the next generation of decision makers (Myers et al. 2004).
We can thus formulate the following hypothesis:

H2: In SCM, a higher level of job skills is associated with a higher level of employee satisfaction

Finally, the human resources department at any organization should focus on recruiting, investing in, and developing talented supply chain professionals to maximize their satisfaction and the company's talent retention. The latter creates a crucial, long-term competitive advantage because relying on a high-skilled employee base has been linked positively to overall improvement of organizational performance (Chowhan, 2016).
Specifically focusing on the SCM literature instead, several authors have promoted the idea of a linkage between the skills of SCM employees and the growth of the supply chain (Fisher et al. 2010; Mendes and Machado, 2015; Huo et al. 2016; Flöthman et al. 2018). Therefore, skilled SCM professionals might enable improvement of supply chains, in terms of the market served and the global reach of supply and distribution networks (Hohenstein et al. 2014; Jordan and Bak, 2016). We can thus formulate the following hypothesis:

H3: In SCM, a greater level of job skills is associated with a greater level of supply chain expansion

METHODOLOGY

To investigate the research model, a survey-based methodology was used, being suitable for testing purposes (Choi et al. 2016). The hypotheses were tested through a questionnaire that collected information about human resources management, employee and SC characteristics, in multiple main cities of Colombia during 2018 and 2019.

Questionnaire design and scale development

The questionnaire collected data from SC professionals pertaining to the four areas included in the model (i.e. the extent of networks of relationships, types of job skills and capabilities, job satisfaction, and company growth). We also collected socioeconomic data and years of experience from the surveyed professionals. In addition, we asked respondents for values and principles followed at their positions, as well as the employee trainings and course offerings at their organizations.
The survey items used to measure the theoretical constructs were derived from an extensive literature review about HRM, according to which the questionnaire was designed. Based on these constructs, we formulated diverse questions to form a large pool of items. Also, we compared these constructs in order to fill the two identified literature gaps (i.e., the existing link between the skills required by regional labor markets and those available from regional human talent; and the relationship between human capital characteristics and SC performance), and we added questions to the survey to fill these gaps. We piloted a preliminary version of the survey that contained 35 questions, which was reduced to 21 items for the sake of respondent time,

uniqueness, and significance of the results. The pre-test allowed us to improve the items of each dimension, alter the five-point Likert scale, and design a better survey. Following the suggestions in literature to reduce non-response, we carefully designed the survey instrument, established the research significance, and balanced survey length (Rogelberg and Stanton, 2007). The final version of the questionnaire in Spanish is presented in Appendix 1.
Table 1 summarizes the main items adopted for each construct and their operationalization in the survey.

Table 1. Operationalization of constructs

First-order construct	Indicators	Label	Scale
Relationship orientation	I have relationships with entrepreneurs in successful start-ups	REL 1	1 = “no relationships” 5 = “with more than 15 people”
	I have relationships with executive managers in medium companies	REL 2	
	I have relationships with decision makers in large firms	REL 3	
Job skills	I have leadership capabilities	SKILL 1	1 = “not at all” 5 = to a large extent
	I have goal orientation	SKILL 2	
	I can manage efficient decision-making and problem-solving processes	SKILL 3	
	I have quantitative skills	SKILL 4	
Personal satisfaction	I am satisfied with my current salary	SAT 1	
	I am satisfied with my non-economic incentives	SAT 2	
	I am satisfied with my work-life balance	SAT 3	
Supply chain expansion	In the last 10 years, my company has changed the number of employees	EXP 1	1 = “0%” 5 = “more than 15%”
	In the last 10 years, my company has changed the number of products or production lines	EXP 2	
	In the last 10 years, my company has changed the number of suppliers	EXP 3	

Data collection and sample characteristics

Logistics professionals face a double challenge because their role encompasses both the diverse functional departments of their company as well as coordinating similar tasks with logistics roles at external supply chain stakeholders (e.g., suppliers, customers, and carriers). Therefore, there are strong interdependencies between SCM and human resources (HR) that require an integrative, cross-functional, multi-company vision and a developed HR ability to create levers to avoid imposing activities and hindering the growth of certain actors of the supply chain (Jurčević et al. 2009).
However, there is still a lack of knowledge and research in this complex topic (Jurčević et al. 2009). Moreover, a few authors (Brauner et al. 2013) argue that incorrect decisions in HR-related

realms lead to lower performance not only for the company, but also for other actors of the supply chain. Furthermore, Latin America faces multiple difficult challenges, such as poor infrastructure, ineffective logistics operations, income inequality, and social and political instability (Velazquez-Martinez et al. 2018). Moreover, Latin America lacks a strong educational system that instills the required abilities in its population that logistics job positions nowadays require.

Developing countries display labor market polarization in which a large number of workers perform manual labor, while in developed regions these tasks are largely handled more efficiently by automated processes and technology. This pattern prevails in most of Latin America as well, due to the low skill levels of workers that makes their jobs lower-paying and non-technologically dependent (Perez-Arrau et al. 2012; Santos, 2016). Unfortunately, this hinders the (macro- and micro-) economic growth of the companies and the country in the long term. Our research picks a Latin American country as a case study and investigates how multiple skill profiles foster the growth of supply chains and the development of companies. This research is relevant to closing the existing research gap in whether SCM professionals might enable supply chain expansion of a company through a better skillset and the relationship network they develop through their jobs.

We conducted the survey at six events organized by LOGYCA, a Colombian organization that administers several international standards, such as barcode, and offers products and services in logistics for over 650,000 different companies across Latin America. These events were in five main cities of Colombia, including the largest metropolitan areas, and were attended by a total of 180 people. The profile of attendees to these events were primarily professionals in logistics, transportation, and supply chain management from service and manufacturing industries. We did not select specific industries because we wanted to guarantee heterogeneity in the results, but we did concentrate on medium and large firms to understand the effect of constructs such as the relationship orientation, as it is in these type of organization where HRM practices are more consolidated and employees have the opportunity to create stronger network of relationships (Blanco-Alcántara et al., 2018). Also, larger firms have the opportunity to employ a larger number and variety of people and to represent a less fragmented, more formal labor market.

At each event, we delivered the survey to the registration table and asked attendees to fill out the survey during the event. Clear instructions were written in the printed version of the survey, and staff from LOGYCA answered questions that the respondents had about the survey. To increase the validity of findings, we opted for a multi-city sample in Colombia. Naturally, larger cities are more oriented to SC activities because urban logistics are more complex and because they are nuclei of economic activities for Colombia. Therefore, this heterogeneity in profile also benefited the results of the survey.

In light of this, we performed the survey in 2018 and 2019, collecting data from 144 questionnaires. After removing responses with missing values on critical items, we obtained a final sample of 131 usable responses, which corresponds to a raw response rate of 73%.

Table 2 reports the relevant descriptive statistics of the sample.

Table 2*.* Sample descriptive

Descriptive	Freq.	%	Descriptive	Freq.	%
Organization Size			*Respondent sex*		
> 250	91	69.5	Male	77	58.8
100-250	11	8.4	Female	46	35.1
50-100	11	8.4	Missing	8	6.1
< 50	18	13.7			
			Respondent position		

Supply Chain Sector			Country and International executive (VP, chief)	28	21.4
Logistic provider	4	3.1	Regional and local executive managers (DCs, Plants, Cross-docking platforms)	28	21.4
Manufacturing	31	23.7	Indoor Firm coordinators (Warehousing, transportation, operations, finance, logistics)	17	13
Retailer	77	58.8	Outdoor Firm coordinators (Procurement, purchases, Customer service, supplier engagement)	23	17.6
Service (Bank, ICT, Finance, Consulting)	9	6.9	Supervisor	28	21.4
Other	10	7.6	Technician	4	3.1
			Other	2	1.5
Geographical scope					
National firm with national operations	40	30.5	*Respondent experience*		
National firm with Latin American operations	44	33.6	4	9	6.9
National firm with global operations	15	11.5	3	5	3.8
International firm with operations in the country	16	12.2	2	58	44.3
Multinational firm with global operations	16	12.2	1	20	15.3
			Missing	39	29.8
Total	131	100		131	100

Bias control

Potential biases were considered in survey and protocol design and in the data analysis. To check non-response bias, we adopted the “continuum of resistance” model (Kypri et al. 2004), thus comparing early and late respondents – where a late respondent is then used as a proxy for a non-respondent. Student's t-tests were performed on early and late waves on all variables, and these tests indicated no statistically significant differences between the groups in both groups.
Further, social desirability in the entire survey bias was reduced through assurance of confidentiality and through questions mixing both the behavior of the organization and its members in general and direct personal behaviors. The institutional items themselves do not relate to personal behaviors or performance and are thus less likely to be affected by a social desirability bias.

Finally, the procedure of the study was the first way to control common method bias (Podsakoff et al. 2003). First, even though the research project was labeled as a comprehensive study to understand human capital management in Latin American supply chains, no reference to the model in *Figure 1* was provided. Second, questions were organized in order to separate the different sections, to prevent respondents from developing theories about possible cause-effect relationships.

Statistical approaches for model testing

Because the objective of our research is theory testing and confirmation, the presented hypotheses were tested using covariance-based structural equation modeling (CB-SEM), which is a common method employed for this type of research (Perols et al. 2013; Hair et al. 2017).

First, to check the reliability of the hypothesized constructs, Confirmatory Factor Analysis was performed (Schreiber et al., 2006). Following indications by Byrne (2013), we also used Average Variance Explained (AVE), Composite Reliability (CR), and Cronbach alpha (CA) to assess construct validity (detailed in section 5.2 below). Acceptable values of CR and CA are above 0.7, while the AVE should be higher than 50%.

The model was then tested using the maximum likelihood (ML) estimation method (White, 1982), because ML provides more realistic indexes of overall fit and less biased parameter values for paths that overlap with the true model, compared to other methods such as generalized least squares and weighted least squares (Olsson et al. 1999; Gómez-Cedeño et al. 2015). The ML estimation assumes that the variables in the model are *conditionally* multivariate normal, which is true for our dataset according to the Doornik-Hansen (χ^2= 104.207; p > χ^2 = 0.000) and Henze-Zirkler (χ^2= 224.383; p > χ^2 = 0.000) tests.

To evaluate the model fit, we use a combination of the chi-square goodness-of-fit statistic and the use of other absolute or relative fit indices (Hu and Bentler, 1999). Regarding the chi-square, there is the need to check for the ratio between the chi-square value and the degrees of freedom in the model, where cut-off values range from <3 to <5, depending on the type of study (i.e., exploratory or explanatory SEM).

With regard to fit indices, they can range from 0 to 1, with values closer to 1 indicating a good fit. Some authors (e.g., Hair et al., 2017) suggest various indices presentation strategies including, among others, the comparative fit index (CFI), the Tucker Lewis index (TLI) and root mean square error of approximation (RMSEA). A satisfactory threshold for CFI and TLI is >0.90 (with a value >0.95 showing excellent fit), whereas RMSEA is supposed to be < 0.05 (Hooper et al., 2008).

Finally, to better explain the results coming from the model tested, we also performed a cluster analysis, a technique that has been largely adopted in HRM research (e.g., Lengnick-Hall et al. 2009). In order to perform the cluster analysis, a two-step clustering algorithm was selected, because it was able to determine the optimal number of clusters by minimizing the variance within each cluster (Punj and Stewart, 1983).

To perform the data analysis, we used Stata v. 15.0. We estimated the measurements, factor analyses, structural models, and cluster analyses. The main conclusion about the statistical experiments is shown below.

Descriptive statistics

Table 3 presents descriptive statistics for the construct indicators.

Table 3. Descriptive statistics for questionnaire items

Items	Mean	St. Dev.
REL 1	2.64	0.938
REL 2	3.29	1.024
REL 3	3.53	0.965
SKILL 1	4.52	0.614
SKILL 2	4.26	0.7
SKILL 3	4.48	0.601
SKILL 4	4.35	0.635
EXP 1	3.47	1.474
EXP 2	3.13	1.358
EXP 3	2.95	1.34
SAT 1	3.66	1.161
SAT 2	4.19	1.001
SAT 3	4.24	0.696

Confirmatory factor analysis

Table 4 presents the results of the confirmatory factor analysis (CFA). All the measurement model fit indicators show a good fit (χ^2/d.f. = 1.12; CFI = 0.987; TLI = 0.982; RMSEA = 0.032). In addition, convergent validity was assessed through significant loadings from all scale items on the hypothesized constructs as well as through the AVE, CR and CA.
AVE ranges were between 49% and 66% (higher or near the 0.5 threshold), and both CR and CA were higher than 0.7 for all the constructs.

Table 4. Confirmatory Factor Analysis

Construct	Factor Loadings	Average Variance Explained	Composite Reliability	Cronbach Alpha
Relationship orientation		66.09%	0.852	0.807
REL 1	0.699			
REL 2	0.914			
REL 3	0.812			
Job skills		49.15%	0.794	0.801
SKILL 1	0.656			
SKILL 2	0.672			
SKILL 3	0.752			
SKILL 4	0.720			

Personal satisfaction		51.84%	0.761	0.812
SAT 1	0.750			
SAT 2	0.789			
SAT 3	0.608			
Supply chain expansion		61.99%	0.827	0.794
EXP 1	0.611			
EXP 2	0.876			
EXP 3	0.848			

As an additional test for discriminant validity, in *Table 5* we report the squared correlation between two latent constructs to their AVE estimates (Fornell and Larcker, 1981). According to this test, the AVE for each construct should be higher than the correlation between each pair of constructs, a condition that is valid for all the constructs.

Table 5*.* Correlation matrix

Variables	*Mean*	*St. dev.*	1	2	3	4
1.Relationship orientation	*3.15*	*0.82*	1			
2. Job skills	*4.40*	*0.49*	0.033	1		
3. Personal satisfaction	*4.03*	*0.81*	0.009	0.089	1	
4. Supply chain expansion	*3.18*	*1.20*	0.007	0.004	-0.002	1

Finally, the common latent factor technique was applied as a further way to address common method bias (Craighead et al. 2011). We found that the common latent variable has a linear estimate of 0.598. This value indicates a variance of 0.357, which is below the threshold of 0.50. Thus, this indicates that common variance does not represent a problem in our study.

Model testing

The postulated path model produced a good fit to the data (χ^2/d.f. = 1.97; CFI = 0.979; TLI = 0.973 RMSEA = 0.039). *Figure 2* reports the results of the hypotheses testing. The structural equation model shows a highly positive and significant relationship between "*Relationship orientation*" and "*Job skills*", thus we can reject the null hypothesis in favor of H1. In turn, higher "*Job skills*" demonstrates having an impact on "*Personal satisfaction*" – thus we can reject the null hypothesis in favor of H2. However, no statistical significance is found for the relationship with "*Supply chain expansion*" – thereby failing to reject the null hypothesis in favor of H3.

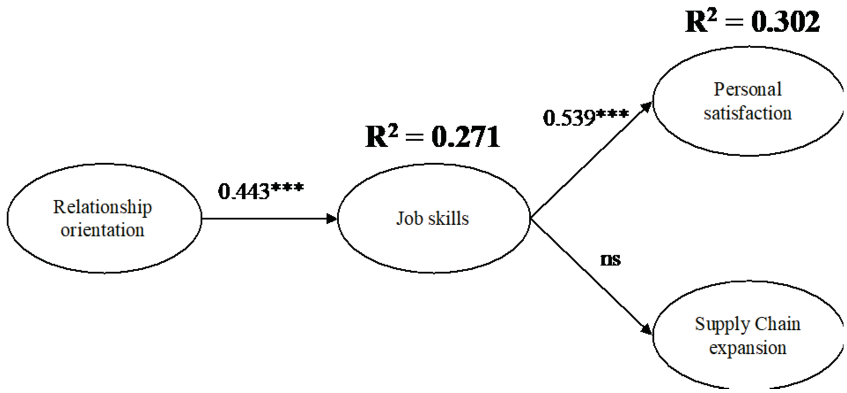


Figure 2. Resultant structural model (**p<0.05; **p<0.01; ***p<0.001)*

Cluster analysis

In performing the cluster analysis we used, as classification variables, the "*Relationship orientation*" and "*Job skills*" factors. The cluster analysis results are reported in *Table 6*.

Table 6. Cluster centroids characteristics (Silhouette coefficient: 0.502; Clusters distance: 1.652)

	Cluster 1		**Cluster 2**		*F*	*P value*
	Mean	*St.dev.*	*Mean*	*St.dev.*		
Relationship orientation	2.41	0.50	3.51	0.71	241.5	0***
Job skills	3.93	0.43	4.62	0.33	3.43	0***
Number of cases	*41*	*32.1%*	*89*	*67.9%*		

The results of the iterative procedure shown at *Table 6* show two different and significant clusters, which can be both differentiated according to the input variables. Cluster reliability was considered acceptable considering the value of the silhouette coefficient – which should be higher than 0.5 – and the cluster distance.
Results show that in cluster 1: *individual – based group*, supply chain professionals tend to have a lower orientation toward having external relationships (below the median, on a 1-5 Likert scale), being characterized by a medium level of skills. In contrast, in cluster 2: *network – based group*, supply chain workers have a higher orientation toward external relationships and an associated higher average level of skills.

After cluster determination, we also ran a one-way analysis of variance (ANOVA) in order to verify if there were statistically significant differences between clusters for "*Personal satisfaction*" and "*Supply chain expansion.*" *Table 7* reports the results below.

Table 7. ANOVA results

	Cluster 1		**Cluster 2**		*F*	*P value*
	Mean	*St.dev.*	*Mean*	*St.dev.*		
Personal satisfaction	3.76	0.8	4.15	0.77	7.34	0***

Supply chain expansion	2.98	1.05	3.26	1.26	1.62	ns

ANOVA at *Table 7* shows, on one hand, that the level of "*Personal satisfaction*" is significantly higher among the network-based group. On the other hand, although the "*Supply chain expansion*" value is higher for the network-based group, it is not statistically significantly greater.

DISCUSSION OF RESULTS

This research analyzes the impact of HRM practices on SCM performance at the individual and organizational levels and confirms the importance, both theoretical and practical, that SCM professionals must establish a powerful network through a relationship orientation in order to enhance their skills. In today's market, where success depends on multiple and complex activities, it is essential for employees to contribute to the competitiveness of their companies and of their supply chains. SCM positions allow employees to interact with talented people inside and outside the company, as well as to access people with multiple hard and soft skills such as leadership and managerial responsibilities (Kiessling et al., 2014). This contact allows individuals and companies to create, through social interactions, a knowledge exchange process contributing to the individual's growth through their daily jobs (Guzzo and Noonan, 1994; Wang-Cowham, 2011).

Finding a strong link between external social relationships and job skills has important implications. At an individual level, SC professionals and managers need to be aware of the opportunities that derive from their social networking in order to boost their skills, and vice versa. Having a large, diverse network of contacts with supplementary, multidisciplinary job profiles will enhance the decision-making processes of the SCM professionals and leverage their learning processes. Thus, at a firm level, giving incentives to SCM employees to build a stronger network could become a valuable organizational practice. In fact, this is already a widespread practice in knowledge-intensive companies operating in complex and turbulent environments such as the computer software, pharmaceuticals, chemicals, and aerospace industries (Wilhelm, 2011; Huggins et al. 2012).

The results support the importance of organizational knowledge exchange as a core concept in KBV. Individuals represent a source of competitive advantage for companies when the employees possess knowledge and skills (representing human capital) that are rare, inimitable and non-substitutable (Kang et al., 2007; Hitt and Ireland, 2002). In addition, these characteristics can be gained not simply through self-development, but also by access to other professionals who already have these skills and play this role for their organization and supply chains through social interaction (i.e., social capital) (Lewin et al. 2011). Therefore, social and human capital should not be treated in isolation (Lepak and Snell, 1999). This is particularly vital for people working in supply chain positions, given the level of complexity that affects both outbound and inbound supply chains. A radical improvement of personal knowledge and skills is possible through continuous interaction with people who have already managed this complexity, and who are able to transfer this experience to other individuals and teams (Mangan and Christopher, 2005; Lengnick-Hall et al. 2013). The aforementioned evidence is also supported by our clustering analysis.

The model's results also support the idea that SCM professionals who have higher skills are also more satisfied with their jobs. Humphrey et al. (2007) found through a large meta-analytical review that motivational and social characteristics (which are closely tied with human capital and social capital characteristics) can explain 34% and 9% of the job satisfaction outcomes of employees, respectively. That is, job satisfaction depends heavily on motivational-human capital characteristics. This evidence is also reinforced through our clustering results' analysis. If we step from the individual-based to the network-based cluster profile, we can see an increment of the level of personal satisfaction with an increase of the average level of employees' job skills.

This justifies investing in the development of skills of supply chain workers, not only to exploit their talent for a more effective SC, but also to maximize potential retention through higher job satisfaction. In the long term, this becomes a sustainable strategy for companies and supply chains that build a baseline of SCM talents that will educate new SCM professionals by training them with the right skills and assuring the transference of acquired lessons and know-how to subsequent generations (Swart et al. 2012; Wood et al. 2016).

Finally, our data was not able to confirm the existence of a positive relationship between the level of job skills and the supply chain expansion, despite the evidence provided by Antoncic and Antoncic (2011). From a theoretical point of view, it is true that some scholars have debated how skilled SCM professionals contribute to company's growth, but it is also true that, according to our definition of supply chain expansion (measured as the increase in the number of employees, products, or suppliers) that relationship is not expected to be linear.

Thus, having skilled employees on its own may not be sufficient to explain company growth, given that the growth may depend largely on the specific characteristics of the company, such as sector, products, location, and size, among others, or on other internal and external variables that arise from the complexity of current businesses practice. For example, scanning the external environment in regard to both competitors and customers from a strategic SCM perspective may serve as a source of competitive advantage (Kiessling et al., 2014). Nevertheless, we hypothesize that using a larger database may dissipate the doubts to conclude with certainty about job satisfaction influence on company growth. For this reason, we are currently running this survey to obtain more observations (at least 100 more in the next couple of months) that will shed light on this issue and determine whether the results are going to change or not.

From an empirical perspective instead, a further scrutiny of our data might better explain this trend. *Table 1, in section 5.1, reports* the descriptive statistics of the items used to measure "*Job skills*" and "*Supply Chain expansion.*" The data show that the relationship is not significant because, on average, the respondents consider the level of skills used for their jobs higher than the growth results. However, we realize that, for each item, the average rate is near or higher than three, which, according to the scale used, still means an expansion between 6-10%. Overall, this indicates that companies and their supply chains still benefit from growth in the case of higher employees' skills; even though, on a 1-5 scale, this outcome is not commeasured to the level of skills in the input.

Therefore, further research is required in that domain by creating a better metric that relates the quality of SCM professionals and their skills to the growth of their company and their supply chains. Lastly, some evidence emerges from cluster analysis: the *network-based group*, characterized by SCM workers with broader external relationships and higher skills, also have a higher average level of growth compared to the *individual-based group* (3.26 vs. 2.96, respectively), suggesting a link may exist overall, even if it is not statistically significant with the sample size of this data set.

Consequently, current research results help to broaden the understanding between HRM-related issues and supply chain outcomes. In short, individuals with higher social capital are more likely to have higher human capital, particularly their in terms of their skills. Meanwhile, job skills are strongly related with employees’ satisfaction. Further research is needed, however, to evaluate the impact of external relations, skills, and satisfaction on supply chain expansion.

CONCLUSIONS AND FUTURE DEVELOPMENTS

This paper´s theoretical contribution helps to clarify the links between HRM constructs in the context of SCM. First, the results provide empirical evidence of the important link between relationship orientation (i.e., social capital) and job skills (i.e., human capital) for SC professionals and managers. The empirical findings also contribute to the existing knowledge base in multiple ways. We find that low- and high-skilled SCM professionals show a different dependence on social relationships. Actually, the individual-based group shows a lower orientation toward building external relationships and is characterized by a medium level of skills; while the network-based group heavily depend on social relationships to improve, train and diffuse knowledge.

Second, it evaluates the impact of job skills on SC outcomes at both the individual (i.e., job satisfaction) and firm level (i.e., supply chain expansion) units of analysis. In this regard, job skills were strongly associated with employees’ job satisfaction but not significantly associated with supply chain expansion. By modeling job skills and satisfaction, our results suggest that SCM professionals having higher skills are also those more satisfied about their jobs. In contrast, variation in job skills was not sufficient to explain variation in supply chain expansion. This may be due to other internal and external variables that arise from the complexity of businesses practice or the need for other supply chain expansion metrics.

Our results suggest that SC managers and professionals should be increasingly aware of the impact of developing significant and positive social relationships with SC members (inside and outside the firm) on their skills and career. Furthermore, higher skills may help them access better job profiles that properly fit their expectations and increase their satisfaction. These findings are relevant because roles of supply managers have changed dramatically (and are currently changing) due to several factors such as new business opportunities, cultural and economic contexts, and a rapidly-evolving data-driven business environment (Kiessling et al., 2014; Waller and Fawcett, 2013). Adaptative capacity then becomes a requirement to update skills that are necessary to succeed in complex and constantly evolving SCM environments.

Moreover, SCM and logistics education decision makers should also be aware of these results. Despite the importance of technical knowledge and skills, soft skills (e.g., managing social relationships, leadership, goal orientation, etc.) are also essential for a satisfying career in SCM. Therefore, education programs in SCM and logistics should include material on these soft skills, which are necessary to succeed in an increasingly complex SC environment. Further research would be valuable on how education should change to promote softer skills, in addition to technical skills, in order to achieve productivity alongside professional satisfaction.

The model’s results also indicate that firms must actively manage their social and human capital in order to have skilled and satisfied employees, which could support SC goals. In fact, HRM activities in SCM should reinforce the evidenced link between Relationships – Skills – Job Satisfaction to drive SCM’s sustainable competitive advantage dimensions. However, further research is needed to broaden the understanding about the impact of employees’ job satisfaction on supply chain expansion and firm performance. This will require novel metrics and methodologies that properly characterize SC outcomes.

Finally, in order to offset the limitations of this study, future research may consider the evaluation of other countries, longitudinal data, broader data on each of the construct elements, and larger sample sizes. Acknowledging the implication of the model's results, the proper analysis extension might minimize cultural differences as biases and illuminate the relationship between social capital and skills level on employees' job satisfaction and companies' expansion in a supply chain context.

REFERENCES

Adler, P.S., and Kwon, S. (2002). Social capital: Prospects for a New Concept. *The Academy of Management Review*, *27(1),* 17–40.

Antoncic, J. and Antoncic, B. (2011). Employee satisfaction, intrapreneurship and firm growth: A model. *Industrial Management and Data Systems, 111(4),* 599–607.

Barnes, J. and Liao, Y. (2012). The effect of individual, network, and collaborative competencies on the supply chain management system. *International Journal of Production Economics*, *140*(2), 888-899.

Barney, J.B. (1991), "Firm resources and sustained competitive advantage", *Journal of Management*, *17(1),* 99-120.

Baron, R. A. and Markman, G. D. (2000). Beyond social capital: How social skills can enhance entrepreneurs' success. *Academy of Management Perspectives, 14(1),* 106-116.

Beske, P., & Seuring, S. 2014. Putting sustainability into supply chain management. *Supply Chain Management: an international journal*, *19*(3): 322-331.

Blanco-Alcántara, D., Díez-Esteban, J. M., & Romero-Merino, M. E. (2018). Board networks as a source of intellectual capital for companies: Empirical evidence from a panel of Spanish firms. *Management Decision*.

Boudreau, J., Hopp, W., McClain, J. O., and Thomas, L. J. 2003. On the interface between operations and human resources management. *Manufacturing & Service Operations Management 5* (3): 179-202.

Brauner, P., Runge, S., Groten, M., Schuh, G. and Ziefle, M. (2013). "Human Factors in Supply Chain Management".. In: Yamamoto S. (eds) *Human Interface and the Management of Information.* Information and Interaction for Learning, Culture, Collaboration and Business,. HIMI 2013. Lecture Notes in Computer Science, vol 8018. Springer, Berlin, Heidelberg

Brettel, M., Friederichsen, N., Keller, M., and Rosenberg, M. 2014. How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *International journal of mechanical, industrial science and engineering 8* (1): 37-44.

Byrne, B. M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Psychology Press.

Cao, L. (2011). Dynamic capabilities in a turbulent market environment: empirical evidence from international retailers in China. *Journal of Strategic Marketing*, *19*(5), 455-469.

Carter, C. R., and, Liane Easton, P. 2011. Sustainable supply chain management: evolution and future directions. *International journal of physical distribution & logistics management 41* (1): 46-62.

Chapman, S., Ettkin, L. P., & Helms, M. M. (2000). Do small businesses need supply chain management?. *IIE Solutions*, *32*(8), 31-31.

Choi, T. M., Cheng, T. C. E. and Zhao, X. (2016). Multi-methodological research in operations management. *Production and Operations Management*, *25*(3), 379-389.

Chowhan, J. (2016). Unpacking the black box: understanding the relationship between strategy, HRM practices, innovation and organizational performance. *Human Resource Management Journal*, *26*(2), 112-133.

Craighead, C. W., Ketchen, D. J., Dunn, K. S. and Hult, G. T. M. (2011). Addressing common method variance: guidelines for survey research on information technology, operations, and supply chain management. *IEEE Transactions on Engineering Management*, *58*(3), 578-588.

Derwik, P. and Hellström, D. (2017). Competence in supply chain management: a systematic review. *Supply Chain Management: An International Journal, 22(2)*, 200-218.

Dobrow Riza, S., Ganzach, Y. and Liu, Y. (2018). Time and job satisfaction: A longitudinal study of the differential roles of age and tenure. *Journal of Management*, *44*(7), 2558-2579.

Dowell, D., Morrison, M. and Heffernan, T. (2015). The changing importance of affective trust and cognitive trust across the relationship lifecycle: A study of business-to-business relationships. *Industrial Marketing Management*, *44*, 119-130.

Dubey, R., and Gunasekaran, A. 2015. Supply chain talent: the missing link in supply chain strategy. *Industrial and Commercial Training 47*(5): 257-264.

Ellinger A.E. and Ellinger A.D. (2014). Leveraging human resource development expertise to improve supply chain managers ' skills and competencies. *European Journal of Training and Development*, 38 (1/2), 118-135.

Elvira, M.M. and Davila, A. (2005). Special research issue on human resource management in Latin America, T*he International Journal of Human Resource Management, 16(12)*, 2164-2172.

Fawcett, S. E. and Waller, M. A. (2013). Considering supply chain management's professional identity: The beautiful discipline (or,“we don't cure cancer, but we do make a big difference”). *Journal of Business Logistics*, *34*(3), 183-188.

Fisher, S. L., Graham, M. E., Vachon, S. and Vereecke, A. (2010). Guest Editors' Note: Don't miss the boat: Research on HRM and supply chains. *Human Resource Management*, *49*(5), 813-828.

Flöthmann, C., Hoberg, K. and Wieland, A. (2018). Competency requirements of supply chain planners & analysts and personal preferences of hiring managers. *Supply Chain Management: An International Journal*, *23*(6), 480-499.

Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 39-50.

Forret, M. (2006). The impact of social networks on the advancement of women and racial/ethnic minority groups. *Gender, Ethnicity, and Race in the Workplace, 3(January),* 149–166.

Fossum, J. and Arvey, R. (1986). Modelling the skills obsolescence process: a psychological/economic integration. *Academy of Management Review*, *11(2)*, 362-374.

Goffnett, S.P., Cook, R.L., Williams, Z. and Gibson, B.J. (2012). Understanding satisfaction with supply chain management careers: an exploratory study. *The International Journal of Logistics Management, 23(1)*, 135-158.

Gómez-Cedeño M., MatuteVallejo J., Guitart-Tarrés L., and Castán-Farrero J.M. (2015). Impact of human resources on supply chain management and performance. *Industrial Management & Data Systems*, 115 (1), 129-157.

Gowen Iii, C. R., and Tallon, W. J. 2003. Enhancing supply chain practices through human resource management. *Journal of Management Development 22*(1): 32-44.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management Journal*, *17*(S2), 109-122.

Gunasekaran, A., Patel, C. and Mcgaughey, R. (2004). A Framework for Supply Chain Performance Measurement. International *Journal of Production Economics, 87,* 333-347.

Guzzo, R. A. and Noonan, K. A. (1994). Human resource practices as communications and the psychological contract. *Human resource management*, *33*(3), 447-462.

Hair, J., Hollingsworth, C. L., Randolph, A. B. and Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems*, *117*(3), 442-458.

Halldorsson, A., Kotzab, H., Mikkola, J. H., and Skjøtt-Larsen, T. 2007. Complementary theories to supply chain management. *Supply chain management: An international journal 12* (4): 284-296.

Handfield, R. B., Cousins, P. D., Lawson, B. and Petersen, K. J. (2015). How can supply management really improve performance? A knowledge-based model of alignment capabilities. *Journal of Supply Chain Management*, *51*(3), 3-17.

Hitt, M.A. and Ireland, D.R. (2007). The Essence of Strategic Leadership: Managing Human and Social Capital. *Journal of Leadership & Organizational Studies, 9(1)*, 3–14.

Hohenstein, N. O., Feisel, E. and Hartmann, E. (2014). Human resource management issues in supply chain management research: a systematic literature review from 1998 to 2014. *International Journal of Physical Distribution & Logistics Management*, *44*(6), 434-463.

Hooper, D., Coughlan, J., and Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*, 53-60.

Hu, L. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6(1)*, 1-55.

Huggins, R., Johnston, A. and Thompson, P. (2012). Network capital, social capital and knowledge flow: how the nature of inter-organizational networks impacts on innovation. *Industry and Innovation*, *19*(3), 203-232.

Hult, G.T.M., Ketchen, D.J., Jr. and Arrfelt, M. (2007). Strategic supply chain management: improving performance through a culture of competitiveness and knowledge development. *Strategic Management Journal, 28(10)*, 1035-1052.

Humphrey, S.E., Nahrgang, J.D. and Morgeson, F.P. (2007). Integrating Motivational, Social, and Contextual Work Design Features: A Meta-Analytic Summary and Theoretical Extension of the Work Design Literature. *Journal of Applied Psychology*, *92*(5), 1332–1356.

Huo, B., Ye, Y., Zhao, X. and Shou, Y. (2016). The impact of human capital on supply chain integration and competitive performance. *International Journal of Production Economics*, *178*, 132-143.

Jordan, C. and Bak, O. (2016). The growing scale and scope of the supply chain: a reflection on supply chain graduate skills. *Supply Chain Management: An International Journal*, *21*(5), 610-626.

Jurčević, M., Ivaković, M., & Babic, D. (2009). The Role Of Human Factors In Supply Chains. Retrieved from https://bib.irb.hr/datoteka/408815.THE_ROLE_OF_HUMAN_FACTORS_IN_SUPPLY_CHAINS.pdf

Kang, S.C., Morris, S. and Snell, S. (2007). Relational archetypes, organizational learning, and value creation: Extending the human resource architecture. *Academy of Management Review*, *32*(1), 236–256.

Kang, S.C. and Snell, S. A. (2009). Intellectual capital architectures and ambidextrous learning: a framework for human resource management. *Journal of Management Studies*, *46(1)*, 65-92.

Kehoe, R. R. and Wright, P. M. (2013). The impact of high-performance human resource practices on employees' attitudes and behaviors. *Journal of management*, *39*(2), 366-391.

Kianto, A., Vanhala, M. and Heilmann, P. (2016). The impact of knowledge management on job satisfaction. *Journal of Knowledge Management*, *20*(4), 621-636.

Kiessling, T., Harvey, M. and Akdeniz, L. (2014). The evolving role of supply chain managers in global channels of distribution and logistics systems. *International Journal of Physical Distribution & Logistics Management*, *44*(8/9), 671-688.

Kypri, K., Stephenson, S. and Langley, J. (2004). Assessment of nonresponse bias in an internet survey of alcohol use. *Alcoholism: Clinical and Experimental Research*, 28(4), 630-634.

Lai Wan, H. (2007). Human capital development policies: enhancing employees' satisfaction. *Journal of European Industrial Training*, *31*(4), 297-322.

Lengnick-Hall, M. L., Lengnick-Hall, C. A., Andrade, L. S. and Drake, B. (2009). Strategic human resource management: The evolution of the field. *Human resource management review*, *19*(2), 64-85.

Lengnick-Hall, M.L., Lengnick-Hall, C.A. and Rigsbee, C.M. (2013). Strategic human resource management and supply chain orientation. *Human Resource Management Review, 23,* 366–377.

Lepak, D.P. and Snell, S.A. (1999). The human resource architecture: Toward a theory of human capital allocation and development. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *24*(1), 31–48.

Lewin, A. Y., Massini, S. and Peeters, C. (2011). Microfoundations of internal and external absorptive capacity routines. *Organization science*, *22*(1), 81-98.

Lutz, H., & Birou, L. 2013. Logistics education: a look at the current state of the art and science. *Supply Chain Management: An International Journal 18* (4): 455-467.

Mangan, J., and Christopher, M. 2005. Management development and the supply chain manager of the future. *The International Journal of Logistics Management*, *16* (2): 178-191.

McAfee, R.B., Glassman, M. and Honeycutt, E. (2002). The Effects of Culture and Human Resource Management Policies on Supply Chain Management. *Journal of Business Logistics, 23,* pp. 1 – 18.

McCallum, S. and O'Connell, D. (2009). Social capital and leadership development. *Leadership & Organization Development Journal, 30(2)*, 152–166.

McCormick, E.J. (1976), "Job and task analysis", in Dunnelle, M.D. (Ed.), *Handbook of Industrial and Organizational Psychology*, Rand McNall, Chicago, IL, 651-696.

Mendes, L. and Machado, J. (2015). Employees' skills, manufacturing flexibility and performance: a structural equation modelling applied to the automotive industry. *International Journal of Production Research*, *53*(13), 4087-4101.

Menon, S.T. (2012). Human resource practices, supply chain performance, and wellbeing. *International Journal of Manpower, 33(7),* 769-785.

Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D. and Zacharia, Z.G. (2001). Defining supply chain management. *Journal of Business Logistics*, *22(2)*, 1-25.

Messersmith, J. G., Patel, P. C., Lepak, D. P. and Gould-Williams, J. S. (2011). Unlocking the black box: Exploring the link between high-performance work systems and performance. *Journal of Applied Psychology*, *96*(6), 1105.

Möller, K. K., & Halinen, A. (1999). Business relationships and networks: Managerial challenge of network era. *Industrial marketing management*, *28*(5), 413-427.

Mumford, M.D., Peterson, N.G. and Childs, R.A. (1999), "Basic and cross-functional skills", in Peterson, N.G. and Mumford, M.D. (Eds), *An Occupational Information System for the 21st Century: The Development of O*NET*, American Psychological Association, Washington, DC, 49-69.

Murphy, P.R. and Poist, R.F. (2007). Skill requirements of senior-level logisticians: a longitudinal assessment. *Supply Chain Management: An International Journal, 12(6)*, 423-431.

Myers, M. B., Griffith, D. A., Daugherty, P. J. and Lusch, R. F. (2004). Maximizing the human capital equation in logistics: education, experience, and skills. *Journal of Business Logistics*, *25*(1), 211-232.

Olsson, U. H., Troye, S. V. and Howell, R. D. (1999), "Theoretic fit and empirical fit: The performance of maximum likelihood versus generalized least squares estimation in structural equation models", *Multivariate Behavioral Research*, Vol. 34 No. 1, pp. 31-58.

Othman R. and Ghani R.A. (2008). Supply chain management and suppliers' HRM practice. *Supply Chain Management: An International Journal*, 13 (4), 259-262.

Ou, C.S., Liu, F.C., Hung, Y.C. and Yen, D.C. (2010). A structural model of supply chain management on firm performance. *International Journal of Operations & Production Management, 30(5)*, 526-545.

Perez-Arrau, G., Eades, E. and Wilson, J. (2012). Managing human resources in the Latin American context: the case of Chile. *The International Journal of Human Resource Management, 23(15)*, 3133–3150.
Perols, J., Zimmermann, C., and Kortmann, S. (2013). On the relationship between supplier integration and time-to-market. *Journal of Operations Management*, *31*(3), 153-167.
Podsakoff, P., MacKenzie, S., Lee, J., and Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88(5)*, 879-903.
Punj, G. and Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of marketing research*, *20*(2), 134-148.
Rogelberg, S. G., & Stanton, J. M. (2007). Introduction: Understanding and Dealing With Organizational Survey Nonresponse. *Organizational Research Methods*, *10*(2), 195–209.
Santos, I. (2016). Labor market polarization in developing countries: challenges ahead. World Bank. Retrieved on March 15th, 2019 from http://blogs.worldbank.org/developmenttalk/labor-market-polarization-developing-countries-challenges-ahead
Scarbrough, H. 2000. The HR implications of supply chain relationships. *Human Resource Management Journal*, *10* (1): 5-17.
Schermuly, C. C. and Meyer, B. (2016). Good relationships at work: The effects of Leader–Member Exchange and Team–Member Exchange on psychological empowerment, emotional exhaustion, and depression. *Journal of Organizational Behavior*, *37*(5), 673-691.
Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of educational research*, *99*(6), 323-338.
Shub A.N. and Stonebraker P.W. (2009). The human impact on supply chains: evaluating the importance of “soft” areas on integration and performance. *Supply Chain Management: An International Journal*, 14 (1), 31-40.
Slone, R., Mentzer, J. and Dittmann, J. (2007). Are you the weakest link in your company’s supply chain? *Harvard Business Review*, *8(9)*, 116-127.
Smith–Doerflein, K., Tracey, M. and Tan, C.L. (2011). Human resource management and supply chain effectiveness: an exploratory study. *International Journal of Integrated Supply Management, 6(3-4)*.
Swart, W., Hall, C. and Chen, H. (2012). Human Performance in Supply Chain Management, *Supply Chain Forum: An International Journal, 13(2)*, 10-20.
Van Hoek, R. (2001). The contribution of performance measurement to the expansion of third party logistics alliances in the supply chain. *International Journal of Operations & Production Management*, *21*(1/2), 15-29.
Velazquez-Martinez, J.C., Yoshizaki, H.T.Y. and Mejia-Argueta, C. (2019). “Introduction to Supply Chain Management and Logistics in Latin America”. In: Velazquez-Martinez, J.C., Yoshizaki, H.T.Y. and Mejia-Argueta, C.(eds). *Supply Chain Management and Logistics in Latin America: A Multi-Country*. Bingley, UK: Emerald Publishing Limited, 1-5.
Wagner, S. M. (2011). Supplier development and the relationship life cycle. *International Journal of Production Economics*, *129*(2), 277-283.
Waller, M. A., and Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics 34* (2): 77-84.
Wanous, J. P., Reichers, A. E. and Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures?. *Journal of applied Psychology*, *82*(2), 247.
Wang-Cowham, C. (2011). Developing talent with an integrated knowledge-sharing mechanism: an exploratory investigation from the Chinese human resource managers' perspective. *Human Resource Development International*, *14*(4), 391-407.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, *50(1)*, 1-25.
Wilhelm, M. M. (2011). Managing coopetition through horizontal supply chain relations: Linking dyadic and network levels of analysis. *Journal of Operations Management*, *29*(7-8), 663-676.
Wood, G., Dibben, P. and Meira, J. (2016). Knowledge transfer within strategic partnerships: the case of HRM in the Brazilian motor industry supply chain. *The International Journal of Human Resource Management, 27 (20)*, 2398-2414.
Wright, P.M., McMahan, G.C. and McWilliams, A. (2001).Human resources and sustained competitive advantage. *International Journal of Human Resource Management*, *5(2),* 301-326.
Wright, C. F. and Kaine, S. (2015). Supply chains, production networks and the employment relationship. *Journal of industrial Relations*, *57*(4), 483-501.

DECISION SCIENCES INSTITUTE

Exploring the Antecedents of Supply Chain Performance with Disruption Considerations: from Dynamic Capability Perspective and Supply Chain Resilience as a Mediator

Shih-Jung Juan
Department of Management Information Systems, National Chengchi University
Taipei, Taiwan, R.O.C.
Email: 104356505@nccu.edu.tw

Woo-Tsong Lin
Department of Management Information Systems, National Chengchi University
Taipei, Taiwan, R.O.C.
Email: lin@mis.nccu.edu.tw

ABSTRACT

Under the globalization and the multi-level complications of supply chains (SCs), SC disruptions increasingly occur to the detriment of SC performance (SCP). Here, we explore the ability of dynamic capability (DC) and SC resilience (SCRES) to support SCP with disruption considerations. The research results show that SC agility (SCA) is the determinant of SCP with disruption considerations. Both SC Collaboration and SCA mediate between knowledge-based DC and SCP with disruption considerations and between resource-based DC and SCP with disruption considerations, and their mediating roles are sequential and interdependent.

KEYWORDS: Dynamic capability, Supply chain resilience, Supply chain performance, Supply chain disruption

INTRODUCTION

With increasing environmental uncertainty and globalization, organizations must now more than ever obtain competitive advantages to achieve success. Competition no longer exists merely among organizations but among SCs (Wu et al., 2014) so organizations must factor in the capacities of upstream suppliers and the needs of downstream customers to rapidly respond to external volatility. In this process, SCs have the key role. Moreover, SCs are more complex as they include multilevel suppliers or customers, increasing their vulnerability. Dynamics of the external environment, such as technical progress, changing customer needs, political considerations and economic factors, increase the possibility of SC disruption. When SC disruption emerges, it is crucial to understand which capabilities a focal firm or a SC must have to remedy this disruption. This study focuses on SCP with disruption considerations, rather than SC disruption alone so its key target is SCP during or following SC disruption, and the antecedents of SCP with disruption considerations are the main focus.

Based on a literature review, two capabilities are proposed. First, SCRES demonstrates its importance in the disruptive situation. Second, from dynamics view, DC is a solution of environmental volatility. When SC disruption occurs, resilience is the most important factor for survival, so SCRES is closer to SCP with disruption considerations than DC. Rojo et al. (2018) note that a hierarchy of firm capabilities is composed of operational routines and DCs. The former is operational functions of the organization, and the latter is modifications of operational routines. They are related in that operational capabilities are the visible outcome of DCs (Liu et al., 2013). DC is a value-creating strategy that can improve operational capability; and over the long-term, it is an evolutionary process supporting an organization’s growth. Juttner and Maklan (2011) note that SCRES consists of four capabilities: SC flexibility, SC velocity, SC visibility and SC collaboration (SCC); and of these flexibility, velocity and visibility are included under “agility”. Therefore, SCRES include SC agility (SCA) and SCC. Liu et al. (2013) consider SCA as an operational capability. Likewise, SCRES is assumed to be an operational capability. Accordingly, we reasonably assume that the DC of a focal firm would be the most influential factor of its SCRES.

Several studies find that SCA significantly affects a firm’s financial or operational performance (Tse et al., 2016; Chan et al., 2017; Martinez-Sanchez & Lahoz-Leo, 2018; Liu et al., 2013; Al-Shboul, 2017; Gligor & Holcomb, 2012). Similarly, Wu et al. (2014), Panahifar et al. (2018) and Pradabwong et al. (2017) demonstrate that collaboration supports a firm’s performance. DC is a factor widely applied to explain variances in performance across competing firms (Liu et al., 2013). To date, given the sparseness of literature on how DC affects SCRES, only Lee and Rha (2016), who regards SC ambidexterity as resiliency of SC, examine the effect of DC on the resiliency of SC. In fact, Lee and Rha (2016) do not focus on the effect of DC on SCRES, but on SC ambidexterity. Moreover, it is unclear whether DCs can directly improve organizational performance. Wilden et al. (2013) find that DCs can directly influence organizational performance only when organizational structure is more decentralized, self-managed, and has local autonomy, or under a more competitive environment. In other words, the performance effects of DCs are contingent on organizational structure and competitiveness of the market.

According to resource-based theory and knowledge-based theory, DCs are divided into resource-based dynamic capability (RBDC) and knowledge-based dynamic capability (KBDC). In this study, RBDC, KBDC and SCA contain complex concepts. According to Teece (2007), RBDC is represented by a framework, consisting of sensing, seizing and transforming capability. In addition, a literature review shows that absorptive capability (AC) (Zahra and George, 2002; Liu et al., 2013), knowledge management process capability (KMPC) (Cantor et al., 2014) and organizational learning capability (Wu & Chen, 2014; Rojo et al., 2018) can represent KBDC. Because agility includes flexibility, velocity and visibility according to Juttner and Maklan (2011), agility is extended to the SC context as SCA. We assume that SC flexibility, SC visibility and SC velocity can represent SCA.

In this study, we explore the roles of two DCs of the focal firm, including RBDC and KBDC, as well as SCRES, comprising SCA and SCC, to see how they enable SCP with disruption considerations, combining SCP and SC disruption, to fill the research gap among these three. Furthermore, KBDC, SC velocity and SCP with disruption considerations are all first proposed as newly conceptualized constructs, and SCA is re-conceptualized, showing the uniqueness of this research model.

LITERATURE REVIEW AND HYPOTHESES

We develop the research model by combining the three research streams of SCRES, DC and SCP with disruption considerations. In this section, we propose our research model, literature review and research hypotheses.

Research Model Development

In this study, we explore the capabilities of successful SC in a dynamic environment. A competitive SC must have resilience to disruptions or vulnerabilities and the focal firm must have sufficient DC to deal with environmental uncertainty. Thus, we combine RBDC with KBDC and SCRES to reduce the impact of SC vulnerability and SC disruption. This can enable an effective model of SC risk management that achieves better SCP with disruption considerations. In this research model, DCs of the focal firm, including RBDC and KBDC, are drivers; SCRES, consisting of SCA and SCC, are mediators; SCP with disruption considerations of the focal firm is a dependent variable and industrial type is the control variable.

Development of Hypotheses

Supply Chain Resilience and Supply Chain Performance with Disruption Considerations

Some SCs recover from risk events more effectively than other SCs, which leads to concern about SCRES (Juttner & Maklan, 2011) though there are few empirical studies of SCRES.

Juttner and Maklan (2011) note that “SCRES addresses the SC’s ability to cope with the consequences of unavoidable risk events in order to return to its original operations or move to a new, more desirable state.”

Chan et al. (2017) note that agility helps deliver value to customers, face changes readily, value human knowledge and skills, and form virtual partnerships. To survive in a volatile business environment, firms should not only be agile themselves but should also be agile within their SC relationships (Tse et al., 2016). Chan et al. (2017) contended that a firm’s level of SCA represents the strength of the interface between the firm and its markets. Swafford et al. (2006) define SCA as the capability of the SC to adapt or respond quickly to a dynamic and unpredictable business environment. An agile supply chain enables its member firms to be more market-sensitive, more capable of synchronizing supply with demand, and better able to achieve shorter cycle times so SCA is widely considered to be the most critical success factor in today’s competitive marketplace (Chan et al., 2017). Tse et al. (2016) regard SCA as a firm’s ability to transform the threats of market uncertainty and SC disruption into competitive opportunities by increasing visibility in inventory and demand levels, and satisfying customer needs with speed and flexibility. This is in line with the view of Juttner and Maklan (2011). Accordingly, we consider SCA to reflect the three dimensions of flexibility, visibility and velocity.

Upton (1997) defines flexibility as “the ability to change or adapt with little penalty in time, effort, cost or performance.” Swafford et al. (2006) posit that flexibility in a firm’s SC process derives from co-alignment of its range and adaptability dimensions. Swafford et al. (2006) propose that SC process flexibilities include procurement/sourcing flexibility, manufacturing flexibility and distribution/logistics flexibility in a firm’s internal SC. Rojo et al. (2018) note that SC flexibility refers to the ability of a SC to react to changes in the environment. Researchers have suggested which dimensions SC flexibility should include, and we here adopt the model of SC flexibility proposed by Moon et al. (2012). A flexible SC can be not only a reactive capability, but also perform a strategic role (Rojo et al., 2018). In uncertain environments, firms can develop a competitive advantage by using flexibility to handle uncertainty and dynamics better than their rivals (Rojo et al., 2018).

Williams et al. (2013) define SC visibility as the access to high quality information that describes various factors of demand and supply. Scholars assume that the critical factors accelerating SC visibility are an automatic system and closer relationships between the SC partners. There is agreement that the resulting SC visibility is useful for high quality of information sharing (Papert et al., 2016; Wei & Wang, 2010; Williams et al., 2013).

Though speed is critical in SCs to sustain competitive advantage, there is currently little consensus over a definition of SC speed. Chan et al. (2017) note that the speed of a firm’s response to key SC outcome measures reduces manufacturing lead-times, increases new product introductions, and improves customer service. Speeds pertinent to SC management

(SCM) include customer response speed (Chiang et al., 2015), purchasing decision-making speed (Kaufmann & Gaeckler, 2015) and product launch speed. Here we combine the SC velocity concept suggested by Chan et al. (2017) and two specific SC speed constructs proposed by Chiang et al. (2015) and Kaufmann and Gaeckler (2015) in our SC velocity construct.

The SC Operations Reference (SCOR) model measures SCP by monitoring and diagnosing overall SC health using APICS (2017). To improve SCP measurement, Beamon (1999) proposes a new framework for SCP measurement including three key elements of strategic goals: resource, output and flexibility. Wu and Chang (2012) using the balanced scorecard propose SCP measurement in the e-SCM context. Wu et al. (2014) further categorize organizational SCP measurement into finance and non-finance in a complementary manner. Here we adopt the financial measures of SCP from Wu et al. (2014) because the research model focuses on evaluating disruption.

SCA reflects the complex coordination and integration among different channel members, which enable firms to change SC practices in response to market changes (Liu et al., 2013). Due to the growing need for timely and cost-effective means of product and service delivery, SCA is necessary for superior firm performance (Liu et al., 2013). SCA mediation between drivers and firm performance is evidenced with significant impact by Chan et al. (2017), Liu et al. (2013), and Gligor and Holcomb, (2012). Even focusing on the effects of disruption, we still argue that SCA is a determinant of SCP. Accordingly, we propose the following hypothesis:

H1, SCA has positive impact on firm’s SCP with disruption considerations.

There are several important reasons for inter-firm collaborations, for example, sharing the cost of large investments, spreading risk, and accessing complementary resources (Wu et al., 2014). Firms may collaborate to access resource combinations or improve capabilities (Pradabwong et al, 2017). Liao, Hu and Ding (2017) note that collaboration is a significant process that leads to value-creation opportunities in SCM. SCC targets mutual benefits and profit maximization for the SC members through a well-coordinated plan and delivery of offerings (Pradabwong et al, 2017). Prior research has identified SCC in various ways, generally concerned with the relationship, the process and mutual benefits gained from collaborating with SC partners (Pradabwong et al, 2017). Angerhofer and Angelides (2006) propose “a collaborative supply chain model” that includes stakeholders, business strategy, processes, enabling technology, topology and levels of collaboration.

In a collaborative SCM, transacting partners (e.g. suppliers and their customers) exchange and integrate information to make strategic or tactical joint decisions (e.g. supply and demand forecasts) (Panahifar et al., 2018). Trading partners with a higher level of collaboration can achieve better operational performance (Panahifar et al., 2018). That collaboration supports a firm’s performance is found by Wu et al. (2014), Panahifar et al. (2018) and Pradabwong et al.

(2017). We also assume that SCC is a precursor of SCP despite SC disruptions. Therefore, we postulate:

H2, SCC has positive impact on firm's SCP with disruption considerations.

Most SCC approaches encourage sharing information (Panahifar et al., 2018). The results of improved information sharing include improved coordination for faster response, increased agility and greater flexibility (Panahifar et al., 2018). SCA requires a firm to closely supervise the legally separate but operationally interdependent partners to maintain a close and coordinated relationship (Liu et al., 2013). This means that SCA refers to how well a firm collaborates with its channel partners in executing SC practice (Liu et al., 2013). Thus, better SCC leads to higher SCA. Gligor and Holcomb (2012) find that SC coordination, cooperation and communication are the determinants of SCA. Accordingly, we propose the following hypothesis:

H3, SCC supports SCA.

<u>Dynamic Capability of the Focal Firm and Supply Chain Resilience</u>

Liu et al. (2013) note that "DCs are defined as the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments and achieve new and innovative forms of competitive advantage." The DC perspective is a widely applied paradigm to explain how performance varies across competing firms (Liu et al., 2013). DCs emphasize the importance of change in the capabilities that support firm-specific resources, and they focus particularly on the development of these resources. Because DCs include both tangible and intangible resources of a firm, we refer to them as resource-based dynamic capabilities (RBDCs). RBDCs include three clusters of activities: sensing, seizing and transforming capabilities (Teece, 2007). Sensing refers to scanning, perceiving, absorbing, and interpretive activities of organization. Seizing refers to the ability of organization to make timely decisions in order to develop new opportunities. Transforming is the ability of organization to reconfigure organizational resources timely and efficiently for matching the changed environment. RBDCs are generated and based only on the combination of these three capabilities, so we can consider RBDCs as a strategic integrating capability.

However, SCA is one type of operational capability (Liu et al., 2013), and in the firm capability hierarchy proposed by Rojo et al. (2018), DC is higher than operational capability. Thus, after a focal firm makes a strategic decision, this strategy would modify operational process. Vanpoucke et al. (2014) reveal that supplier integrative capability (i.e. RBDC) significantly influences process flexibility. Moreover, Dubey et al. (2018) show that SC connectivity is a predictor of SC visibility, while Zhang and Cao (2018) find that inter-organizational systems are a predictor of SCC. Hence we regard SC connectivity and inter-organizational systems as one type of assets within SC and propose the following hypotheses:

H4, A firm's RBDC supports SCA.

H5, A firm's RBDC supports SCC.

A DC with intangible resources such as knowledge is also a source of competitive advantage. From the organizational learning perspective, recognizing the new external knowledge from SC partners is a very important process within knowledge acquisition. Furthermore, an organization utilizes the knowledge management (KM) system to organize and disseminate useful knowledge among its employees. And the organization continues to learn through external challenges. In this study, we formulate KBDC with AC, KMPC and organizational learning.

Zahra and George (2002) consider AC as "a DC pertaining to knowledge creation and utilization that enhances a firm's ability to gain and sustain a competitive advantage". KMPC is the ability of an organization to acquire, create, transfer, integrate, share and apply knowledge related resources and activities across functional boundaries to generate new knowledge. It should be recognized that organizational learning as a group-based learning approach originates from individual learning (Wu & Chen, 2014).

An organization with knowledge-based dynamic learning capability will generate collective memory that is embedded in the processes and culture of the organization. Therefore, an organization with KBDC would make better strategic decisions in response to a changing environment than its competitors. These effective strategies will in turn support their SC operations and executive factors such as SC flexibility, SCA and SCC. Liu et al. (2013) find that AC has a significant effect on SCA, while Rojo et al. (2018) show that both operational AC and organizational learning are the determinants of SC flexibility. Cantor et al. (2014) show that the KM process is a predictor of joint planning with suppliers. Accordingly, we propose the following hypothesis:

H6, A firm's KBDC facilitates SCA.

H7, A firm's KBDC facilitates SCC.

Mediating Effects

According to the hierarchy of firm capabilities proposed by Rojo et al. (2018), because DC pertains to the high-level strategic capability of a firm, SCA and SCC are the ordinary operational capability of a firm. When SC disruption occurs, DC can affect SCP only through SCRES because a system without the element of resilience cannot prevent or cope with the results of a disruption (Ponomarov & Holcomb, 2009). In contrast, SCA and SCC are both elements of SC resilience. Therefore, we postulate the following:

H8a, SCA mediates the association between RBDC and SCP with disruption considerations.

H8b, SCA mediates the relatedness between KBDC and SCP with disruption considerations.
H9a, SCC mediates the association between RBDC and SCP with disruption considerations.
H9b, SCC mediates the relatedness between KBDC and SCP with disruption considerations.

Control Variable

Firms in highly dynamic industries, such as high-tech industries, have shorter product lifecycles, in which time-to-market is critical (Wu et al., 2014). These firms show higher revenue change and customer turnover compared to those in less dynamic industries (Wu et al., 2014). Some studies suggest that industry type is an important variable to control the performance of SC practice (Wu et al., 2014). Thus, a firm's SCP under disruption should also consider industrial type.

METHODOLOGY

We collected research data through a survey, and used PLS-SEM for data analysis. The analysis unit is focal firm. As described below, the questionnaire was mostly adopted from previous studies apart from two items of SC velocity and one item of seizing capability, which were developed by the authors.

Measurement

The questionnaire has two parts. The first part is the basic data for firms and respondents; the second part uses a 7-point Likert scale (ranging from 1 = strongly disagree to 7 = strongly agree) to analyze five constructs: RBDC, KBDC, SCA, SCC and SCP with disruption considerations. We considered their sub-constructs and the indicators of these constructs and selected the unambiguous constructs without overlapping sub-constructs as our constructs.

SC flexibility uses the scale from Rojo et al. (2018). SC visibility uses the scale from Dubey et al. (2018). SC velocity adapts the scales from Chiang et al. (2015), and Kaufmann and Gaeckler (2015), together with the supply chain speed construct self-developed from Kaipia (2008) and adapted from Lee and Rha (2016). SCC uses the scale from Wu et al. (2014). RBDC uses the scale from Lee and Rha (2016). AC adapts the scale from Garcia-Morales et al. (2007). KMPC uses the scale from Wu and Chen (2014). Organizational learning capability uses the scale from Rojo et al. (2018). SCP with disruption considerations adapts scales from the financial measures of Wu et al. (2014).

The questionnaire design considered the problem of common method bias to elicit subjective responses from the informants. Podsakoff et al. (2003) notes that common method bias is due to measurement methods rather than the construct of interest. We thus separated the items of the questionnaire into four sections, according to the content and the length of each section. Each section was preceded by a brief introduction to transfer the informants' attention and avoid fatigue. In the questionnaire preface, we assured informants of their anonymity and that response data was only for analysis. Informants did not need to record their email address to answer this online questionnaire.

Subject

Many industries in Taiwan are transforming into multinational corporations with increasingly international contacts and thus face increasing risk of SC disruption. For the context of this study, B2B manufacturing firms are more appropriate to be the subject. We investigated B2B focal firms from the 2018 list of manufacturing firms published by the Taiwan Stock Exchange Corporation and OTC Corporation. For the pilot phase, we obtained the 2018 list of public companies published by the Taiwan Stock Exchange Corporation (TSEC) available online and extracted manufacturing firms. A total of 630 manufacturing public companies received emails with an online questionnaire hyperlink. There are thus a total of 30 respondents for analysis in the pilot phase. In the formal investigative phase, the same invitation emails were followed up with three follow-up emails. Next, over-the-counter-listed (OTC-listed) companies were searched from the OTC Exchange website and 1,115 manufacturing listed companies were accessed. These companies also received this invitation email with an online questionnaire hyperlink and two follow-up emails were sent when there was no response. The investigation period of these two phases was from March 12th to April 19th of 2019. There were 72 responses from the two phases and the initial analysis was based on these 72 samples. The samples in the pilot phase were put into the initial analysis. The respondent firms are mostly high-tech and middle and small corporations but have high annual revenue. The informants are mostly upper-managers, especially vice general managers over one-fourth, and their service years are mostly over 20 years.

Survey Procedures

The survey had three stages, including the pretest and pilot test of the questionnaire and the formal survey. In the pretest stage, the questionnaire was evaluated by five professors from four universities of Taiwan and the upper managers of three manufacturing firms to confirm its surface validity. In the pilot survey procedure, 30 samples were used to assess the questionnaire. The item with the highest factor loading from every first-order sub-construct of three original second-order constructs (i.e. KMPC, SC flexibility and SC velocity) was selected to reflect these three first-order sub-constructs. This completed the formal questionnaire, as shown in the Appendix. In the formal survey procedure, the final questionnaire was our measurement instrument.

Measurement Model

This study uses structural equation modeling with Smart PLS 2.0 M3 to estimate the research model. PLS-SEM is appropriate for this study for the following reasons (Ali & Park, 2016; Hair et al., 2017):

1. This is an exploratory study. The hypothesis associations between dynamic capabilities (i.e. RBDC and KBDC) and SC resilience (i.e. SCA and SCC) are in the early stage of theory development.
2. The structural model is complex, including both first-order and second-order reflective constructs.
3. The sample size (n=72) is relatively small and the data are not all normally distributed.

This study follows a two-step approach to analyze and interpret the PLS-SEM results: (1) assess measurement model, and (2) test the structural model. We implement PLS algorithms in the first stage.

Scale Refinement and Reliability

In these results, the meanings of the items of the three sub-constructs of RBDC and KBDC seem partially redundant or duplicated. Thus, we deleted the fourth item from sensing, the third item from seizing and the fifth item from transforming in the RBDC construct. The second and third items from absorptive capability, the third and fourth items from knowledge management process capability, and the first and third items from organizational learning in the KBDC construct were all eliminated to increase validity.
Cronbach's alpha is the most generally accepted criteria for internal consistency reliability. Cronbach's alpha values of all second-order sub-constructs and first-order constructs range from 0.852 to 0.978 which is satisfactory for a threshold value 0.7.

Validation

Scale validation proceeds in the two phases of convergent validity and discriminant validity analyses.
Convergent validity is assessed by three criteria (Wu & Chen, 2014) as following:
1) All item loadings should be larger than 0.7;

2) Composite reliability (CR) for each construct should be greater than 0.8;
3) Average variance extracted (AVE) for each construct should be larger than 0.5.

The factor loadings of all items range from 0.753 to 0.981, the CR for all constructs from 0.9 to 0.983, and AVE from 0.694 to 0.955. These three criteria are all satisfactory. Discriminant validity between constructs is assessed according to the recommendation of Fornell and Larcker (1981) that the square root of AVE for each construct should be larger than its correlations with all other constructs. The lowest square root of AVE 0.833 (corresponding to SC flexibility) is higher than the highest correlation between any pair of constructs 0.829 (between seizing and transforming). Hence, the discriminant validity meets the criterion.

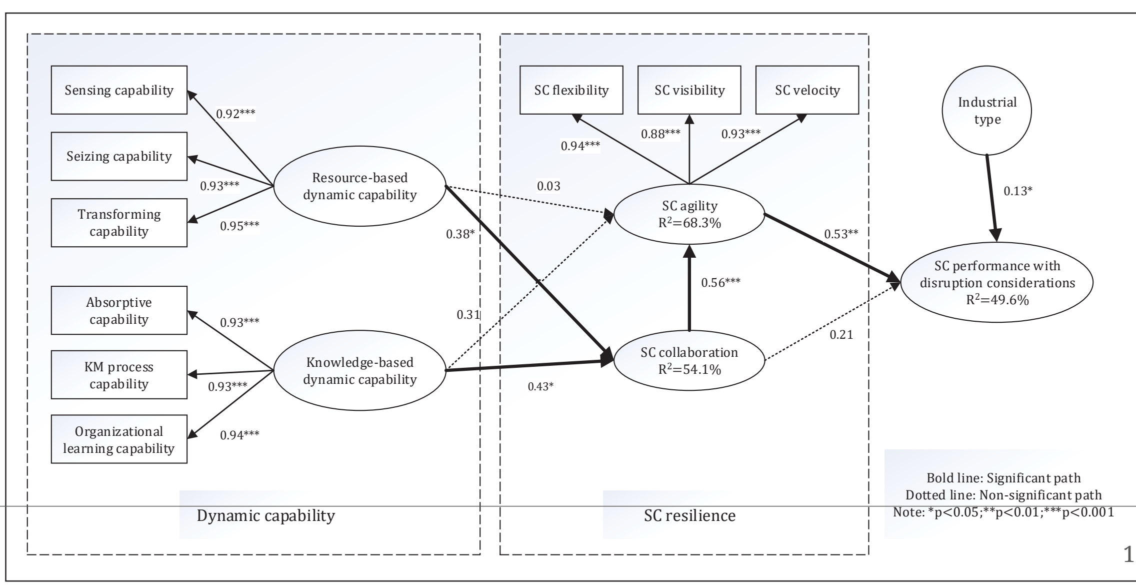
DATA ANALYSIS AND RESULTS

In this stage, we analyze hypothesized relationships in our research model and the mediating effects of SC resilience, and then discuss the findings of this research.

Structural Model: Path Analysis or Direct Effects

We verify causal relationships of the hypotheses in our research model using PLS for our data to estimate the standardized path coefficients and their statistical significance for the hypothesized paths and compare relative effect sizes for common dependent variables. It is also necessary to examine the relative importance of the first-order sub-constructs in reflecting the second-order construct, namely the sub-construct factor loadings of RBDC, KBDC and SCA. We ran the bootstrapping algorithm with 1,000 resamplings (Hair et al., 2017, p.97, 195). The results of the analysis, including path coefficients, path significances, sub-construct factor loadings, and variance explained (R^2 values) for each dependent variable, are shown in Figure 1.

Figure 1 Structural Model



For the path between sub-construct and construct, sensing capability, seizing capability and transforming capability all show high predominant loadings, reflecting RBDC (factor loadings = 0.92, 0.93 and 0.954). And they all have high variances (R^2= 84.6%, 86.5% and 91%) explained. Transforming capability is the highest, followed by seizing capability and then sensing capability. Of these, transforming capability is the key factor of RBDC. Similarly, AC, KMPC and organizational learning capability all represent high notable loadings, reflecting KBDC (factor loadings = 0.926, 0.934 and 0.939). In addition, they all have high variances (R^2= 85.8%, 87.3% and 88.1%) explained. Organizational learning capability is the highest, followed by KMPC and then AC. Of these, organizational learning capability is the key factor of KBDC. Likewise, SC flexibility, SC visibility and SC velocity all indicate highly significant loadings, reflecting SCA (factor loadings = 0.94, 0.881 and 0.93). They also all have high variances (R^2= 88.3%, 77.6% and 86.5%) explained. SC flexibility is the highest, followed by SC velocity and then SC visibility. Of these, SC flexibility is the key factor of SCA.

For dependent variables, only SCA (β=0.53, path coefficient) has significant and positive impact on SCP with disruption considerations, so H1 is supported. SCC (β=0.21) does not have significant and positive effect on a firm's SCP with disruption considerations, so H2 is not supported. SCP with disruption considerations has 49.6% of its variance explained. However, SCC (β=0.56) has significant and positive effect on SCA, thus supporting H3. SCA has 68.3% of its variance explained. But RBDC (β=0.03) and KBDC (β=0.31) do not have significantly positive impacts on SCA, thus not supporting H4 and H6. RBDC (β=0.36) and KBDC (β=0.43) are both the prominent determinants of SCC, so H5 and H7 are supported. They jointly explain 54.1% of the variance in SCC. Industry type (β=0.13) has significant impact on SCP with disruption considerations.

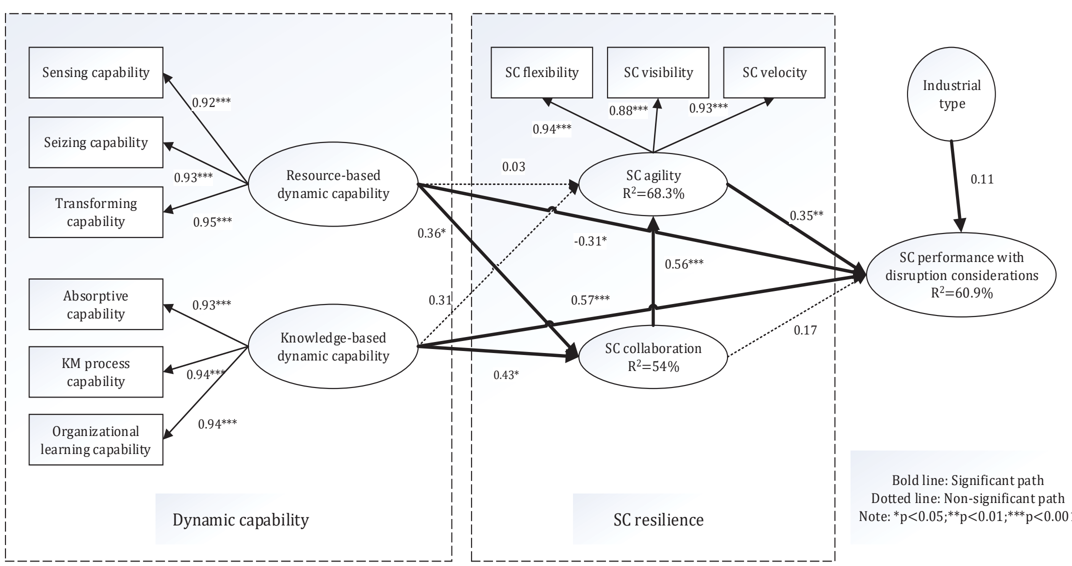
Mediating Effects

The mediating effect is supported when the path from independent variable to the mediator and the path from mediator to the dependent variable are both significant. In this case, the condition on the direct path from the independent variable to the dependent variable will be considered (Hair et al., 2017: p.228-235). Since the association from RBDC to SCA is not significant, and even though the effect of SCA on SCP with disruption considerations is prominent, the direct relationship between RBDC and SCP with disruption considerations is not considered, thus not supporting H8a. Similarly, because the link from KBDC to SCA is not significant, and although the effect of SCA on SCP with disruption considerations is significant, the direct relationship between KBDC and SCP with disruption considerations is not considered, so not supporting H8b. Likewise, since the relationship between SCC and SCP with disruption considerations is not significant, H9a and H9b are not supported.

Three mediating effects are found, including the mediating effect of SCA from SCC to SCP with disruption considerations, the two SCC from RBDC to SCA and from KBDC to SCA. Since the effect of SCC on SCP with disruption considerations is not significant, SCA has full mediating effect for the link from SCC to SCP with disruption considerations. Similarly, since the effect of RBDC on SCA is not significant, SCC has full mediating effect for the association from RBDC to SCA. Likewise, because the effect of KBDC on SCA is not significant, SCC has full mediating effect for the relationship between KBDC and SCA.

Even though H8a, H8b, H9a and H9b are not supported, we still examine the two relationships between RBDC and SCP with disruption considerations and between KBDC and SCP with disruption considerations. The alternative model is proposed as shown in Figure 2. In Figure 2, RBDC (β=-0.31) and KBDC (β=0.57) are both the important predictors of SCP with disruption considerations but the effect of RBDC on SCP with disruption considerations is negative. Both of them and SCA jointly explain 60.9% variance of SCP with disruption considerations. The explained variance of SCP with disruption considerations in the alternative model is 11.3% much more than that in the original model. There is a partial mediating role of SCC and SCA in enabling SCP with disruption considerations.

Figure 2 Alternative Model



Findings and Discussions

In this research model, the first important finding is that the effect of KBDC on SCP with disruption considerations has the highest direct significance. This accounts for the fact that

KBDC is the most important determinant of SCP with disruption considerations. In other words, DC underlying knowledge is superior to SCA. This indicates that through AC, KMPC and organizational learning, organizations can recover from SC disruption more than through SCA. In this study, we integrate AC, KMPC and organizational learning into KBDC from the previous works. It especially stresses the importance of KBDC on SCP in the situation of SC disruption. The second important finding is that the better RBDC, the worse SCP with disruption considerations. Namely, RBDC is the suppressor of SCP with disruption considerations rather than an enabler. These two findings show that KBDC as the source of competitive advantage, integrating AC, KMPC and organizational learning, is a more sophisticated capability for addressing the situation of SC disruption than RBDC. It indicates that learning is imperative for organization, which always needs more time before the success of transforming organization. That a focal firm can succeed in addressing SC disruption must learn to generate effective DC which intangible resource like knowledge is particularly essential.

For SCP with disruption considerations, apart from KBDC, the second important predictor is SCA consistent with the results of the previous researches. In this model, SCC is the only precursor of SCA. Because SCA reflects the connective capability among SC partner network, the success of SCC would drive SCA to response market change and to address SC disruption. Only focal firm's DC cannot trigger SCA no matter which DC is, but through SCC, RBDC and KBDC are the predictors of SCA, thus RBDC and KBDC are the indirect predictors of SCA. This is the third important finding. Only SCC cannot enable SCP with disruption considerations. SCC can improve SCP with disruption considerations only through SCA, so SCC is the indirect determinant of SCP with disruption considerations. KBDC and RBDC are both determinants of SCC and KBDC is more important than RBDC for SCC. This stresses the importance of KBDC once more for SCC. This is the fourth important finding.

In our initial hypotheses, H8 presume that SCA mediates the associations between RBDC or KBDC and SCP with disruption considerations. H9 presume that SCC mediates the associations between RBDC or KBDC and SCP with disruption considerations. In these four hypotheses, SCA and SCC are independent and not related. After the data analysis, we find that the mediating roles of SCC and SCA between RBDC or KBDC and SCP with disruption considerations are sequential and dependent rather than respective and independent. In other words, the effects of RBDC and KBDC on SCP with disruption considerations can be through SCC first and then SCA. This is the fifth important finding of this research.

Accordingly, there are three strong associations in this research model. The strongest path lies in the link from KBDC to SCP with disruption considerations. The second strongest connection is the association from KBDC through SCRES, i.e. SCC first and then SCA, to SCP with disruption considerations. The third strong relationship is from RBDC through SCRES, i.e. SCC first and then SCA, to SCP with disruption considerations.

CONCLUSION

We combine the perspectives of three research streams, including DC, SCRES and SCP with disruption considerations in research framework. In this framework, DCs are drivers, composed of KBDC and RBDC; SCRES is mediator, composed of SCA and SCC; and SCP with disruption considerations is a dependent variable. The initial research findings demonstrate moderately explained variance in SCP with disruption considerations, showing that the mediating role of SCRES and the drivers of DC are almost consistent with our original formulation, apart from three direct linkages from SCC to SCP with disruption considerations, from RBDC to SCA and from KBDC to SCA. The reasons are inferred as following:
(1) Dependent variable contains two concepts, i.e. SCP and SC disruption, thus the effect of SCC on SCP with disruption considerations is not significant: Given scarce extant literature regarding the effect of SCC on SC disruption, most of them are about the effect of SCC on SCP. Because SC disruption is a negative concept, only SCC cannot trigger SCP with disruption considerations.
(2) SCA is one type of dynamic capability, thus the effects of RBDC and KBDC on it are not significant: Rojo et al. (2018) noted that a flexible SC can be not only a reactive capability, but also perform a strategic role. Thus, we can infer that SCA reflected by SC flexibility could not only operational capability but also strategic capability, i.e. dynamic capability. The reason that DCs are not the predictors of SCA is clear.
There are also some limitations in this study. First, this is cross-sectional investigation with a self-reported questionnaire which includes all independent variables and dependent variables, so common method bias is possible. Especially, the dependent variable considers the risk of disruption in SCP. If there is the above case, future researchers can separate independent variables and dependent variable into two investigative phases for the same respondent to complement the cross-sectional investigative method. Second, there are potential endogenity in RBDC. It could avoid this problem to add time-series review into transforming capability items (e.g. the items of organizational learning capability have three years review).

ACKNOWLEDGMENT

The authors are grateful to the Ministry of Science and Technology, R.O.C., for the partial financial support to this work under the grant MOST 108-2410-H-004-205.

REFERENCES

Al-Shboul, M. (2017). Infrastructure framework and manufacturing supply chain agility: The role of delivery dependability and time to market. *Supply Chain Management: An International Journal*, 22(2), 172-185.

Ali, M., & Park, K. (2016). The mediating role of an innovative culture in the relationship between absorptive capacity and technical and non-technical innovation. *Journal of Business Research*, 69, 1669-1675.

Angerhofer, B. J., & Angelides, M. C. (2006). A model and a performance measurement system for collaborative supply chains. *Decision Support Systems*, 42, 283-301.

Beamon, B. M. (1999). Measuring supply chain performance. *International Journal of Operations & Production Management*, 19(3), 275-292.

APICS (2017). *Supply Chain Operations Reference Model.* 12.

Cantor, D. E., Blackhurst, J., Pan, M., & Crum, M. (2014). Examining the role of stakeholder pressure and knowledge management on supply chain risk and demand responsiveness. *The International Journal of Logistics Management,* 251, 202-223.

Chan, A. T. L., Ngai, E. W. T., & Moon, K. K. L. (2017). The effects of strategic and manufacturing flexibilities and supply chain agility on firm performance in the fashion industry. *European Journal of Operational Research,* 259, 486-499.

Chiang, A. H., Chen, W. H., & Wu, S. (2015). Does high supply chain integration enhance customer response speed? *The Service Industries Journal,* 35(1-2), 24-43.

Dubey, R., Altay, N., Gunasekaran, A., Blome, C., Papadopoulos, T., & Childe, S. J. (2018). Supply chain agility, adaptability and alignment: Empirical evidence from the Indian auto components industry. *International Journal of Operations & Production Management,* 38(1), 129-148.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equations with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50.

Garcia-Morales, V. J., Ruiz-Moreno, A., & Llorens-Montes, F. J. (2007). Effects of technology absorptive capacity and technology proactivity on organizational learning, innovation and performance: An empirical examination. *Technology Analysis & Strategic Management,* 19(4), 527-558.

Gligor, D. M., & Holcomb, M. C. (2012). Antecedents and consequences of supply chain agility: Establishing the link to firm performance. *Journal of Business Logistics,* 33(4), 295-308.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares StructuralEquation Modeling (PLS-SEM).* 2nd ed. Los Angeles: SAGE Pub.

Juttner, U., & Maklan, S. (2011). Supply chain resilience in the global financial crisis: An empirical study. *Supply Chain Management: An International Journal,* 16(4), 246-259.

Kaipia, R. (2008). Effects of delivery speed on supply chain planning. *International Journal of Logistics: Research and Applications,* 11(2), 123-135.

Kaufmann, L., & Gaeckler, J. (2015). On the relationship between purchasing integration and purchasing decision-making speed. *International Journal of Physical Distribution & Logistics Management,* 45(3), 214-236.
Lee, S. M., & Rha, J. S. (2016). Ambidextrous supply chain as a dynamic capability: Building a resilient supply chain. *Management Decision,* 54(1), 2-23.
Liao, S. H., Hu, D. C., & Ding, L. W. (2017). Assessing the influence of supply chain collaboration value innovation, supply chain capability and competitive advantage in Taiwan's networking communication industry. *International Journal of Production Economics*, 191, 143-153.
Liu, H., Ke, W., Wei, K. K., & Hua, Z. (2013). The impact of IT capabilities on firm performance: the mediating roles of absorptive capacity and supply chain agility. *Decision Support Systems,* 54, 1452-1462.
Martinez-Sanchez, A., & Lahoz-Leo, F. (2018). Supply chain agility: A mediator for absorptive capacity. *Baltic Journal of Management*, 13(2), 264-278.
Moon, K. K. L., Yi, C. Y., & Ngai, E. W. T. (2012). An instrument for measuring supply chain flexibility for the textile and clothing companies. *European Journal of Operational Research,* (222), pp. 191-203.
Panahifar, F., Byrne, P. J., Salam, M. A., & Heavey, C. (2018). Supply chain collaboration and firm's performance: The critical role of information sharing and trust. *Journal of Enterprise Information Management,* 31(3), 358-379.
Papert, M., Rimpler, P., & Pflaum, A. (2016). Enhancing supply chain visibility in a pharmaceutical supply chain: Solutions based on automatic identification technology. *International Journal of Physical Distribution & Logistics Management*, 46(9), 859-884.
Ponomarov, S. Y., & Hocomb, M. C. (2009). Understanding the concept of supply chain resilience. *The International Journal of Logistics management*, 20(1), 124-143.
Pradabwong, J., Braziotis, C., Tannock, J. D. T., & Pawar, K. S. (2017). Business process management and supply chain collaboration: Effects on performance and competitiveness. *Supply Chain Management: An International Journal,* 22(2), 107-121.
Rojo, A., Stevenson, M., Montes, F. J. L., & Perez-Arostegui, M. N. (2018). Supply chain flexibility in dynamic environments: The enabling role of operational absorptive capacity and organizational learning. *International Journal of Operations & Production Management,* 38(3), 636-666.
Swafford, P. M., Ghosh, S., & Murthy, N. (2006). The antecedents of supply chain agility of a firm: Scale development and model testing. *Journal of Operations Management*, 24(2), 170-88.
Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal,* 28, 1319-1350.

Tse, Y. K., Zhang, M., Akhtar, P., & MacBryde, J. (2016). Embracing supply chain agility: An investigation in the electronics industry. *Supply Chain Management: An International Journal*, 21(1), 140-156.
Upton, D. M. (1997). The management of manufacturing flexibility. *California Management Review*, Winter, 72-89.
Vanpoucke, E., Vereecke, A., & Wetzels, M. (2014). Developing supplier integration capabilities for sustainable competitive advantage: A dynamic capabilities approach. *Journal of Operations Management* 32, 446-461.
Wei, H. L., & Wang, E. T. G. (2010). The strategic value of supply chain visibility: Increasing the ability to reconfigure. *European Journal of Information Systems*, 19, 238-249.
Wilden, R., Gudergan, S. P., Nielsen, B. B., & Lings, I. (2013). Dynamic capabilities and performance: Strategy, structure and environment. *Long Range Planning*, 46(1/2), 72-96.
Williams, B. D., Roh, J., Tokar, T., & Swink, M. (2013). Leveraging supply chain visibility for responsiveness: The moderating role of internal integration. *Journal of Operations Management*, 31, 543-554.
Wu, I. L., & Chen, J. L. (2014). Knowledge management driven firm performance: The roles of business process capabilities and organizational learning. *Journal of Knowledge Management,* 18(6), 1141-1164.
Wu, I. L., Chuang, C. H., & Hsu, C. H. (2014). Information sharing and collaborative behaviors in enabling supply chain performance: A social exchange perspective. *International Journal of Production Economics,* 148, 122-132.
Wu, I. L., & Chang, C. H. (2012). Using the balanced scorecard in assessing the performance of e-SCM diffusion: A multi-stage perspective. *Decision Support Systems*, 52, 474-485.
Zahra, S., & George, G. (2002). Absorptive capacity: A review, reconceptualization and extension. *Academy of Management Review,* 27(2), 185-203.
Zhang, Q., & Cao, M. (2018). Exploring antecedents of supply chain collaboration: Effects of culture and inter-organizational system appropriation. *International Journal of Production Economics,* 195, 146-157.

APPENDIX

Formal Questionnaire

Part 1: Basic information
Firm characteristics
1. Industry type
2. Annual revenue (NT$ millions) Last year
3. Number of employees (persons)
4. Number of suppliers

Respondent's characteristics
1. Working experience (years)
2. Education level
3. Gender
4. Position
Part 2: Constructs and item measures
Resource-based Dynamic Capability
Sensing capability
SEN1: We can perceive demand shifts and changes in customer preference before competitors do.
SEN2: We can feel major potential opportunities and threats.
SEN3: We have good observation and judgment ability.
SEN4: We have frequent interactions with other partners to acquire new knowledge related to product development, process innovation, or logistics and distribution practices.#
Seizing capability
SEI1: We can quickly deal with conflicts in the strategic decision-making process.
SEI2: Under any circumstance we can make timely decisions to deal with environmental change.
SEI3: We can reconfigure resources in time to address environmental change.#
SEI4: We can quickly decide our technological innovation and development.+
Transforming capability
REC1: We can successfully realign or reinvent organization in response to (or in anticipation of) market change.
REC2: We can successfully reconfigure organizational resources to come up with new productive assets.
REC3: We are able to engage in resource re-combinations to better match the product-market areas.
REC4: We are able to align skills to meet the current customer's needs.
REC5: We can effectively integrate and combine existing resources into novel combinations.#
Knowledge-based Dynamic Capability
Absorptive capability
AC1: We have a clear division of roles and responsibilities to acquire, analyze, interpret, understand and utilize new external knowledge into its operations.
AC2: We have the necessary skills to implement new acquired knowledge better than its rivals.#

AC3: We have considerable competences for knowledge development.#
AC4: We can recognize new external knowledge and has the competences to absorb, adapt and use this knowledge.
Knowledge management process capability
KC2: We can codify acquired knowledge into accessible and applicable formats.
KT2: We can distribute relevant knowledge throughout the organization.
KI2: We can integrate different sources and types of knowledge.#
KA2: We can apply knowledge to solve new problems.#
Organizational learning capability
OL1: We have acquired and used much new and relevant knowledge that provided competitive advantage over the last three years.#
OL2: We have acquired important capacities and skills that provided competitive advantage over the last three years.
OL3: The improvements of my firm have been influenced fundamentally by new knowledge entering the organization over the last three years.#
OL4: My firm is an organization that fosters learning.
Supply Chain Agility
Supply chain flexibility
SF3: There are multiple suppliers that provide major materials/components/products in our SC.
OSF4: Our SC can adjust manufacturing facilities and processes.
DF2: Our SC can change delivery modes.
ISF1: The information systems of our SC can support the transportation and distribution management.
Supply Chain Visibility
VIS1: Inventory levels are visible throughout the SC.
VIS2: Demand levels are visible throughout the SC.
Supply Chain Velocity
PDS2: Our SC reacts faster to purchasing changes than our competitors.
CRS1: Our SC can quickly modify products to meet customers' requirements.
SCS2: Our SC can quickly deliver products to customers before the due date.+
Supply Chain Collaboration
COL1: My firm and SC partners set up a communication plan for action.
COL2: My firm and SC partners collaborate in developing new market and customer response.
COL3: My firm and SC partners collaborate in designing their processes or products.

COL4: My firm and SC partners collaborate in implementing their operational activities.
COL5: My firm and SC partners have frequent interaction when problems occur.
Supply Chain Performance with Disruption Considerations
FM1: Our supply chain can improve return on investment under the risk of supply chain disruption.
FM2: Our supply chain can improve return on assets under the risk of supply chain disruption.
FM3: Our supply chain can improve sales growth under the risk of supply chain disruption.
FM4: Our supply chain can improve market share under the risk of supply chain disruption.
FM5: Our supply chain can improve production and inventory cost under the risk of supply chain disruption.
Note: # indicates deleted item; + indicates self-developed item.

DECISION SCIENCES INSTITUTE

Examining the Antecedents and Consequences of "The Speed to Recall" on Recall Completion

Anto Verghese
University of Wisconsin- Whitewater
Email: verghesa@uww.edu

David Peng
University of Houston
Email: xpeng@bauer.uh.edu

Arunachalam Narayanan
University of North Texas
Email: Arunachalam.Narayanan@unt.edu

ABSTRACT

Automotive recalls have impacted the companies and the customers involved. To this account, we examine the antecedents of the speed to recall—assessed via National Highway Traffic Safety Administration (NHTSA) investigation time and manufacturers' time to recall— and, subsequently, examine the impact of the speed to recall on recall completion ratio, a measure that captures the effectiveness of the recall process. Our study contributes to the literature and practice, alike, by highlighting factors that affect the speed and effectiveness of recall process while demonstrating the significance of reducing NHTSA investigation time and time to recall in improving recall completion ratio.

KEYWORDS: Automotive recalls, Recall completion ratio, Time to recall, Investigation time, Secondary data

INTRODUCTION

The National Highway Traffic Safety Administration (NHTSA) reports an increasing trend in automotive recalls. In the past few years, the number of vehicles recalled have exceeded the number of vehicles sold (Gorzelany, 2018), and the cost of recalls in 2016 amounted to $22.1 billion dollars (Jibrell, 2018). However, despite the recalls issued, there are approximately 70 million cars on the road that have open recalls (Gillis, 2018). Vehicles with open recalls not only can have safety concerns to their drivers and passengers, but also may pose a real threat to other vehicles on the road. In addition, automotive recalls have drawn the attention of lawmakers. On December 4th, 2015 the US government signed into law Fixing America's Surface Transportation Act (FAST Act), under which automakers must report recalls and their completion rates to U.S. Congress. And, despite the significance of recall completion rates, we have limited understanding about the speed of the recall process that might influence recall completion rates.

Recalls can have a significant impact on affected firms' performance and threaten their very survival (Eilert et al., 2017). For instance, Takata had to file for bankruptcy due to the recalls involving their airbags and the costs incurred in the process (Tajitsu, 2017). Hence, NHTSA and auto manufacturers involved have the motivation to undertake a comprehensive investigation prior to issuing a recall. Besides, auto manufacturers might want to strategically time their recalls to minimize the negative consequences of issuing a recall. In one instance, it is documented that Toyota saved $100 million by delaying a full recall (McCurry, 2010). There is also significant variability in the NHTSA investigation times and auto manufacturers' time to recall (*source* NHTSA), which could influence the effectiveness of recalls completion.

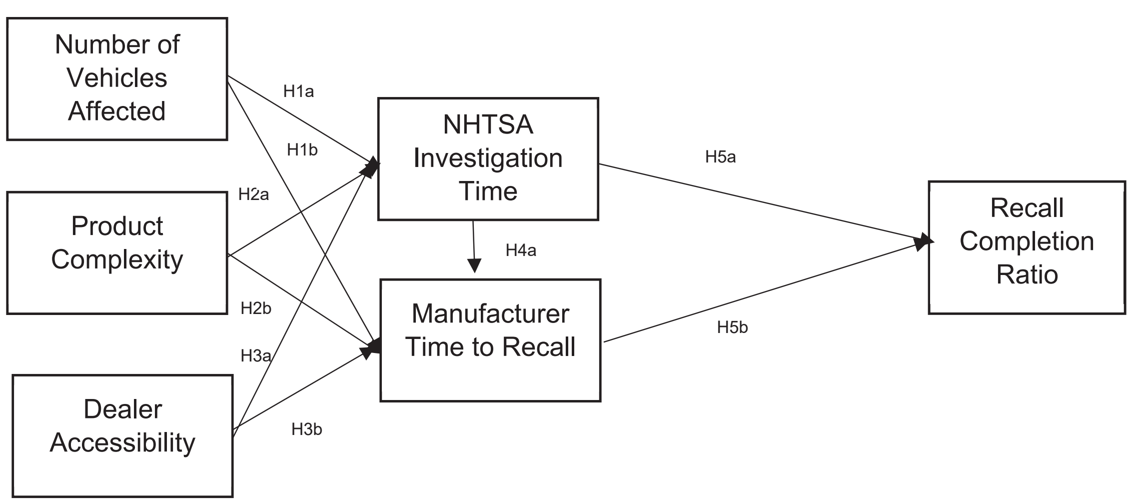
In this study, we examine two distinct components of the speed to recall— NHTSA investigation time and auto manufacturers' time to recall— as factors that might influence recall completion ratio. Besides, we utilize the behavioral theory of the firm to examine the antecedents to NHTSA investigation time and auto manufacturer time to recall. The behavioral theory of the firm suggests that "firms selectively attend to market information, conduct limited search, and find satisficing solutions" (Eilert et al., 2017, p.113). In this regard, we examine the number of vehicles affected, product complexity, and dealership accessibility as factors that can influence the time to recall.

RELEVANT LITERATURE REVIEW

The studies on automotive recalls can be broadly classified into two categories, one examining the effects of recall and the other examining the drivers of recall. Studies that examine the effects of recall largely examine its impact on variables such as firms' financial performance and brand sales. For instance, Liu and Shankar (2015) examine the impact of recalls on brand sales and advertising effectiveness while Borah and Tellis (2016) examine the effect of online negative chatter associated with recalls across other models of the same brand and across brands within same segments and their impact on sales and stock market performance. More recently, Topaloglu and Gokalp (2018) examined how brand concept moderates the relationship between recalls and product sales.

Among studies that examine the drivers of recall, Shah, Ball, and Netessine (2018) examine the plant level drivers of automotive recall. Specifically, they examine the impact of plant variety, product variety, and capacity utilization as the drivers of automotive recalls. Eilert et al. (2017) examine the effect of severity of problem on time to recall contingent of brand reliability. They also examine the effect of time to recall on stock market performance. And, Ni and Huang (2018) examine the effect of the source of defect (i.e., design or manufacturing), parties responsible for defect, types of defect, the number of models involved, and the number of prior recalls on time to recall. None of these studies look at NHSTA or dealer accessibility in relation to time to recall, it must be noted that mandatory recalls are initiated by NHSTA and the relevant customer information and recall service process are completed by the dealers. These two organization are important entities in the automotive recall process.

Figure-1: Proposed Model



Although the extant literature has informed us on the implications of recalls and some factors that can influence recalls. To the best of our knowledge, there is no extant study that has examined the impact of the speed of recall on recall completion ratio, which is significant from a legal, a financial, and a customer perspective. This study also contributes to extant literature by examining two distinct components to the speed of recall that has been largely overlooked, and their antecedents (see Figure-1).

HYPOTHESES DEVELOPMENT

Number of Vehicles Affected → Investigation Time; Time to Recall

As customers submit complaints, NHTSA screens the complaints for safety issues, if deemed serious then it would result in a formal investigation. However, if the number of vehicles affected is large, there is likely to be greater public pressure towards speeding up the formal investigation. Similarly, if the number of vehicles involved is large the manufacturer is likely to issue the recall quickly to salvage brand reputation. Besides, as the number of vehicles involved increases, the chance of injury or a crash related to the recall issue at hand increases, which can then cause legal and financial problems to the organizations involved. Therefore, we propose, the following hypotheses.

H1a: The number of vehicles affected negatively influences NHTSA investigation time.
H1b: The number of vehicles affected negatively influences manufacturers' time to recall.

Product Complexity → Investigation Time; Time to Recall

An engine in car can be strategic product while a windshield wiper can be a commodity product. Strategic products are typically more complex in nature than commodity products, and, hence, it is likely to take longer for both, NHTSA and the manufacturer to determine the root-cause of the problem associated with strategic products vis-à-vis commodity products. As the complexity of the product increases, it requires greater skill, time, and effort to determine the root-cause of the associated problem. Moreover, determining accountability associated with a complex product is going to be relatively more difficult. Hence, we propose the following hypotheses.

H2a: Product complexity positively influences NHTSA investigation time.
H2b: Product complexity positively influences manufacturers' time to recall.

Dealer Accessibility → Investigation Time; Time to Recall

Dealers serve as the primary source of data regarding customer reported problems during an investigation for NHTSA and to the manufacturer. Since, dealers deal directly with customers, they are often very knowledgeable regarding customer reported problems, which can then further help with formal investigations and the manufacturers' recall process in a timely manner. Moreover, since dealer interact with customers on a regular basis, they have access to timely information regarding customer contact details, which can be instrumental while the manufacturer contemplates to issue a recall. Thus, we propose the following hypotheses:

H3a: Dealer accessibility negatively influences NHTSA investigation time.
H3b: Dealer accessibility negatively influences manufacturers' time to recall.

Investigation Time → Time to Recall

With the NHTSA investigation time becoming longer, it is more likely that the manufacturer will have less flexibility to delay issuing the recall, which can be due to the need to preserve brand reputation. Moreover, as NHTSA investigation time increases, it gives the manufacturer greater time to gather relevant information regarding the potential recall and effectively plan resource deployment, which can then help quicken its decision to issue a recall if needed. Thus, we propose the following hypothesis:

H4: Investigation time negatively influences manufacturers' time to recall.

Investigation Time & Time to Recall → Recall Completion Ratio

As the NHTSA investigation time and manufacturers' time to recall increases, the eventual recall notices might not reach the correct customer due to inaccurate databases regarding customers. Moreover, increased investigation time and manufacturers' time to recall are likely to convey a lack of seriousness regarding the recall to the customer. With increased time to recall, customers are likely to have experienced the problems attributed to the recall and might make a self-determination not to expend resources such as their time and effort towards addressing the issue. Therefore, we propose the following hypotheses:

H5a: NHTSA investigation time can negatively influence recall completion ratio.
H5b: Manufacturers' time to recall can negatively influence recall completion ratio.

DATA AND ANALYSIS

Data Collection and Analysis Software

The recall data are obtained from NHTSA while the data regarding dealerships are obtained from a third-party vendor. The data collected involved recall investigations during the period of 2012-2018. Besides, we only focused on mandatory recalls to examine the effect of NHTSA investigation times. The data was compiled using Microsoft Access and analyzed employing the path analysis structural equations modeling (SEM) technique in STATA 13 as the variables use in our study were observed and not latent. A total of 176 unique mandatory recalls were considered in this study.

Measures Used

In this study, our primary dependent variable is captured as the ratio of the number of vehicles remedied to total number of vehicles involved. NHTSA investigation time is recorded as the time since a formal investigation began related to the recall to when the manufacturer was notified, and the manufacturer time to recall is the duration since the manufacturer was notified to when the customer was intimated regarding the recall by the manufacturer. The measure for number of vehicles affected by the recall was obtained from the NHTSA website, and product complexity was coded as either a 'zero' or 'one' depending upon whether it was a commodity type product or a strategic product respectively. And as noted earlier, an engine in car can be strategic product while a windshield wiper can be a commodity product. We use the dealership count for each brand at the time of investigation being opened as a proxy measure for dealership accessibility. All variables were standardized prior to our SEM analysis in STATA.

Controls

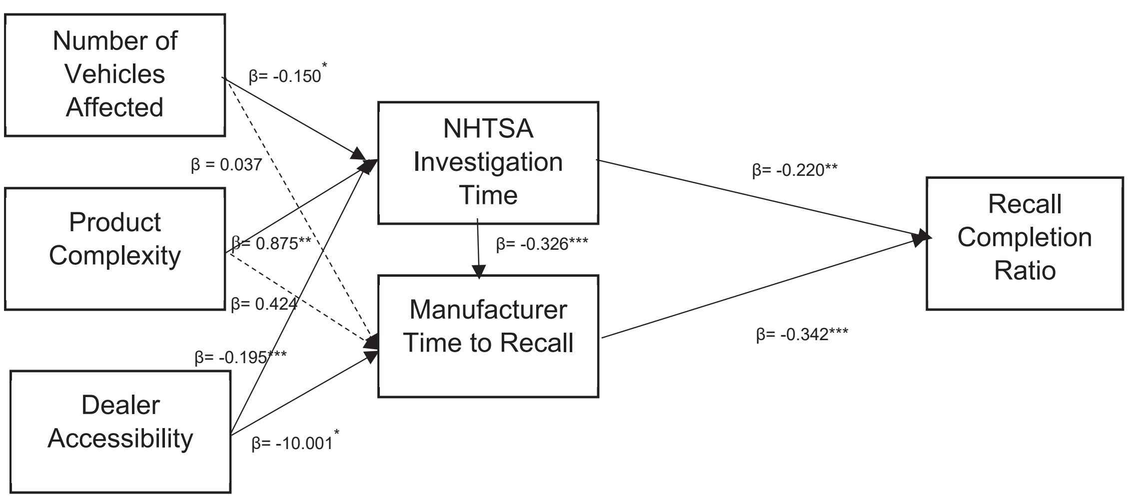
In this study, we control for brand, time fixed effects, the severity of the problem, and the source of the problem (Ni and Huang, 2018). The brand variable was dummy coded, and so were the severity variable and the source variable. Severity dummy variable received either a value zero or a one depending upon whether a fire or a crash was reported related to the recall. If no fire or crash was reported the severity dummy variable received a zero else one. The source of the problem variable received either a value zero or one depending upon if the problem was due to the supplier or the parent company respectively.

Table-1: SEM Path Analysis Results

	Hypothesized Effects	Std. Coefficient
H1a	Number of vehicles affected → Investigation time	-0.150 *
H1b	Number of vehicles affected → Time to recall	0.037
H2a	Product Complexity → Investigation time	0.875**
H2b	Product Complexity → Time to recall	0.424
H3a	Dealer Accessibility → Investigation time	-0.195**
H3b	Dealer Accessibility → Time to recall	-10.001*
H4	Investigation time → Time to recall	-0.326***
H5a	Investigation time → Recall Completion Ratio	-0.220**
H5b	Time to recall → Recall Completion Ration	-0.342***

Notes:
Fit indices (overall): CFI= 0.94; RMSEA= 0.06; R^2 = 0.74
*p<0.05; **p<0.01; ***p<0.001
Control Variables: Brand, time, severity of the problem, source of problem

Figure-2: Model Results



*p<0.05; **p<0.01; ***p<0.001

RESULTS

The results of our analysis are presented in Table-1 and Figure-2. The model fit indices suggest that the model has a satisfactory fit (*CFI = 0.94*, *RMSEA = 0.06; SRMR=0.01; Coefficient of*

Determination= 0.74). H1a and b predict that the number of vehicles affected will negatively influence the investigation time and time to recall (i.e., higher the number of vehicles affected, the investigation time and time to recall will be shorter). We find statistical support for H1a (β= -0.150*) while H1b is not supported (β= 0.037). Similarly, H2 a and b posit that the product complexity will positively influence investigation time and time to recall. We find strong statistical evidence to support H2a (β= 0.875**) while H2b (β= 0.424) is not supported. H3 a and b pertain to the negative relationship between dealer accessibility and investigation time, and time to recall respectively. We find evidence to support both, H3a and b (β= -0.195**; β= -10.001*). In addition, H4 that negatively relates investigation time to time to recall (β= -0.326**) is statistically supported. Moreover, our relationships as proposed in H5 a and b are statistically significant (β= -0.220**; β= -0.342***), which suggests that greater investigation times and time to recall can negatively impact recall completion ratio.

DISCUSSION AND LIMITATIONS

In this research we find the number of vehicles affected in the recall is associated with shorter the investigation time at NHSTA but does not affect the manufacturers' time to recall. This finding suggests two things, first when there is a greater number of vehicles involved NHTSA is more likely to deploy greater amounts of resources towards completing the formal investigation process more quickly, and second, larger or smaller recalls are both treated in the same way by manufacturer as they affect brand image alike. For example, Volkswagen had a highly publicized emissions recall related to its diesel vehicles in US, however the number of diesel vehicles sales was only a fraction of its total sales in US (Ewing, 2015). Also, the complexity of the product positively affects the investigation time at NHSTA, however has no effect on the ability of the manufacturers' time to recall. Specifically, longer investigation times with NHTSA for concerns with strategic parts can be attributed to greater difficulty in identifying the root cause of the problem or need for communication with other stakeholders to accurately determine accountability. We also find that having a higher number of dealers in the network significantly reduces the time to recall and investigation time. This is a significant finding, since having access to customers' current up to date information is essential in completing the recall process and the dealers are the source of such information.

As expected, the NHTSA investigation time does affect the manufacturers' time to recall; the longer the time to investigate, the shorter is the time to recall. Moreover, the investigation time and the time to recall negatively affect the recall completion ratio. Our analysis based on secondary data shows that NHSTA can be a significant bottleneck in the mandatory recall process in automotive industry. Therefore, the resources (staffing and others) and the processes at the NHSTA are significant determinants in the recall process.

There are some limitations in this research, as with any research which uses secondary data. First, this study is limited to the mandatory recall process. Second, the research is based on the automotive industry in US, and the process and practice could be different in other nations, for example the role of entities such as NHSTA. Third, this study does not explicitly control for customer behavior.

REFERENCES:

Bibrell, A. (2018). "Auto recall bill grew 26% to $22 billion in 2016, study says" Autonews (January), https://www.autonews.com/article/20180130/RETAIL05/180139974/auto-recall-bill-grew-26-to-22-billion-in-2016-study-says (Accessed on March 12, 2019).

Borah, A., & Tellis, G. J. (2016). Halo (spillover) effects in social media: do product recalls of one brand hurt or help rival brands?. *Journal of Marketing Research*, *53*(2), 143-160.

Eilert, M., Jayachandran, S., Kalaignanam, K., & Swartz, T. A. (2017). Does it pay to recall your product early? An empirical investigation in the automobile industry. *Journal of Marketing*, *81*(3), 111-129.

Ewing, J. (2015). "Volkswagen says 11 million cars worldwide are affected in diesel deception," The New York Times (September), (Accessed on February 15, 2019).

Gillis, J. (2018). "Over 70 Million Vehicles On The Road With Open Recalls,"ConsumerFed (September), https://consumerfed.org/press_release/over-70-million-vehicles-on-the-road-with-open-recalls/ (Accessed on February 17, 2019).

Gorzelany, J. (2018), "Auto Recalls Hit A Four-Year Low Last Year, But Still Exceed Units Sold," Forbes (March), https://www.forbes.com/sites/jimgorzelany/2018/03/13/auto-recalls-hit-a-four-year-low-last-year-still-exceed-units-sold/#b459d3538567 (Accessed on May 27, May 2018).

Liu, Y., & Shankar, V. (2015). The dynamic impact of product-harm crises on brand preference and advertising effectiveness: An empirical analysis of the automobile industry. *Management Science*, *61*(10), 2514-2535.

McCurry, J. (2010), "Toyota Memo Shows $100m Savings over Floor Mat Recall," The Guardian (February), http://www. theguardian.com/business/2010/feb/22/toyota-memo-safety (Accessed on December 14, 2018).

Ni, J., & Huang, X. (2018). Discovery-to-recall in the automotive industry: A problem-solving perspective on investigation of quality failures. *Journal of Supply Chain Management*, *54*(2), 71-95.

Shah, R., Ball, G. P., & Netessine, S. (2016). Plant operations and product recalls in the automotive industry: An empirical investigation. *Management Science*, *63*(8), 2439-2459.

Tajitsu, N. (2017). "Japanese airbag maker Takata files for bankruptcy, gets Chinese backing," Reuters (June), Available at https://www.reuters.com/article/us-takata-bankruptcy-japan/japanese-airbag-maker-takata-files-for-bankruptcy-gets-chinese-backing-idUSKBN19G0ZG (Accessed on January 10, 2019).

Topaloglu, O., & Gokalp, O. N. (2018). How brand concept affects consumer response to product recalls: A longitudinal study in the US auto industry. *Journal of Business Research*, *88(1)*, 245-254.

DECISION SCIENCES INSTITUTE

Investigation of absorptive capacity, ambidexterity and supply chain innovation

Jan Stentoft
University of Southern Denmark
Email: stentoft@sam.sdu.dk

Christopher Rajkumar
University of Southern Denmark
Email: cra@sam.sdu.dk

Ram Narasimhan
Michigan State University
Email: narasimh@msu.edu

Antony Paulraj
Nottingham University Business School China
Email: antony.paulraj@nottingham.edu.cn

ABSTRACT

Based on 187 survey responses, this paper examines the relationship between absorptive capacity and ambidexterity and the relationships between these concepts and Supply Chain Innovation (SCI). A positive relationship between absorptive capacity and ambidexterity and between absorptive capacity and SCI is found. Data reveal no relationship between ambidexterity and SCI.

KEYWORDS: Absorptive capacity, Ambidexterity, Innovation, Survey research, Structural equation modeling

INTRODUCTION

Supply chain innovation (SCI) has become imperative for any firm to survive in today's dynamic and competitive marketplace (Zimmermann et al., 2016). It is widely acknowledged that supply chains provide vital resources for creating competitive advantages (Narasimhan & Narayanan, 2013). Innovation processes are important both from a single company perspective and from a network perspective with a focus on shared processes (Arlbjørn & Paulraj, 2013; Ojha et al., 2016; Wagner, 2012). Supply chain innovation, over the last decade, has received increased academic attention in extant literature (Arlbjørn et al., 2011; Zimmermann et al., 2016, Wong & Ngai); however, with varied perspectives as to its conceptualization and operationalization. Although extant literature on SCI has grown considerably with respect to its content and processes (Gao et al, 2017; Wong & Ngai, 2019), empirical research on SCI is still sparse and additional work is needed (Grawe, 2009). Researchers have also called for development of measurement scales for operationalizing SCI (Wong & Ngai, 2019).

In this paper, we rely on an understanding of SCI as being comprised of the three components: *business processes, network structure, and technology* (Arlbjørn et al., 2011). SCI may lead to transformations in terms of engagement in a number of actions or projects that might lead to changes in a firm's supply chain, e.g. restructuring of customer and/or supplier portfolios, planning and distribution of goods (Arlbjørn & Paulraj, 2013; Hong & Roh, 2015).

Literature on supply chain innovation acknowledges two factors that are essential for successful outcome in innovation efforts. First absorptive capacity – which is the ability to acquire, disseminate and explicit knowledge – is a requirement for innovation (Cohen & Levinthal, 1990; Kavin & Narasimhan, 2018; Narasimhan & Narayanan, 2013; Zahra & George, 2002). Absorptive capacity is deemed important within a supply chain context because the most relevant information for the firm is often found outside the firm; however, information acquisition may be difficult. Such external information requires effective dissemination and application (Hong & Roh, 2015; Narasimhan & Narayanan, 2013; Preston et al, 2017). Second, the success of any innovation effort depends on the firms' ability to be ambidextrous (i.e. simultaneously conduct operational and development tasks) (Gibson & Birkinshaw, 2004; Im et al, 2019; Narasimhan & Narayanan, 2013). The underlying premise of ambidexterity is that companies that effectively manage both exploratory and exploitive activities will achieve a higher level of supply chain innovations (Fernhaber & Patel, 2012; Kristal et al, 2010). According to Gibson & Birkinshaw (2004) and Jansen et al (2006), ambidextrous companies are more innovative. Research has found a positive relationship between firms having a high level of absorptive capacity and ambidexterity (Solís-Molina et al, 2018).

This research offers two main contributions. First, a theoretical framework for supply chain innovation is presented. Although, other researchers have studied the relationship between absorptive capacity and ambidexterity, there is a paucity of literature linking absorptive capacity, ambidexterity and SCI. Our study fills this gap in current literature. Second, empirical data about the relationship between absorptive capacity and ambidexterity and SCI is analyzed. The paper is organized as follows. First, the theoretical foundation of the paper (research model) and research hypotheses are discussed. The next section describes the sampling frame and empirical data that were collected. The findings from statistical analysis are presented next. Finally, the conclusions drawn from this research along with its limitations, implications, and directions for future research are discussed.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Absorptive Capacity and Ambidexterity

The concept of absorptive capacity was introduced by Kedia & Bhagat (1988) and further explored by Cohen & Levinthal (1990). These authors argue that the ability to exploit external knowledge is a critical component of internal innovative capabilities. The ability to evaluate and utilize knowledge outside the company is a function of the level of prior related knowledge such as basic skills, shared language, and scientific or technological developments in a given field. Based on this, Cohen & Levinthal (1990, p. 128) define absorptive capacity as an "ability to recognize the value of new information, assimilate it, and apply it to commercial ends". Absorptive capacity is concerned about how much knowledge can be absorbed in the company. It can be viewed as a measure of an organization's ability to learn. According to Lane et al (2006), it can be divided into three phases: 1) acquire or search for knowledge, 2) store and share knowledge and 3) analyze and use knowledge.

Ambidexterity refers to the tension induced by the capability to concurrently exploit and explore (March, 1991); it is introduced as a central management task to secure long term survival. Duncan (1976) was the first to introduce the management dilemma of ambidexterity in a discussion on the need for a management structure between the first phases of innovation (i.e. exploration) and the implementation (i.e. exploitation) phases. Extant literature on ambidexterity discusses different forms of ambidexterity (i.e. sequential, simultaneous and contextual) (Gibson & Birkinshaw, 2004; O'Reilly & Tushman, 2008; 2013). Sequential ambidexterity treats operation and development tasks are separated in time. Simultaneous ambidexterity (also named structural ambidexterity) means that the company works on operations and development tasks at the same time. Contextual ambidexterity is focused on individuals or groups that make prioritizations between operations and development tasks (Gibson & Birkinshaw, 2004). In this paper, we don't adopt a specific perspective of ambidexterity but approach the concept as a general dilemma pertaining to prioritization of operational and development tasks in the supply chain.

While absorptive capacity reflects the ability of a company to value and integrate external knowledge for pursuing SCIs, ambidexterity addresses how the knowledge is to be used (Fernhaber & Patel, 2012). Both absorptive capacity and ambidexterity are deemed critical for the success of SCIs. Several studies have found that being ambidextrous requires a firm to expand on their capacity to acquire and process knowledge, i.e. to master absorptive capacity (Datta, 2011; Enkel et al, 2017; Rothaermel & Alexandre, 2009). Accordingly, we posit the following hypothesis:

H1: Absorptive capacity is positively associated with ambidexterity

Supply Chain Innovation
Supply chain innovation, in this paper, is viewed as a higher-order factor covering business processes, network structure and supply chain technology (Arlbjørn et al, 2011; Arlbjørn & Paulraj, 2013). Thus, it is a multi–dimensional construct that also encompasses technology and process innovations (Gao et al, 2017; Hazen et al, 2012; Kwak et al, 2018). Supply chain innovation can be incremental or radical and it can take place within the company or in collaboration with external supply chain partners. We briefly discuss each of these aspects of SCI.

Business Processes
The first component in the supply chain innovation framework proposed by Arlbjørn et al. (2011) is business processes which incorporate activities that generate a distinct output (product/process/service) of value to the customers (Davenport, 1993). In extant SCM literature, it is well-recognized that process thinking constitutes one of the backbones of supply chain management (Ellram & Cooper, 2014; Lambert & Cooper, 2000; Mentzer et al, 2001). The driving force for employing customer-oriented business processes within the firm as well as among partners of the supply chain is to both practice business more effectively and efficiently, and to organize inter-firm relationships (Arlbjørn et al, 2011). This paper uses the eight business processes developed by the Global Supply Chain Forum (Lambert & Cooper, 2000; Lambert & Enz, 2017).

Network Structure
The network structure aspect includes supply chain partners, variety of process networks, and structural dimensions. The structural component deals with both horizontal and vertical structure along with the horizontal position of the firm (Arlbjørn et al, 2011; Lambert & Enz, 2017). Process networks focus on different degrees of resources used toward integrating and managing

processes within and among supply chain partners (Arlbjørn et al., 2011; Arlbjørn & Paulraj, 2013; Munksgaard et al, 2014). In addition, network structure refers to effective networks in which new value creation depends on knowledge resources and core competences that could enhance the capability of the firm to sustain long-term relationship with supply chain partners (Bellamy et al, 2014; Cao & Zhang, 2011; Narasimhan & Narayanan, 2013).

Technology

Munksgaard et al (2014) state that supply chain technologies might be practiced separately or together with other technologies or with the other components of the framework to realize supply chain innovations. The technology component does not relate to the relevant technology in particular; instead it also includes the innovative use of technology within the context of supply chain (Arlbjørn et al, 2011; Gao et al, 2017; Hazen et al, 2012; Kwak et al, 2018). It is important to stress that SCI is not about the relevant technology itself (e.g. big data, automation and additive manufacturing, and other disruptive technologies) (Tan et al, 2015; Vyas, 2016). But it is in the novel use of technology within a supply chain context (Stentoft et al, 2016). Technology enables firms in engaging their customers more instantaneously online and adding more product variety (Arlbjørn et al, 2011).

Extant research has found a positive relationship between absorptive capacity and innovation (Albort-Morant et al, 2018; Berghman et al, 2012; Kostopoulos et al, 2011). Other studies that relate absorptive capacity to innovations within supply chains are Autry & Griffs (2008), Chapman et al, (2003), Gao et al, (2017), Grawe (2009); Narasimhan & Narayanan (2013), and Preston et al, (2017). Furthermore, extant research has also posited the positive relationships between ambidexterity and an innovation outcome (e.g. Im et al, 2019; Narasimhan & Narayanan, 2013; Wong et al, 2013). While most of these research relate absorptive capacity and ambidexterity to product innovations, we believe that both absorptive capacity and ambidexterity are strong proponents of developing an innovation culture within firms as well as supply chains. Accordingly, in line with extant research, we forward the following two hypotheses for formal testing:

H2: Absorptive capacity is positively associated with supply chain innovation (SCI)
H3: Ambidexterity is positively associated with supply chain innovation (SCI)

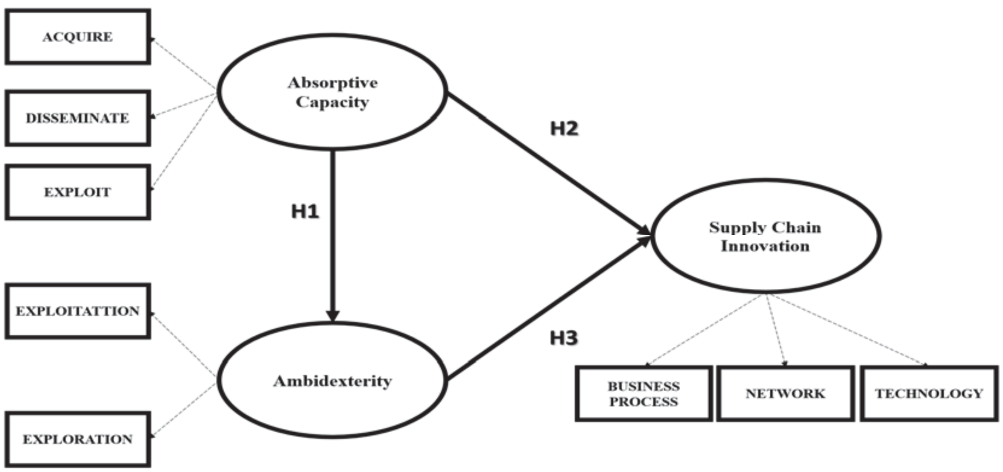
METHOD

This paper examines the relationship between ambidexterity, absorptive capacity, and supply chain innovation. Figure 1 shows the research framework of this paper.

Data Collection

This paper uses a questionnaire–survey that was distributed among Danish manufacturing firms with at least 50 employees in autumn 2016. The population of the companies was identified using the Danish company database “Names and numbers, business” (NN Markedsdata, 2016). The database allowed searching for these companies in a structured manner and the process resulted in an initial sample of 1,580 companies. The selected companies were then contacted and asked to identify the person with the overall responsibility for supply chain management. This process provided us with a net population of 879 companies that agreed to participate in the study. The email link to the electronic questionnaire (SurveyXact, 2016) was sent to all the participating companies. Reminder e–mails were also sent to increase the response rate and allow comparison of early and late responses (before and after the initial deadline). This process finally resulted in 187 companies that provided valid responses with a response rate of approximately 21.3%. The questions included in the survey questionnaire are grounded in literature and was further pretested using industry representatives. The questionnaire included questions related to company profile, supply chain innovation, learning practices, ambidexterity, and performance. The questions were of two types: 1) yes or no questions and 2) statements on a 5–point Likert Scale.

Figure 1: Research framework



Measures

<u>Dependent Variable</u>

Supply chain innovation construct was measured using the scales adapted from Arlbjørn et al (2011) and Cooper et al (1997); this construct included twenty-six 5-point Likert scale questionnaire items (see Appendix A). This construct includes three factors: business processes, network structure, and technologies. Business process (BP) factor includes indicators such as customer relationship management (BPCRM), supplier relationship management (BPSRM), customer service management (BPCSM), demand management (BPDEM), order fulfilment (BPORF), manufacturing flow management (BPMFM), product development and commercialization (BPPDC), and returns management (BPREM). Network structure (NS) factor includes indicators such as internal functions (NSINT), customers (NSCUS), suppliers (NSSUP), third party providers (e.g. logistics providers) (NS3PL), competitors (NSCOM), consultants (NSCON), universities (NSUNI), and public agencies (NSPUB). These represented the diverse sources of external information. Technology (TE) factor includes indicators such as planning and execution systems (TEPLA), identification systems (TEIDF), communication systems (TECOM), analytics technology (TEANA), electronic marketplaces (TEELM), advanced manufacturing technologies (TEAMT), advanced materials (TEADM), big data (TEBIG), and drones (TEDRO).

<u>Independent Variables</u>

The factors exploitation and exploration related to the construct ambidexterity were adapted from Kristal et al (2010); this construct included eight 5-point Likert scale questionnaire items (see Appendix A). The construct ambidexterity–exploitation includes indicators such as “improving our existing supply chain technologies (SCOIT1)”, “reducing waste in our existing supply chain processes (SCOIT2)”, “developing stronger competencies in our existing supply chain processes (SCOIT3)”, “making minor improvements in our existing supply chain processes (SCOIT4)”; and ambidexterity–exploration includes indicators such as “experiment with new supply chain technologies (SCORE1)”, “develop new supply chain processes (SCORE2)”, “developing new competencies in our supply chain processes (SCORE3)”, “implementing large scale improvements (SCORE4)”.

Absorptive capacity was measured using scale adapted from Ho & Lu (2015), Patel et al (2012), Setia & Patel (2013) Tu et al (2006), and Zacharia et al (2011); this construct included seventeen 5-point Likert scale questionnaire items (see Appendix A). This construct includes three factors: acquire, disseminate, and exploit. Acquire (ACQR) includes indicators such as “committed to looking for and developing new ideas in our supply chain (ACQR1)”, “created an environment that encourages identification and development of new and useful ideas (ACQR2)”, “frequently look for external sources of new knowledge and skills (ACQR3)”, “’seek to learn from tracking new market trends in our industry (ACQR4)”, “seek to learn from trying out new technologies (ACQR5)”, “seek to learn from other supply chain actors (ACQR6)””, and collect frequently operations related information in our supply chain (ACQR7)”. Disseminate (DISM) factor includes indicators such as “record and store newly acquired supply chain knowledge for future reference (DISM1)”, “share practical experiences with each other (DISM2)”, “periodically meet to discuss consequences of new products and processes (DISM3)”, “common language regarding our products and processes in our company (DISM4)”, “frequent interactions between departments in the firm to acquire new knowledge (DISM5)”, “employees are engaged in cross-functional work (DISM6)”, “clear division of roles regarding sharing newly acquired supply chain ideas and knowledge (DISM7)”. Exploit (EXPL) factor includes indicators such as “constantly consider how to exploit newly acquired knowledge, skills, and technologies in our supply chain (EXPL1)”, “quickly recognize the usefulness of new external operational knowledge to existing operational knowledge (EXPL2)”, “proficient in integrating newly acquired knowledge into current ways of doing things (EXPL3)”.

Descriptive Statistics

For statistical testing of the hypotheses, we utilized factor analysis and structural equation modeling. Given that the data were collected from a single respondent within each company, common method bias might be a concern. This could be tested using Harman’s single factor test (Podsakoff et al, 2003) which suggests entering all measurement items into a single factor analysis. Common method bias occurs when only a single factor emerges while carrying out factor analysis or if a single factor explains most of the variance in the data. An unrotated factor analysis using the eigen value greater than one criterion revealed eight different factors explaining 64% of the variance and the first factor accounted for only 27% of the variance. Since a single factor did not emerge and the first factor did not explain most of the variance, we could conclude that common method variance might not be an issue in the data. Next, the validity and reliability of the measurement instrument was assessed. The content validity was established by grounding the survey instrument in literature and was further confirmed by pre–testing the survey instrument with the help of scholars and practitioners. Based on their suggestions, the survey instrument was revised with minor changes to the questionnaire.

A confirmatory factor analysis (CFA) was performed to check the model fit, reliability, and construct validity. Model fit was assessed for reliability and validity by using a series of fit measures, where each measure provides different information for evaluating the model (Hair et al, 2006). Table 1 summarizes the overall fit statistics of the measurement model which includes higher–order constructs such as supply chain innovation and ambidexterity.

Table 1: Statistical fit of the overall measurement model	
RMSEA	0.088
90% confidence interval of RMSEA	(0.071;0.10)
χ^2	179.67
CFI	0.98
NNFI	0.97
Standard RMR	0.049
Goodness of Fit Index	0.88

The overall model fit can be tested using the comparative fit index (CFI), non–normed fit index (NNFI), root mean square error of approximation (RMSEA), and normed chi–square (i.e., χ^2 per degree of freedom) (Hair et al, 2006). In general, if the values of CFI and NNFI are higher than 0.90, then it represents good fit and if the value of RMSEA is typically less than 0.10 then it is an acceptable model. The normed chi–square (χ^2 divided by degrees of freedom) estimates the relative efficiency of competing models. In our case, the value of normed chi–square is 2.43, which indicates that it is a reasonable fit (Papke-Shields et al, 2002).

Table 2 presents the coefficients of the items on the higher–order constructs and Cronbach alpha value (reliability). The coefficients of the retained items were higher than the cut off value of 0.50 (Hair et al, 2006). The analysis also confirmed that the items do not cross-load indicating a good fit to the measurement model. In other words, the items were loading high only on their underlying factors. Reliability was tested using Cronbach alpha (α) (Nunnally & Bernstein, 1994) and the value of reliability coefficient need to be 0.70 or higher (Fornell & Larcker, 1981; Nunnally & Bernstein, 1994). The result indicates that Cronbach's alpha values (α) of the factors are greater than 0.70 and therefore, the theoretical constructs are reliable.

Table 2: Confirmatory Factor Analysis

Constructs	**Standardized Coefficients**	**t-value**	**Cronbach alpha**
Supply Chain Innovation (SCI)			
BPCRM	1.00	-	0.62
BPSRM	1.23	7.41	0.71
BPCSM	1.01	6.37	0.58
BPDEM	0.99	6.00	0.54
BPORF	1.10	6.27	0.57
BPMFM	1.07	6.67	0.63
BPPDC	1.10	6.56	0.60
BPREM	*Item removed*		
NSINT	1.00	-	0.59
NSCUS	1.06	5.69	0.55
NSSUP	1.41	6.73	0.75
NS3PL	1.12	5.70	0.56
NSCOM	*Item removed*		
NSCON	*Item removed*		
NSUNI	*Item removed*		
NSPUB	*Item removed*		
TEPLA	1.00	-	0.62

TEIDF	0.87	5.94	0.53
TECOM	1.15	7.62	0.74
TEANA	1.34	8.04	0.84
TEELM		*Item removed*	
TEAMT		*Item removed*	
TEADM		*Item removed*	
TEBIG	0.77	5.76	0.51
TEDRO		*Item removed*	
Ambidexterity			
SCOIT1	1.00	-	0.82
SCOIT2	0.77	10.05	0.67
SCOIT3	0.76	11.08	0.73
SCOIT4	0.76	11.22	0.73
SCORE1	1.05	12.91	0.81
SCORE2	1.02	12.96	0.81
SCORE3	0.91	11.95	0.77
SCORE4	1.08	12.78	0.80
Absorptive Capacity			
ACQR1	1.00	-	0.74
ACQR2	1.10	9.66	0.75
ACQR3	0.90	8.54	0.66
ACQR4	0.88	8.62	0.67
ACQR5	0.91	8.25	0.64
ACQR6	0.72	7.41	0.58
ACQR7		*Item removed*	
DISM1		*Item removed*	
DISM2		*Item removed*	
DISM3	1.00		0.75
DISM4	0.79	8.43	0.64
DISM5	1.05	11.07	0.84
DISM6	0.83	9.27	0.71
DISM7	0.78	8.15	0.62
EXPL1	1.00	-	0.85
EXPL2	1.07	14.69	0.89
EXPL3	0.95	12.64	0.79

RESULTS

The results of the confirmatory factor analysis (CFA) are presented in Table 2. CFA results also indicate a strong evidence of convergent validity since the t–value ranges from 5.69 to 12.96 at the significant level of 0.01 (in general, t–values should exceed the threshold value of 1.96 at the significant level of 0.05). Moreover, the first indicator of all the constructs will not have a t–value because we intentionally made the constructs scale–invariant by fixing the first indicator to 1 (e.g. Kim, 2014).

Structural equation modeling was performed to test the proposed hypotheses and the results of the regression are presented in Table 3.

Table 3: Research model results

			Estimate	S.E.	C.R.	P
AMBIDEXTERITY (ABX)	<---	ABSORPTIVE CAPACITY (ABS)	.968	.118	8.176	***
SUPPLY CHAIN INNOVATION (SCI)	<---	ABSORPTIVE CAPACITY (ABS)	.486	.127	3.831	***
SUPPLY CHAIN INNOVATION (SCI)	<---	AMBIDEXTERITY (ABX)	.176	.093	1.884	.060
Acquire (ACQR)	<---	ABSORPTIVE CAPACITY (ABS)	1.000			
Disseminate (DISM)	<---	ABSORPTIVE CAPACITY (ABS)	.982	.096	10.197	***
Exploit (EXPL)	<---	ABSORPTIVE CAPACITY (ABS)	1.250	.114	10.971	***
Exploration (SCORE)	<---	AMBIDEXTERITY (ABX)	1.000			
Exploitation (SCOIT)	<---	AMBIDEXTERITY (ABX)	.996	.090	11.118	***
Business Process (BP)	<---	SUPPLY CHAIN INNOVATION (SCI)	1.000			
Network Structure (NS)	<---	SUPPLY CHAIN INNOVATION (SCI)	1.247	.164	7.612	***
Technology (TE)	<---	SUPPLY CHAIN INNOVATION (SCI)	.961	.164	5.858	***

The results indicate that hypotheses H1and H2 are statistically significant and the significance level is at 99.9% (p–value < 0.001). Therefore, it can be concluded that the proposed hypotheses are supported. In other words, it could be suggested that there is a positive relationship between absorptive capacity and ambidexterity, and absorptive capacity and supply chain innovation. However, the analysis does not indicate a relationship between ambidexterity and supply chain innovation, failing to provide support for hypothesis H3.

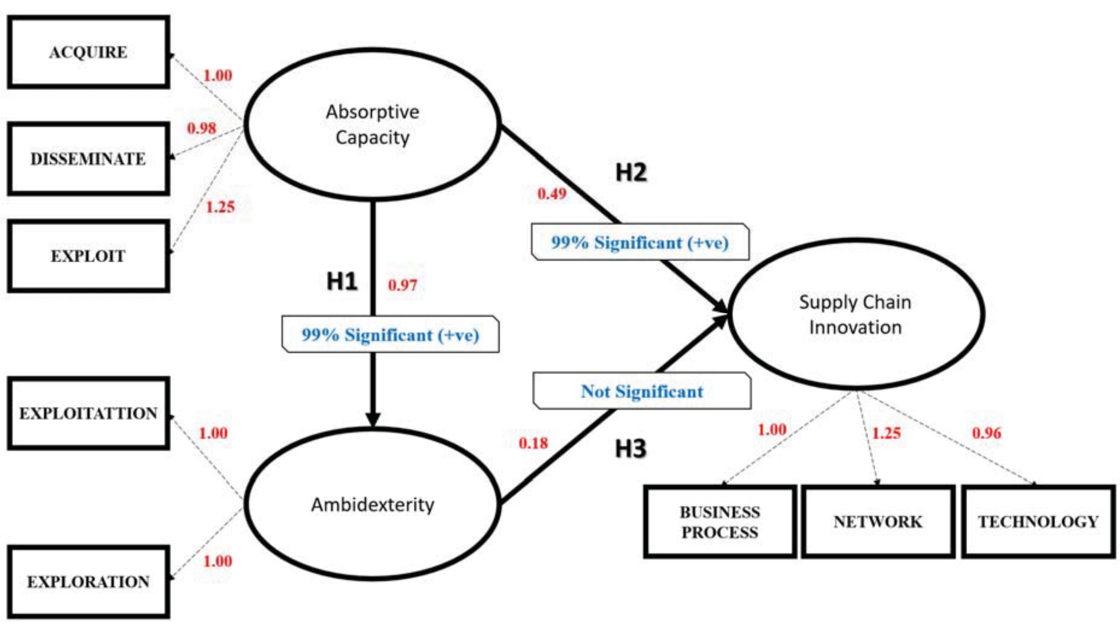
DISCUSSION

The purpose of this research is to investigate the association between absorptive capacity, ambidexterity, and supply chain innovation. Prior research observes the influence of absorptive capacity and ambidexterity on innovation efforts (e.g. Narasimhan & Narayanan, 2013) and underlines that the perception of several innovation is pursued in daily operations as continuous innovation. Bearing in mind, our research extends this understanding toward supply chain innovation. Supply chain innovation is much more complex than product innovation and manages the processes involved in the whole supply chain (Arlbjørn et al., 2011). Supply chain business processes, supply chain networks, and supply chain technologies can be also continuously improved or incorporated in daily operations.

On the whole, our findings show that absorptive capacity positively affects both ambidexterity and supply chain innovation, and ambidexterity does not affect supply chain innovation. First, focusing on the link between absorptive capacity and ambidexterity, the result suggests that absorptive capacity positively affect ambidexterity. Therefore, it could be suggested that firms being ambidextrous can significantly develop their ability to acquire, disseminate, and exploit knowledge to realizing absorptive capacity (e.g. Enkel et al, 2017). In other words, firms that are better able to balance activities pertaining to absorptive capacity are more likely to continuously explore new opportunities/knowledge and increase their firms' learning competencies. Second, the result reveals that there is a positive link between absorptive capacity and supply chain innovation. Extant literature has recognized the link between absorptive capacity and innovation (e.g. product innovation) (Albort-Morant et al, 2018; Berghman et al, 2012). However, it has not explored the link between absorptive capacity and supply chain innovation. Therefore, our research investigates this unexplored relationship and reveals that firms' capability of acquiring, disseminating, and exploiting knowledge increases the practice of supply chain innovation. Supply chain innovation, as mentioned earlier, is a complex process and requires firms to not only acquire new knowledge, but also disseminate and exploit the acquired knowledge to explore new opportunities for supply chain innovation. Borrowing literature on the association between absorptive capacity and innovations in supply chain (Gao et al, 2017; Narasimhan & Narayanan, 2013; Preston et al, 2017), this research confirms that firms' absorptive capacity increases the practice of supply chain business processes, supply chain networks, and supply chain technology. Third, with reference to the link between ambidexterity and supply chain innovation, interestingly, the result indicate that ambidexterity does not affect supply chain innovation. It is believed that being able to innovate in supply chain innovation is determined through the firm's ability to balance processes pertaining to daily operations and development (Slone et al., 2010; Stank et al., 2011). Moreover, extant literature has always focused on the link between ambidexterity and innovation (e.g. Blome et al., 2013; Kortmann et al., 2014; Markides, 2013). For this reason, our research made an attempt to explore the link between ambidexterity and supply chain innovation and found that firms being ambidextrous does not impact supply chain innovation practices. One of the reasons for this result could be that innovation (e.g. product innovation) can be seen explicitly while supply chain innovation includes innovations (e.g. process innovation) that can be realized only through different processes.

The overall result of the research model is presented in Figure 2. This result could be groundwork for further exploration of the proposed model with different dimensions.

Figure 2: Result of the research model



CONCLUSION

This paper has set out to analyze and discuss the relationship between absorptive capacity, ambidexterity and SCI. The paper presents novel data to a research area that lacks empirical data in the context of SCI. Data from the presented survey reveals a positive relationship between absorptive capacity and ambidexterity and thus, a support for hypothesis 1. The companies that acquire, disseminate and exploit supply chain relevant knowledge fully, perceived their ambidexterity to be higher compared to firms that were not as adept in enhancing their absorptive capacity. Furthermore, hypothesis 2 is also supported, i.e. that the firms that perceived their absorptive capacity to be high tended to pursue SCIs. However, the analysis does not find support for a relationship between ambidexterity and companies' efforts with SCIs. This result is surprising and may indicate that SCIs also are being carried out as part of daily operations and not only in development-oriented projects. These results are important for companies to consider if, and when, they are planning transformation processes towards more digitized processes. The results have implications for theory and practice. For theory, this study confirms the positive relationship between absorptive capacity and ambidexterity but not that SCIs are dependent on companies' ambidextrousness. Companies are pursuing SCIs regardless of their perceived ability to balance operational and development-oriented activities in their supply chains. It indicates that such innovations become a reality through development-oriented operations. For practice, these results indicate the importance for companies to search for relevant SCM knowledge and disseminate and exploit this knowledge to implement SCI that assists companies in fulfilling strategic objectives and in maintaining or even improving competitiveness.

The paper is not without limitations. First, the paper is based on perceptions from a single respondent from each company. Future research should include multiple respondents from each company. Second, data are from Danish companies. Future research should include respondents from other countries to make cross country comparisons. Finally, this paper analyzes SCI as one construct consisting of the three components; business processes, network structure and technology. Future research can decompose the construct to analyze the innovations more specific within each of the components.

REFERENCES

Albort-Morant, G., Leal-Rodríguez, A., & De Marchi, V. (2018). Absorptive capacity, relationship learning mechanisms as complementary drivers of green innovation performance. *Journal of Knowledge Management*, 22 (2), 432-452.

Arlbjørn, J. S., & Paulraj, A. (2013). Special topic forum on innovations in business networks from a supply chain perspective: Current status and opportunities for future research. *Journal of Supply Chain Management*, 49(4), 3–11.

Arlbjørn, J. S., de Haas, H., & Munksgaard, K. B. (2011). Exploring supply chain innovation. *Logistics Research*, 3 (1), 3–18.

Autry, C. W., & Griffis, S. E. (2008). Supply chain capital: The impact of structural and relational linkages on firm execution and innovation. *Journal of Business Logistics*, 29 (1), 157-174.

Bellamy, M. A., Ghosh, S., & Hora, M. (2014). The influence of supply network structure on firm innovation. *Journal of Operations Management*, 32 (6), 357-373.

Berghman, L., Matthyssens, P., & Vandenbempt, K. (2012). Value innovation, deliberate learning mechanisms and information from supply chain partners. *Industrial Marketing Management*, 41(1), 27-39.

Blome, C., Schoenherr, T., & Kaesser, M. (2013). Ambidextrous governance in supply chains: the impact on innovation and cost performance. *Journal of Supply Chain Management*, 49(4), 59-80.

Cao, M., & Zhang, Q. (2011). Supply chain collaboration: Impact on collaborative advantage and firm performance. *Journal of Operations Management*, 29(3), 163–180.

Chapman, R.L., Soosay, C., & Kandampully, J. (2003). Innovation in logistic services and the new business model. *International Journal of Physical Distribution & Logistics Management*, 33 (7), 630-650.

Cohen, W.M., & Levinthal, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35 (1), 128-152.

Cooper, M.C., Lambert, D.M. & Pagh, J.D. (1997). Supply chain management: more than a new name for logistics. *The international journal of logistics management*, 8 (1), 1-14.

Davenport, T.H. (1993). *Process innovation: Reengineering work through information technology*. Boston: Harvard Business Press.

Duncan, R. B. (1976). The ambidextrous organization: Designing dual structures for innovation. In: Kilmann, R. H., Pondy, I. R., & Slevin, D. (Eds.). *The management of organization design: Strategies and implementation*. New York: North Holland, 167-188.

Ellram, L.M, & Cooper, M. C. (2014). Supply chain management: It's all about the journey, not the destination. *Journal of Supply Chain Management*, 50 (1), 8-20.

Enkel, E., Heil, S., Hengstler, M., & Wirth, H. (2017). Exploratory and exploitative innovation: To what extent do the dimensions of individual level absorptive capacity contribute?. *Technovation*, 60, 29-38.

Fernhaber, S.A., & Patel, P.C. (2012). How do young firms manage product portfolio complexity? The role of absorptive capacity and ambidexterity, *Strategic Management Journal*, 33 (13), 1516-1539.

Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1), 39-50.

Gao, D., Xu, Z., Ruan, Y.Z., & Lu, H. (2017). From a systematic literature review to integrated definition for sustainable supply chain innovation (SSCI). *Journal of Cleaner Production*, 142, 1518–1538.

Gibson, C.B., & Birkinshaw, J. (2004). The antecedents, consequences, and mediating role of organizational ambidexterity. *Academy of Management Journal*, 47 (2), 209-226.

Grawe, S. J. (2009). Logistics innovation: A literature–based conceptual framework. *The International Journal of Logistics Management*, 20 (3), 360–377.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. & Tatham, R.L. (2006). *Multivariate Data Analysis*. Uppersaddle River: Pearson Prentice Hall.

Hazen, B. T., Overstreet, R. E., & Cegielski, C. G. (2012). Supply chain innovation diffusion: Going beyond adoption. *The International Journal of Logistics Management*, 23 (1), 119–134.

Ho, H., & Lu, R. (2015). Performance implications of marketing exploitation and exploration: Moderating role of supplier collaboration. *Journal of Business Research*, 68 (5), 1026-1934.

Im, G., Rai, A., & Lambert, L. S. (2019). Governance and resource-sharing ambidexterity for generating relationship benefits in supply chain collaborations. *Decision Sciences*, forthcoming.

Jansen, J. J. P., Van Den Bosch, F. A. J., & Volberda, H. W. (2006). Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. *Management Science*, 52 (11), 1661-1674.

Kedia, B. L., & Bhagat, R. S. (1988). Cultural constraints on transfer of technology across nations: Implications for research in international and comparative management. *Academy of Management Review*, 13 (4), 559–571.

Kim, D. Y. (2014). Understanding supplier structural embeddedness: A social network perspective. *Journal of Operations Management*, 32(5), 219-231.

Kortmann, S., Gelhard, C., Zimmermann, C., & Piller, F. T. (2014). Linking strategic flexibility and operational efficiency: The mediating role of ambidextrous operational capabilities. *Journal of Operations Management*, 32(7-8), 475-490.

Kostopoulos, K., Papalexandris, A., Papachroni, M., & Ioanno, G. (2011). Absorptive capacity, innovation, and financial performance. *Journal of Business Research*, 64 (12), 1335-1343.

Kristal, M. M., Huang, X. & Roth, A. V. (2010). The effect of an ambidextrous supply chain strategy on combinative competitive capabilities and business performance. *Journal of Operations Management*, 28, 415–429.

Kwak, D.–W. (2018). Investigating the relationship between supply chain innovation, risk management capabilities and competitive advantage in global supply chains. *International Journal of Operations & Production Management*, 38 (1), 2–21.

Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial Marketing Management*, 29 (1), 65-83.

Lambert, D. M., & Enz, M. G. (2017). Issues in supply chain management: Progress and potential. *Industrial Marketing Management*, 62, 1-16.

Lane, P.J., Koka, B.R., & Pathak, S. (2006). The reifcation of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of Management Review*, 31 (4), 844-863.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, 14 (8), 95-112.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2 (1), 71–87.

Markides, C. C. (2013). Business model innovation: what can the ambidexterity literature teach us?. *Academy of Management Perspectives*, 27(4), 313-323.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zachariah, Z. G. (2001). Defining supply chain management. *Journal of Business Logistics*, 22 (2), 1–25.

Munksgaard, K. B., Stentoft, J., & Paulraj, A. (2014). Value–based supply chain innovation. *Operations Management Research*, 7 (3–4), 50–62.

Narasimhan, R., & Narayanan, S. (2013). Perspectives on supply network–enabled innovations. *Journal of Supply Chain Management*, 49 (4), 27–42.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychological Theory*. McGraw–Hill, New York.

O'Reilly III, C. A., & Tushman, M. L. (2008). Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in Organizational Behavior*, 28, 185–206.

O'Reilly III, C. A., & Tushman, M. L. (2013). Organizational ambidexterity: Past, present, and future. *The Academy of Management Perspectives*, 27 (4), 324-338.

Papke‐Shields, K.E., Malhotra, M.K., & Grover, V. (2002). Strategic manufacturing planning systems and their linkage to planning system success. *Decision Sciences*, 33(1), 1-30.

Patel, P. C., Terjesen, S., & Li, D. (2012). Enhancing effects of manufacturing flexibility through operational absorptive capacity and operational ambidexterity. *Journal of Operations Management*, 30 (3), pp. 201-220.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5), 879-903.

Preston, D. S., Chen, D. Q., Swink, M., & Meade, L. (2017). Generating supplier benefits through buyer-enabled knowledge enrichment: A social capital perspective. *Decision Sciences*, 48 (2), 248-287.

Rothaermel, F. T. & Alexandre, M. T. (2009). Ambidexterity in technology sourcing: The moderating role of absorptive capacity. *Organization Science*, 20 (4), 759–780.

Setia, P., & Patel, P.C. (2013). How information systems help create OM capabilities: Consequents and antecedents of operational absorptive capacity. *Journal of Operations Management*, 31 (6), 409-431.

Slone, R. E., Dittmann, J. P., & Mentzer, J. T. (2010). *The New Supply Chain Agenda: The Five Steps that Drive Real Value*. Boson: Harvard Business Press.

Stank, T. P., Dittmann, P., & Autry, C. W. (2011). The new supply chain agenda: A synopsis and directions for future research. *International Journal of Physical Distribution & Logistics Management*, 41 (10), 940–955.

Stentoft, J., Mikkelsen, O. S., & Jensen, J. K. (2016). Offshoring and backshoring manufacturing from a supply chain innovation perspective. *Supply Chain Forum: An International Journal*, 17(4), 190-204.

Tan, K. H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, 165, 223-233.

Tseng, C., Pai, D., & Hung, C. (2011). Knowledge absorptive capacity, innovation performance in KIBS. *Journal of Knowledge Management*, 15 (6), 971-983.

Tu, Q., Vonderembse, M. A., Ragu-Nathan, T. S., & Sharkey, T. W. (2006). Absorptive capacity: Enhancing the assimilation of time-based manufacturing practices. *Journal of Operations Management*, 24 (5), 692-710.

Vyas, N. (2016). Disruptive technologies enabling supply chain evolution. *Supply Chain Management Review*, 20 (1), 36-41.

Wong, D. T. W. & Ngai, E. W. T. (2019). Critical review of supply chain innovation research. *Industrial Marketing Management*, forthcoming.

Zacharia, Z. G., Nix, N. W., & Lusch, R. F. (2011). Capabilities that enhance outcomes of an episodic supply chain collaboration. *Journal of Operations Management*, 29 (6), 591-603.

APPENDIX A

Supply Chain Innovation (scale developed from Arlbjørn, 2011)

Business Processes
To what extent is your company pursuing innovations in the following supply chain management business processes? (Five-point Likert scale)

- Customer Relationship Management (CRM)
- Supplier Relationship Management (SRM)
- Customer Service Management (CSM)
- Demand Management (DeM)
- Order Fulfilment (OrF)
- Manufacturing Flow Management (MFM)
- Product Development and Commercialization (PDC)
- Returns Management (ReM)

Network Structure
To what extent does your company innovate together with the following supply chain network actors? (Five-point Likert scale)

- Internal functions
- Customers
- Suppliers
- Third party providers (e.g. logistics providers)
- Competitors
- Consultants
- Universities
- Public agencies

Technology

- To what extent does your company work with the following technologies in your supply chain?
- Planning and execution systems (e.g. enterprise resource planning systems, advanced planning systems, material requirements systems)
- Identification systems (e.g. barcodes, radio frequency identification)
- Communication systems (e.g. electronic data interchange, web-based communication tools, mobile communication solutions, cloud technology)
- Analytics technology (e.g. business intelligence, statistics and analytics software, algorithms)
- Electronic marketplaces (e.g. e-portals, e-auctions, supplier collaboration tools)
- Advanced manufacturing technologies (e.g. advanced robotics, 3D printing)
- Advanced materials (e.g. ultra-light or high-strength materials)
- Big data
- Drones

Ambidexterity (representative references Kristal et al., 2010; Levinthal & March, 1993; March, 1991)

Exploitation

- We continuously focus on improving our existing supply chain technologies processes
- We are focusing on reducing operational redundancies in our existing supply chain processes
- We are focusing on developing stronger competencies in our existing supply chain processes
- We have a constant focus on making improvements in our existing supply chain processes

Exploration

- We continually experiment with new supply chain solutions
- We continually develop new supply chain processes
- We are focusing on exploring new competencies in our supply chain processes
- We have a constant focus on novel approaches to make improvements

Absorptive capacity (based on Ho & Lu, 2015; Patel et al., 2012; Setia & Patel, 2013; Tu et al., 2006; Zacharia et al., 2011).

Acquire

- We are committed to look for and develop new ideas in our supply chain
- We have created an environment that encourages identification and development of new and useful ideas
- We frequently look for external sources of new knowledge and skills
- We seek to learn from tracking new market trends in our industry
- We seek to learn from trying out new technologies
- We seek to learn from other supply chain actors (e.g. customers and suppliers)
- We collect frequently operations related information in our supply chain

Disseminate

- We record and store newly acquired supply chain knowledge for future reference
- We share practical experiences with each other
- We periodically meet to discuss consequences of new products and processes
- We have a common language regarding our products and processes in our company
- We have frequent interactions between departments in the firm to acquire new knowledge
- Employees are engaged in cross-functional work
- We have a clear division of roles regarding sharing newly acquired supply chain ideas and knowledge

Exploit

- We constantly consider how to exploit newly acquired knowledge, skills, and technologies in our supply chain
- We quickly recognize the usefulness of new external operational knowledge to existing operational knowledge
- We are proficient in integrating newly acquired knowledge into current ways of doing things

DECISION SCIENCES INSTITUTE

Supply Chain Relationship Building and Knowledge Sharing, Big Data Capability, and Product Improvement

Canchu Lin
Carroll University
Email: clin@carrollu.edu

Anand Kunnathur
University of Toledo
Email: Anand.Kunnathur@utoledo.edu

ABSTRACT

To extend big data research, this study examines Big Data capability's impact on product improvement and explores supply chain relationship building and knowledge sharing as two antecedents of Big Data capability. The model is tested with survey data. Findings of this study provide empirical support to the proposed theoretical model.

KEYWORDS: Supply chain relationship building, Knowledge sharing, Big Data capability, Product improvement

INTRODUCTION

Big data as a hot topic has attracted an increased amount of attention from industry practitioners as well as scholars. Recent research has noted its impact on market and financial performance (Gupta & George, 2016; Wamba et al., 2017). It can be easily understood that market performance contributes to financial performance. But market performance depends on product performance, as it is usually operationalized as how fast new products are introduced to and how well current products sell in the market (Wamba et al., 2017). However, prior and current research has rarely explored how big data contributes to product performance, especially existing product performance. Yet, incremental product improvement is critical to firm performance, as it helps cost reduction and repositioning (Holahan, Sullivan, & Markham, 2014). Further, because of brand commitment, customers prefer incremental upgrade to an existing product to a radical new product (Kirshner, Levin, & Nediak, 2017). Given these benefits, it is imperative to explore how big data enhances incremental product improvement.

Besides capturing big data's impact on firm performance, prior and current research has also offered another insight that organizational practices surrounding big data should be conceptualized as dynamic capabilities, which helps to explicate why there are cases of success as well as failure in big data practices (Braganza et al., 2017; Chen, Preston, & Swink, 2015; Gupta & George, 2016; Wamba et al., 2017). However, it is less known what factors contribute to the development of such big data related dynamic capabilities in firms.

Further, related to that gap is a question of whether such contributing factors exist outside the firm. Firms are embedded in supply chains. There are reported cases of how big data enhances supply chain performance in the literature such as forecasting cycle time (Wang & Zhang, 2016), production or service scheduling (Barton & Court, 2012), and demand forecasting and inventory management (McAfee & Brynjolfsson, 2012; Waller & Fawcett, 2013b). Yet few studies have examined whether supply chains might be an important source of resources that can facilitate the development of big data dynamic capabilities.

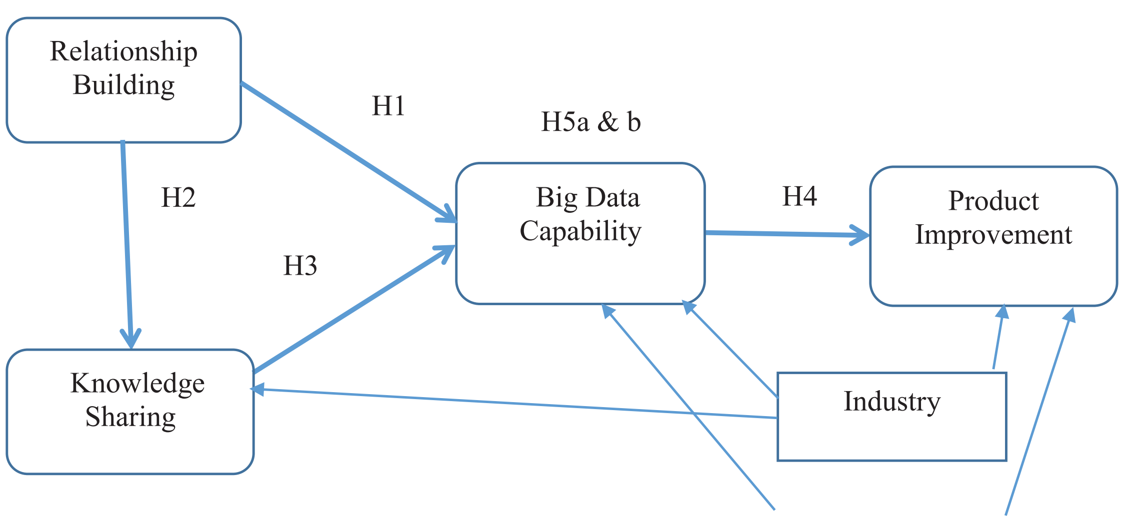
To fill these research gaps, this study attempts to empirically investigate big data's role in a firm's incremental product improvement performance and the influence of supply chain dynamics on this role. Supply chain dynamics in this study refers to two concepts: supply chain relationship building (RB) and knowledge sharing (KS) in the supply chains. Consistent with the literature, this study conceptualizes big data initiatives or practices in organizations as a dynamic organizational capability, which is named as Big Data capability (BDC). Based on the organizational capabilities perspective, we first argue that BDC contributes to a firm's performance in product improvement. Then drawing on the theories of dynamic capabilities, resource dependence, and social capital, we articulate an argument that supply chain dynamics (RB and KS) contributes to the development of BDC. We further show that BDC mediates the effects of supply chain dynamics on product improvement.

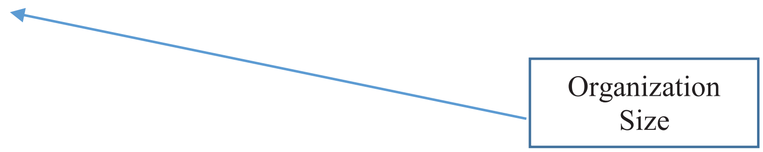
While accomplishing the above-mentioned objectives, this study contributes to the literature in three ways. First, it will show that when big data's initiatives and practices in firms develop into a BDC, it will enhance firms' product improvement performance. Second, this study will extend the literature by demonstrating that RB and KS contribute to the development of BDC. Our third contribution lies in revealing that BDC mediates the effects of RB and KS on product improvement. Findings from this study will help to renew our understanding of dynamic capabilities, especially BDC, as they will show that they are developed not in a vacuum but with external assistance despite that they are conceptualized as firm specific and path-dependent (Teece, Pisano, & Shuen, 1997; Teece, 2007).

LITERATURE REVIEW

The theoretical framework of this study is built on a strong rationale. To address environmental changes and keep their competitive advantage sustainable, firms have to develop dynamic capabilities (Girod & Whittington, 2017; Peteraf, Di Stefano, & Verona, 2013; Schilke, 2014), which help them to manage evolutionary changes and make incremental improvements in their products (Tushman & O'Reilly, 1996). But to develop dynamic capabilities needs resources including knowledge, which often times resides in buyer-supplier relationships (Villena, Revilla, & Choi, 2011). Thus, our research model is constructed as follows: supply chain relationship building and knowledge sharing contribute to Big Data capability, which, then enhances product improvement (see Figure 1).

Figure 1: Conceptual Model





Big Data Capability

According to the dynamic capabilities perspective, to develop and maintain their competitive advantage, firms should develop dynamic capabilities (Collins, 1994; Teece et al., 1997; Teece, 2007). Besides helping to perform the basic functions of the firm (Collins, 1994; Winter, 2003), dynamic capabilities are managerial (Kor & Mesko, 2013), concerned with change, and more importantly, of strategic insight (Collins, 1994; Teece et al, 1997, Teece, 2007). Further, dynamic capabilities are unique to each firm and strongly path-dependent in firms (Teece et al., 1997; Teece, 2007). A dynamic capability involves three components: sensing, seizing, and reconfiguring (Teece, 2007). Accordingly, sensing means to identify and shape opportunities and threats, seizing is to deploy and utilize resources and assets to address those opportunities, and reconfiguring is to transform, redeploy, and redevelop organizational competencies for continuous renewal. Reviewing some case studies, Lin, Kunnathur, and Li (in press) showed that all these three components are identifiable in big data practices.

More importantly, development of a dynamic capability dealing with information technology issues, is an emergent process: employees acquire skills, knowledge, and competencies in their interactions with material resources, which are further integrated, organized, and utilized through organizational processes governed by structural rules; these integrated and synergized skills, knowledge, and competencies, aligned with firms' strategy, then grow into dynamic capabilities (Peppard & Ward, 2004). Viewed from this perspective, a big data related dynamic capability can be developed in firms, which consists of skills, knowledge, and behavioral resources and competencies regarding big data practices, developed in processes of dealing with big data by using material resources such as IT hardware and software, to accomplish strategic objectives (Chen et al., 2015; Gupta & George, 2016; Lin, Kunnathur, & Li, in press; Wamba et al., 2017).

Big data material resources include raw data and technological tools that are used to collect, store, and analyze data, and present analysis results to a right audience. Big data as raw data exists in multiple formats such as numerical, text, visual, and audio. Big data is also carried in structured as well as unstructured formats (Chen, Chiang, & Storey, 2012). Moreover, big data is produced rapidly in multiple sources. Further, big data flows out in big volumes. These constitute the three basic characteristics of big data: volume, velocity, and variety (McAfee & Brynjolfsson, 2012). Technological tools include both hardware and software, including but not limited to IT infrastructure, computing equipment, and all software used for collection, storage, and analysis of big data such as Hadoop, NoDQL and SAS (Chang, Kauffman, & Kwon, 2014). Some of these material resources are possessions of a firm, while others may only be acquired externally. The non-material resources are individual employees' experiences, skills, knowledge of using and handling the material resources, and behaviors such as managing information security, and policy compliance. These non-material resources can also be acquired both internally and externally. Some typical big data non-material resources include data analytics, text analytics, web analytics, and mobile analytics (Chen et al., 2012).

In the spirit of the dynamic capabilities theory (Teece et al., 1997; Teece, 2007; Teece, Peteraf, & Leih, 2016), development of a big data capability is an emergent process. Non-material resources are then turned into big data competencies when they are used in organizational processes where the technical skills, experiences, and knowledge of big data interact with firm-based skills, knowledge, and behaviors such as communication, relationship building, and rule compliance. These big data competencies are shown in multiple areas such as enhancement of big data or IT infrastructure, development of an IT and specifically big data management capability, and employee acquisition of soft skills to handle big data related tasks (Wamba et al., 2017). As these skills, knowledge, and behaviors are firm specific, the big data competencies are therefore firm specific. Finally, when big data competencies are aligned with the firm's business strategy, and shaped by the firm's big data related investments, they grow into its Big Data capability. Thus, *Big Data capability* is defined as a bundle of socially complex routines and processes in which a firm's accomplishment of integrating its material and non-material resources to develop socially complex routines and processes of identifying, collecting, storing, and analyzing big data for achieving its strategic as well as operational objectives. It involves an integration of a firm's dataset, toolset, skillset, and mindset (Pigni,Piccoli, & Watson, 2016). While the dataset and toolset are material resources, skillset non-material resources (skills, experiences, and competencies), mindset indicates its strategic positioning and maneuvering of all these resources.

Product Improvement

Product improvement is a construct of firm performance in this research model. It includes overall product cost reduction, capital investment reduction, lowering of overall business risks (Lemon, 1991), and enhanced product functionality and customer acceptance (Pisano & Wheelwright, 1995). Further, according to McGinnis and Mele Vallopra (1996), established products can benefit from the cumulative effect of gradual process improvement over time. Thus, *product improvement* in this study refers to making changes to a current or existing product for the purpose of meeting customers' needs and creating more business value. Current or existing product performance, unlike new product development, represents a firm's competence-enhancing innovation capability (Tushman & Anderson, 1986), and reflects, to some extent, incremental innovation performance (Christensen, 1997), both pivotal to firms.

Supply Chain Relationship Building

In the process of developing Big Data capability, firms must practice organizational learning, which is performed both internally and externally (Argote et al., 2003; Salomon & Martin, 2008). Supply chains are a good site for external organizational learning. There is strong evidence showing that learning from supply chain partners benefits firms (Azadegan & Dooley, 2010). But effectiveness of organizational learning among supply chain partners depends on the quality and strength of their relationships (Cheung, Myers, & Mentzer, 2010). This calls for relationship building among these partners.

Another driving force behind supply chain relationship building is search for or acquisition of resources. Firms rarely possess all kinds of resources and can rarely develop new resources in an isolated vacuum in their process of developing their Big Data capability. Thus, to get access to resources, firms have to develop relationships with other firms. A supply chain is a candidate site for doing so. This idea has been well elaborated and strongly supported in the supply chain management, especially supply chain relationship building research literature. For example, Ireland et al. (2002) showed that inter-firm relationships mainly serve as a means to acquire external resources. These relationships are especially helpful for facilitating acquisition of resources necessary for filling a particular ‘resource gap’ (Grant, 1991), i.e., the gap between a firm’s strategic goals and its current resource endowments (Mathews, 2003). According to Leonard-Barton (1992), firms enter into business relationships with their supply chain partners mainly for strengthening functions beyond their own capabilities and/or core competencies. Given such economic benefits of buyer-supplier relationships, firms are motivated to enhance inter-organizational communication, and develop relational competency (Paulraj, Lado, & Chen, 2008), which are active means of relationship building.

Not just constituting resources, buyer-supplier relationships are also an important source of competitive advantage (Autry & Golicic, 2010; Krause, Handfield, & Tyler, 2007). This is consistent with the goal of Big Data capability, which, as a dynamic capability, is to develop and maintain a competitive advantage (Teece et al., 1997; Teece, 2007). Although definitions of inter-firm relationship vary across fields of research and studies even within supply chain management research, past studies do point to some common attributes of a positive relationship. These include: trust, commitment, collaboration, and long-term orientation (Griffith, Harvey, & Lusch, 2006; Johnston, McCutcheon, Stuart, & Kerwood, 2004; Monczka, Petersen, Handfield, & Ragatz, 1998; Tangpong, Michalisin, & Melcher, 2008). Thus, *supply chain relationship building* in this study is defined as SC partners’ mutual engagement to develop a relational tie that is characterized by trust, commitment, collaboration, and long-term orientation.

Knowledge Sharing in Supply Chains

As data analysis, a core part of BDC, is an integral part of knowledge creation (Davenport & Prusak, 1998), Big Data capability must be closely related to knowledge management. Knowledge management refers to a conceptual framework that treats an organization or a network of organizations as a system where people, processes, resources are integrated to achieve sustainable results by improving performance through learning and innovation (Seleim & Khalil, 2011; Wang, 2011). Knowledge management as a broad concept covers multiple processes such as knowledge acquisition, creation, retention, and transfer (Argote, McEvily, & Reagans, 2003). In this study, we focus on one element of knowledge management: knowledge sharing, as it is critical to product improvement (Su, Ahlstrom, Li, & Cheng, 2013). We treat knowledge sharing as a process in which both a focal firm and its supply chain partners rely on each other closely to share knowledge together. Big Data capability needs knowledge as its input in its process of producing new knowledge. Just as supply chains are a site where resources reside, they are also important sources of knowledge as well. For this reason, this study treats knowledge sharing as another contributor to the development of Big Data capability in the overall research model.

Once knowledge is produced or imported in supply chains, it must be transferred to where it is needed. In prior research, knowledge sharing, especially, conceptualized along the knowledge-based view (Grant, 1996), mostly refers to an internal integration mechanism facilitating the dissemination and synthesis of knowledge scattered in individuals as well as organizational units through organizational processes and routines (Kale & Singh, 2007; Verona, 1999). Yet, this study examines knowledge sharing in the supply chain context. Thus, *knowledge sharing in supply chains* is defined as the intensity of supply chain partners’ exchange of their knowledge (about customers, market, products, manufacturing, and technologies, etc.) with each other (Ho & Ganesan, 2013).

THEORETICAL DEVELOPMENT: HYPOTHESES

Linkage between SC Relationship Building, Knowledge Sharing, and Big Data Capability

Potentially useful knowledge and expertise always reside outside the firm (Bercovitz & Feldman, 2007). We argue that supply chain RB and KS facilitate the development of BDC because it lubricates acquisition of resources such as knowledge, an important input for developing a dynamic capability. This argument is made based on the following theories: dynamic capabilities, resources dependence, and social capital theories. First, according to Teece (2007), to develop dynamic capabilities, firms must do sensing, which is to identify and shape opportunities. More specifically, firms must invest in research activity, investigate customer needs and technological possibilities, and closely watch and monitor possible changes in the environment, which include, for example, changes in industry structure, instability of market demand (Schilke, 2014), and possible supplier and competitor responses (Teece, 2007). Information and knowledge is critical to effective analysis of these changes in the environment and more importantly, detection of opportunities. Firms can certainly rely on their own research and learning to create knowledge and opportunities. But exogenous knowledge can also open up opportunities (Teece, 2007). This suggests that firms, when developing their dynamic capabilities, even though unique and firm specific, must have an external orientation, a mindset of opening themselves to the outside, without which sensing is impossible to complete. Thus, the dynamic capabilities theory suggests that in their process of developing their own dynamic capabilities, firms are motivated to interact with other firms to gain external knowledge for opening up business opportunities. Such interactions happen in their supply chains. Thus, supply chain RB would enhance knowledge sharing and acquisition that is critical to the development of dynamic capabilities.

Second, in a similar fashion, the theory of resources dependence implies that firms must connect with other firms for resources such as knowledge when developing their dynamic capabilities. This theory is rooted in the open system theoretical tradition that maintains that firms must engage in exchanges with their environment to gain resources (Scott, 1987). To survive and prosper, firms must have critical resources. This is especially so when competition is strong. A firm may want to gain enough market power to deter its competitors from moving in or hope to create a skill or resource gap between the firm and its competitors. In doing either of these two, the firm does not perceive that it has all necessary resources to keep it in a competitive position. Thus, to gain and then enhance their competitive advantage, firms must acquire critical resources, or, using the theory's terms, make themselves in a state of resource dependence (Barringer & Harrison, 2000). To reduce such a resource dependence, firms can form an alliance or participate in interorganizational relationships with other firms (Mitchell & Singh, 1996). Firms enter into partnership relationships because they can take advantage of their complementary resources (Barringer & Harrison, 2000). This explication of why firms participate in interorganizational relationships from the resource dependence perspective can certainly apply to the case of BDC. When developing their BDC, firms need critical resources, which can be acquired by forming an alliance with their suppliers and customers.

A third theory that can be used to explain a perceived relationship between RB and BDC is the social capital theory. Based on this theory, social relationships help to build and gain social capital, which refers to a valuable asset that helps to gain access to resources made available through the relationships (Coleman, 1990). Social capital has three dimensions: cognitive, relational, and structural (Nahapiet & Ghoshal, 1998). The cognitive dimension represents shared meaning and understanding between actors in a relationship (Nahapiet & Ghoshal, 1998). It also helps to diminish misunderstandings, open discussion, and promote frequent communication among them (Tsai & Ghoshal, 1998). Subsequently, a shared vision and congruent goals will develop, which then help different actors in the relationship to integrate knowledge (Inkpen & Tsand, 2005). Additionally, these shared meaning, understanding, vision, and congruent goals would foster information sharing and resource exchange (Atuahene-Gima & Murray, 2007).

Next, the relational dimension refers to trust, friendship, goodwill, respect, and reciprocity developed in the interactions of the actors that help to sustain the relationships (Nahapiet & Ghoshal, 1998). This dimension entails the foundation of the relationship built over time (Villena, Revilla, & Choi, 2011). These ingredients of relational social capital allow actors to reliably expect to obtain and use the information and resources made available in the relationships (Nahapiet & Ghoshal, 1998). Lastly, the structural dimension refers to the embeddedness of actors in a social system or relationships, or the tie strength that positively relate to accessibility to information, knowledge, and other types of resources (Kemper, Schilke, & Brettel, 2013). As an integration of these three dimensions, social capital in general bring in information and resource benefits to organizations that would enhance their dynamic capabilities (Kemper, Schilke, & Brettel, 2013). This is done through employee behaviors.

Previous research (e.g., Lopez-Cabrales, Valle, & Herrero, 2006) shows that core employees play an important role in developing dynamic capabilities. More specifically, one group of such core employees, top managers, bring in social capital to their firms, which then facilitates effective access to the knowledge and resources necessary for developing superior organizational dynamic capabilities such as marketing and R&D capabilities (Kemper, Schilke, & Brettel, 2013). The literature provides evidence supporting the relationship between cognitive social capital and dynamic capabilities. For example, Tripsas and Gavetti (2000) showed that managerial cognitive representations may play a central role in terms of constraining organizational behavior, and ultimately, the development of a firm’s capabilities.

Similarly, relational social capital would facilitate the development of dynamic capabilities as well. Gibbons and Henderson (2012) show that how well top managers help to develop or diffuse organizational capabilities depends on their capability to build and refine relational contracts, which is a managerial practice aiming to seek collaboration with informal communication characterized by credibility and clarity. As agents representing their firms, top managers build and gain social capital in their interactions with external entities. Thus, relational contracts also work in their external communication, especially in interorganizational relationships. For example, Gulati and Nickerson (2008) found that high levels of interorganizational trust, an attribute of relational contracts or social capital, lead to choice of an informal but less costly governance mode for the interorganizational network and more importantly, tend to enhance the network’s exchange performance. As social capital is built in collaborative buyer-supplier relationships (Autry & Griffis, 2008; Cousins, Handfield, Lawson, & Petersen, 2006; Cousins & Menguc, 2006; Krause, Handfield, & Tyler, 2007; Lawson, Tyler, & Cousins, 2008; Min, Stephen, & Chen, 2008), the social capital theory suggests a strong link between supply chain RB and BDC.

H1: Supply chain relationship building is positively related to the development of Big Data capability.

Linkage between Supply Chain Relationship Building and Knowledge Sharing in Supply Chains

Most of the theories earlier reviewed share a central proposition that investment in relationship building, knowledge exchange, and combining resources lead to supernormal profit (Krause, Handfield, & Tyler, 2007). This suggests that RN and KS are inherently related. We argue that among those theories, the social capital theory, is most relevant to our proposition that there is a connection between supply chain RB and KS among supply chain partners. As the cognitive dimension of social capital contributes to diminishing misunderstanding and enhancing shared meaning, vision, and congruent goals (Nahapiet & Ghoshal, 1998; Tsai & Ghoshal, 1998), it facilitates information sharing, and knowledge exchange (Atuahene-Gima & Murray, 2007) and integration (Inkpen & Tsand, 2005). Similarly, the relational dimension, i.e., goodwill, friendship, and trust developed between actors of the relationship, makes it more likely for them to share information and knowledge (Nahapiet & Ghoshal, 1998). Likewise, the structural dimension suggests that the great number of ties and their strength enable actors with social capital to expand their access to information and knowledge (Kemper, Schilke, & Brettel, 2013).

Empirical research in the area of buyer-supplier relationships tends to support these theoretical implications. An extensively cited example is the case of Toyota, where information exchange and knowledge sharing are highly promoted in buyer-supplier relationships (Dyer & Nobeoka, 2000). Further, supplier trust leads to shared planning activities (in which both buyers and suppliers are involved) such as joint decision making (Johnston, McCutcheon, Stuart, & Kerwood, 2004). Joint decision making entails information sharing and knowledge exchange. Similarly, Cheung, Myers, and Mentzer (2010) showed that interorganizational factors such as trust, complementarity, and compatibility, ingredients of social capital, contribute to relationship learning, which involves information sharing, joint sense making, and knowledge integration. Corsten, Gruen, and Peyinghaus (2011) even suggested that to build relationships, and more specifically, to foster mutual trust, which in return nurtures information exchange and knowledge sharing between SC partners, mutual identification should be encouraged. These research findings give us support to articulate the following hypothesis:

H2: Supply chain relationship building is positively related to knowledge sharing in supply chains.

As shown above, all the three theories of dynamic capabilities, source dependence, and social capital suggest that positive supply chain relationships contribute to the development of dynamic capabilities. Further, the above review shows that both the dynamic capabilities and source dependence theories suggest that resources are integral to the development of dynamic capabilities. Knowledge is an important resource. Moreover, the review of the social capital literature indicates that social capital facilitates information and knowledge sharing among supply chain partners. Combining all these findings, we propose the following hypotheses:

H3: Knowledge sharing in supply chains is positively related to the development of Big Data capability.

Linkage between Big Data Capability and Product Improvement

Distinct capabilities can be developed in different areas of business activities that enable firms to generate competitive advantage (Eisenhardt & Martin, 2000; Schilke & Goerzen, 2010). This observation certainly applies to the big data area. Organizational capabilities in general positively impact performance. After a meta-analysis of 115 studies, Karna, Richter, and Riesenkampff (2015) found that both ordinary and dynamic capabilities influence financial performance in both stable and changing environments. Further, dynamic capabilities such as marketing and R&D capabilities differentiate firms in organizational performance (Kemper, Schilke, & Brettel, 2013). Big Data capability inherently involves both such capabilities as it deals with analyzing the market as well as developing tools, technologies, and competencies to analyze big data, similar functions of R&D. Most importantly, as it is revealed in previous reviews of case studies, Big Data practice helped to improve functional performance, enhance service quality, and reduce cost, etc (Ylijoki & Porras, 2016). These are important instances of product improvement based on our definition. Thus, to extend such observations from case studies, we propose the following hypothesis:

H4: Big Data capability is positively related to product improvement.

Mediation Effect of Big Data Capability

As it is shown early, supply chain relationships can bring firms social capital, which constitutes resources. These resources are critical to firm performance such as product improvement. However, according to the resource-based view (Barney, 1991) and the dynamic capabilities theory (Teece, 2007; Teece et al., 1997), when resources turn into organic elements of a dynamic capability, they help to boost firm performance. This suggests that dynamic capabilities tend to enhance the effect of resources stemming from social capital on firm performance. As demonstrated above as well, knowledge, like social capital, is also a resource. Thus it can be reasoned that dynamic capabilities also enhance the effect of knowledge sharing on firm performance. Empirical research has already shown that a path of social capital-dynamic capabilities-performance exists. More specifically, dynamic capabilities, i.e., marketing and R&D capabilities, fully mediate the relationship between social capital and performance (Kemper, Schilke, & Brettel, 2013). This leads us to pose the following two hypotheses:

H5a: Big Data capability mediates the relationship between supply chain relationship building and product improvement.

H5b: Big Data capability mediates the relationship between knowledge sharing in supply chains and product improvement.

METHODS

Questionnaire Development and Measures

The measurement scale for the constructs of knowledge sharing is adapted from existing scales used in previous studies, whereas the scales measuring Big Data capability, relationship building, and product improvement were developed in this study. The process for developing the three scales measuring Big Data capability, relationship building, and product improvement involved four phases: 1) item generation, 2) pre-pilot study, 3) pilot study, and 4) large-scale data analysis and instrument validation. Churchill's (1979) approach for developing and testing new scales was followed in the development of these scales. The first phase was a thorough review of past studies from which the content domain of the constructs was determined for this current research project.

The pre-pilot study was conducted with ten academicians who published work on the topic and eight industry experts with experience in Big Data. Feedback and comments from the academicians were used to check face validity of the newly drafted measurement items and improve the items. Using an adapted version of the Q-sort approach (Churchill, 1979), we asked the eight industry experts to read the construct and its dimensions, their definitions, and the items measuring them, and then match the scale items with the construct dimensions that they think the items measure well. This is to assess substantive validity of the scale items. Items with a 0.5 or higher value of substantive validity were retained. Some items were reworded based on comments and feedback from both the academicians and industry experts.

After the pre-pilot study, a small-scale pilot study was conducted to further improve the newly developed measurement scales. The pilot study questionnaire was posted on a research website (www.qualtrics.com). A list of potential participants was created using the LexisNexis database, with the following information: position title, specialty, mail address, telephone number, email address, firm size, SIC code, business description, and sales/revenue information. In the process of creating the list, a variety of job titles were used as key words such as: “Chief Operations Officer”, “VP Manufacturing”, “Manufacturing Manager”, “Plant Manager”, “Manufacturing Director”, “Production Manager”, and “Plant Operations Manager”. “Manufacturing” was used as the limiter in the search of the firm representatives to form the entries on the list. A total of 45 responses were obtained in this pilot study, of which only 42 were complete and usable. The 42 complete responses were analyzed using SPSS. Further deleting items was based on the results of this pilot study (Cronbach’s alpha, Corrected Item-Total Correlation (CITC), and factor loading values). The scales were then further refined based on the pilot study results. The fourth phase was a large-scale survey questionnaire administration.

The items measuring the construct of Big Data capability were generated based on the results of a thematic analysis of qualitative interviews which are reported in another paper, recent discussions of Big Data in the popular press, business journals for practitioners, and academic business literature, especially the works of Chen et al. (2012), McAfee and Brynjolfsson (2012), Chang et al. (2014), Peppard and Ward (2004), and Teece et al. (1997). The final version of the Big Data capability scale has a total of 16 items measuring Big Data operations, updating IT infrastructure, advanced analytics, and strategic uses of Big Data. However, the confirmatory factor analysis of the large-scale survey data used in this study yielded a single dimension structure for the Big Data capability construct. But the results also supported retaining all the 16 items in the scale. The relationship building (RB) construct has eight original items, measuring the four dimensions of trust, commitment, collaboration, and long-term orientation of the relationships with buyers and suppliers. The items developed for this study are based on previous studies that measured these dimensions, including Griffith et al. (2006) for long-term orientation, Johnston et al. (2004) for trust (they measured trust with two variables (dependability and benevolence)), Monczka et al. (1998) for trust, collaboration (coordination), and commitment. Each of these previous studies focused only on one or two of these four dimensions of relationship building, not the entire construct. They used multiple items to measure each of these four dimensions. If these items are all included in our measure for the relationship building construct, it would be an excessively long list of items. Further, the names and meanings of these four dimensions each are already familiar to respondents. Thus, we chose to use the exact words of these dimensions (trust, commitment, collaboration, and long-term orientation) in the items to measure their levels in the relationships with both buyers and suppliers. The product improvement (PI) construct originally has three measurement items. The scale measuring knowledge sharing is adapted from Ho and Ganesan (2013). The measurement items for all the constructs are listed in the Appendix.

In addition, consistent with previous studies on Big Data capabilities (see, e.g., Gupta & George, 2016), we treat industry and organization size as two control variables in this study. Industry is a nominal variable. Based on the responses from survey participants, the following industries were identified: manufacturing, IT, finance, automotive, telecommunications, construction, logistics, retail, media, healthcare, aerospace, utilities, chemical, consulting, and consumer products. Organization size, which is measured by the number of employees, ranges from 10 to 100,000, with a mean of 4603.

Respondents and Data Collection Procedure

Before the large-scale data collection process began, attention and efforts were directed on determining respondents for this study. It was decided that respondents for this study must be employees with enough knowledge about the phenomenon under study, i.e., Big Data practices in their organizations. Thus, the key informant approach was followed in the data-collection process (Phillips & Bagozzi, 1986). Based on interviews with industry experts in the instrument development and initial validation stage of this study, we learned that employees with high-level managerial positions in information technology, operations, and business analytics were potential key informants for this study. Thus, employees with the following positions were contacted for an invitation of participating in this study as survey respondents: CEO, COO, CIO, CTO, VP of Manufacturing, Operations Manager, IT Manager, Supply Chain Manager, Manufacturing Manager, Manufacturing Director, Plant Manager, Production Manager, and Analytics Manager.

Given the difficulties of locating and gaining access to enough key informants, we outsourced data collection to Qualtrics.com, a commercial research firm, which has a large list of business panel members who represent different firms in different industries. Email invitations were sent to its panel members by Qualitrics.com with the statement of the purpose of the study. They were directed to an online survey posted on Qualtrics.com's website.
It becomes increasingly popular among academicians to outsource data collection to commercial research companies. Yet serious concerns and challenges associated with such a practice emerged. All such concerns and challenges point to data quality. To ensure data quality, we followed recommendations from Schoenherr, Ellram, and Tate (2015). More specifically, we took the following measures to ensure data quality. First, we informed Qualtrics.com representatives of the purpose of this research project, and the detailed requirements for potential respondents. These requirements helped Qualtrics.com to locate a sample of potential respondents. Second, following Schoenherr et al's (2015) practice, we did not provide an initial statement at the beginning of the survey showing the purpose of the project so as to avoid the possibility that respondents form a pattern in their responses that reflect what they think we want as shown in the purpose statement. Third, we used screening questions at the beginning of the survey to ensure that only the respondents with the required characteristics take the survey and all other potential respondents would be disqualified. One such screening question, for example, was whether their organization has Big Data practice. Fourth, Qualtrics.com implemented some attention filters throughout the survey questionnaire to help to ensure that respondents focus their attention on the questions. Also, Qualtrics.com set a mechanism that automatically removes responses that were done in less than 9 minutes. This technical add-in helped to avoid those "speeders" (Schoenherr et al., 2015). Fifth, locations of all respondents and the ip addresses were recorded so as to check whether there were repeat respondents. Sixth, a "Not Applicable" response option was provided to each questionnaire item except for the open-ended and screening questions.

A total of 2700 panel members that were identified as managers in the manufacturing, operations, and IT industry, were targeted as potential respondents for this survey. With members whose contact information was not clear removed, a new total of 2200 remained in the list and an email was sent to them to solicit their participation in this study. A total of 273 responses was obtained. After a careful examination, 22 problematic responses were deleted. Thus, the final sample for this study was 251 responses. The response rate for this study was 12%.

Measurement Model

The measurement model consists of a confirmatory factor analysis that covers four constructs: Big Data capability, relationship building, knowledge sharing, and product improvement. All these constructs were modeled with reflective measurement items. Items were examined based on loadings and the overall model fit indicators such as CFI, TLI, NFI, and RMSEA (Byrne, 2010). Final measurement items, including their mean, standard deviation, loading, critical ratio, standard error, and R^2 values are shown in Table 1.

Table 1: Confirmatory Factor Analysis: Construct Descriptive Statistics, Factor Loadings, and Cronbach α Values

Construct	Mean	Standard Deviation	Loading	CR	SE	R^2
Big Data Capability (α = .92)						
BDC1	5.38	0.66	0.57	8.09	0.08	0.32
BDC2	5.23	0.78	0.58	9.18	0.08	0.34
BDC3	4.32	0.81	0.51	7.44	0.09	0.26
BDC4	5.21	0.87	0.58	8.30	0.10	0.34
BDC5	4.20	0.83	0.66	9.14	0.10	0.44
BDC6	5.16	0.87	0.66	9.30	0.10	0.44
BDC7	5.13	0.95	0.67	9.27	0.11	0.45
BDC8	5.12	0.90	0.63	9.46	0.11	0.40
BDC9	5.12	0.92	0.67	8.96	0.11	0.45
BDC10	4.99	1.01	0.61	8.63	0.13	0.37
BDC11	5.06	0.99	0.70	9.23	0.13	0.49
BDC12	4.94	1.09	0.58	8.20	0.13	0.34
BDC13	5.14	0.97	0.65	10.20	0.10	0.42
BDC14	5.09	1.02	0.71	10.00	0.12	0.50
BDC15	5.19	0.94	0.65	10.16	0.10	0.42
BDC16	5.05	0.90	0.67	10.10	0.10	0.45
Relationship Building (α = .93)						
RB1	5.34	1.01	0.84	14.50	0.06	0.71
RB2	4.28	1.02	0.80	14.35	0.06	0.64
RB3	4.21	1.03	0.78	13.80	0.07	0.61
RB4	5.14	0.92	0.78	13.83	0.06	0.61
RB5	4.37	0.92	0.76	15.20	0.05	0.58
RB6	5.31	0.96	0.82	14.01	0.06	0.67
RB7	5.21	1.02	0.78	13.11	0.07	0.61
RB8	5.22	0.92	0.79	12.93	0.06	0.62
Knowledge Sharing (α = .78)						
KS1	4.98	0.97	0.58	7.83	0.12	0.34
KS2	5.19	0.81	0.65	7.40	0.13	0.43
KS3	4.21	0.86	0.69	7.89	0.14	0.48
KS4	4.28	0.90	0.58	6.35	0.14	0.34
KS5	4.32	0.82	0.65	7.39	0.13	0.42

KS6	4.24	0.74	0.54	6.63	0.11	0.29
Product Improvement (α = .73)						
PI1	5.15	0.93	0.67	10.14	0.10	0.45
PI2	5.12	0.89	0.70	10.00	0.10	0.49
PI3	5.11	0.90	0.71	10.14	0.10	0.50

Assessment of validity and reliability is based on recommendations by Anderson and Gerbing (1988). Reliability was assessed by looking at the Cronbach alpha value. The alpha values for these constructs range from 0.73 to 0.93, all above 0.70 (see Table 1). Furthermore, composite reliability serves a similar purpose but is considered a more rigorous reliability measure in structural equation modeling (Raykov, 1998). The composite reliability values of all the constructs are close to and even greater than the threshold value of 0.70 (See Table 2) (Bagozzi & Yi, 1988; Gefen, Straub, & Boudreau, 2000; Nunnally, 1978). These results indicate that items measuring the same construct are highly interconnected and that the constructs have good construct reliability.

Table 2: Composite Reliability, Correlations, and Average Variance Extracted

	Composite Reliability	BDC	KS	RB	PI
BDC	0.68	0.41			
KS	0.68	0.67**	0.38		
RB	0.66	0.43**	0.34**	0.61	
PI	0.73	0.68**	0.61**	0.37**	0.48

**: Correlation is significant at the 0.001 level (2-tailed).
Average variance extracted values are printed in diagonal.

Similarly, construct validity was assessed in several ways. First, content validity was ensured by designing, developing, and refining measurement items based on literature reviews, interviews of practitioners and academics knowledgeable in the content domain, and cross checking with these practitioners and academics. Next, convergent validity was assessed by looking at each indicator's loading. As seen from Table 1, loadings of all measurement items are higher than 0.5, the cutoff point traditionally used in the literature.

Traditionally, discriminant validity was assessed by examining the cross-loadings of the items measuring each construct, and whether the square roots of the average variance extracted (AVE) for each construct are greater than the corresponding correlation coefficients (Fornell & Larcker, 1981). Recently, Hensler, Ringle, and Sarstedt (2015) identified flaws of the Fornell and Larcker's (1981) method and proposed using a new criterion called the heterotrait-monotrait ratio (HTMT). This ratio is based on the average of the correlations of indicators across constructs measuring different concepts relative to the average of the correlations of indicators within the same construct. Based on this approach, the HTMT ratio below 0.85 demonstrates sufficient discriminant validity. We ran this test on all the four constructs, and the HTMT values were all below 0.85, suggesting that they all have sufficient discriminant validity.

Additionally, the uni-dimensionality of the constructs was established by confirmatory factor analysis, with all item loadings well above the suggested threshold of 0.50 (Byrne, 2010). Finally, construct validity was established by satisfactory content validity, uni-dimensionality, reliability, and convergent and discriminant validity (O'Leary-Kelly & Vokurka, 1998). Composite reliability, and correlations and the AVE values are all presented in Table 2.

The measurement model test results showed that the model overall has good fit. The CMIN/DF ratio (= 1.42) shows superior model fit. The CFI (= 0.95) and TLI (= 0.95) values indicate good fit, as these two values are close to the 0.95 cutoff line. The NFI value (= 0.86) suggests acceptable model fit. Lastly, the RMSEA (= 0.03) shows superior model fit.

Following prior studies (e.g., Swafford, Ghosh, & Murthy, 2006), we evaluated non-response bias, using a *t*-test. *T*-tests were performed comparing the responses of the first 30 and last 30 respondents with five items measured with an interval scale as dependent variables, which were randomly selected from the questionnaire. No significant difference was found between the two groups at the 0.05 level in all the five *t*-tests, suggesting no evidence of non-response bias in the data.

We addressed common-method bias in the questionnaire administration in several ways. First, we followed a rigorous approach for informant selection by ensuring that respondents are in the right positions in their companies, which indicate that they are fully informed of the subject matters investigated in the study. Second, while common-method bias is of greater importance in research related to issues of social desirability, this study focuses on topics that do not incur social desirability, further minimizing the potential for common-method bias. In addition, following recommendations provided in Podsakoff, MacKenzie, and Lee (2003), common-method bias was empirically assessed. More specifically, Harman's single-factor test was conducted. The single-factor model exhibited significantly worse fit than the measurement model, suggesting that common-method bias is not a serious concern in this study (Podsakoff & Organ, 1986).

Next, following the procedures of Schwarz et al. (2017) and Williams, Hartman, and Cavazotte (2010), we created three models: baseline, common latent factor, and marker variable, and then compared the models using chi-square difference test. As shown in Table 3, all the three models have good goodness of fit. The comparison test between baseline and common latent factor models shows that there is no significant difference between the two models, suggesting that common method is not a serious issue. However, the comparison between baseline and marker variable models indicates presence of common method bias in the dataset. But these tests did not show the extent to which common method bias exists in the dataset. Furthermore, the assumption underlying these model tests is that the difference between the two models in comparison is significantly different from zero. Therefore, following the approach recommended by Gaskin (2012), we noted that, from the marker variable model, the common factor unstandardized loading is -.042, with a square of .002, which is far less than 50%. Based on these two statistics, we can conclude that common method bias, is not a serious concern in the dataset.

Table 3: Chi-Square, Goodness-of-Fit Values, and Model Comparison Tests

Model	**$\chi 2$**	**df**	**CFI**	**RMSEA**
Baseline	376.33	246	.95	.04
Common Latent Factor	376.00	239	.95	.04
Marker Variable	474.60	314	.94	.04
Chi-Square Model Comparison Tests ΔModels	$\Delta\chi 2$	Δdf	P	
Baseline vs. Common Latent Factor	0.33	7	1.00	
Baseline vs. Marker Variable	98.28	68	.01	

Structural Model

In testing the structural model, all the constructs and their pre-specified relationships were entered in the model. All the following constructs with their measurement items still remain in the model: Big Data capability, relationship building, knowledge sharing, and product improvement. The final results show that the model has a good fit. The CMIN/DF ratio (= 1.46) shows superior model fit. Both the CFI (= 0.97) and TLI (0.97) values indicate strong goodness-of-fit. The NFI (= 0.88) value is a little bit lower, suggesting acceptable model fit. The RMSEA (= 0.04) suggests superior model fit. Thus, overall, these indicators show a good model fit.

RESULTS

The hypotheses covered in the structural model test are as follows: H1, H2, H3, H4, and H5a and H5b. Hypothesis 1 posits that supply chain relationship building is positively related to Big Data capability. The test results indicate that supply chain relationship building has a strong positive relationship with Big Data capability. Thus, H1 ($\beta = 0.19$, $p < 0.0001$) is supported. Further, hypothesis 2 posits that supply chain relationship building is positively related to knowledge sharing in supply chains. The test results ($\beta = 0.38$, $p < 0.0001$) show that H2 is also supported. Similarly, hypothesis 3 suggests that knowledge sharing is positively related to Big Data capability. The results ($\beta = 0.74$, $p < 0.0001$) show that H3 is supported. Hypothesis 4 indicates that Big Data capability positively contributes to product improvement. The results show support to this hypothesis ($\beta = 0.99$, $p < 0.0001$).

Finally, the two mediation hypotheses (H5a and H5b) were tested. Traditionally, the Baron and Kenny (1986) method has been used extensively to test mediation hypotheses in social science research. However, serious concerns were raised regarding its validity (Zhao, Lynch, & Chen, 2010). To overcome some of the key weaknesses of the Baron and Kenny (1986) method, a more advanced method, the bootstrapping method, which provides more accurate and powerful mediation testing, has been proposed (Shrout & Bolger, 2002). In addition to greater statistical power, the bootstrapping method provides two advantages. One is that indirect effect can be measured directly (Hayes, 2009). The other advantage is that it does not assume normality in the mediation effect. This adds to the validity of the mediation test results as indirect effects may not necessarily exhibit a normal distribution (MacKinnon et al., 2002).

In running the bootstrapping mediation tests, we followed the convention of setting confidence interval at the 95% level and the number of resamples at 5,000. The mediation test results are shown in Table 4. Hypotheses 5a states that Big Data capability mediates the effect of supply chain relationship building on product improvement. The results show that the mediation effect is significant, as 0 does not fall between the lower bound of 0.06 and upper bound of 0.22. Thus, hypothesis 5a is supported. Additionally, the results show that this is a full mediation effect, as supply chain relationship building does not have a significant direct effect on product improvement. Hypothesis 5b indicates that Big Data capability mediates the effect of knowledge sharing in supply chains on product improvement. The bootstrapping mediation test results indicate that this mediation effect is significant, as 0 does not fall between the lower bound of 0.22 and the upper bound of 0.40. Therefore, hypothesis 5b is supported. Further, the results show that this is a partial mediation effect. But, the direct effect of knowledge sharing in supply chains on product improvement is significant only at the 0.05 level with $\beta = 0.12$. The effect was smaller than that of the mediation effect. Based on these test results, the conceptual framework is adjusted (as shown in Figure 2).

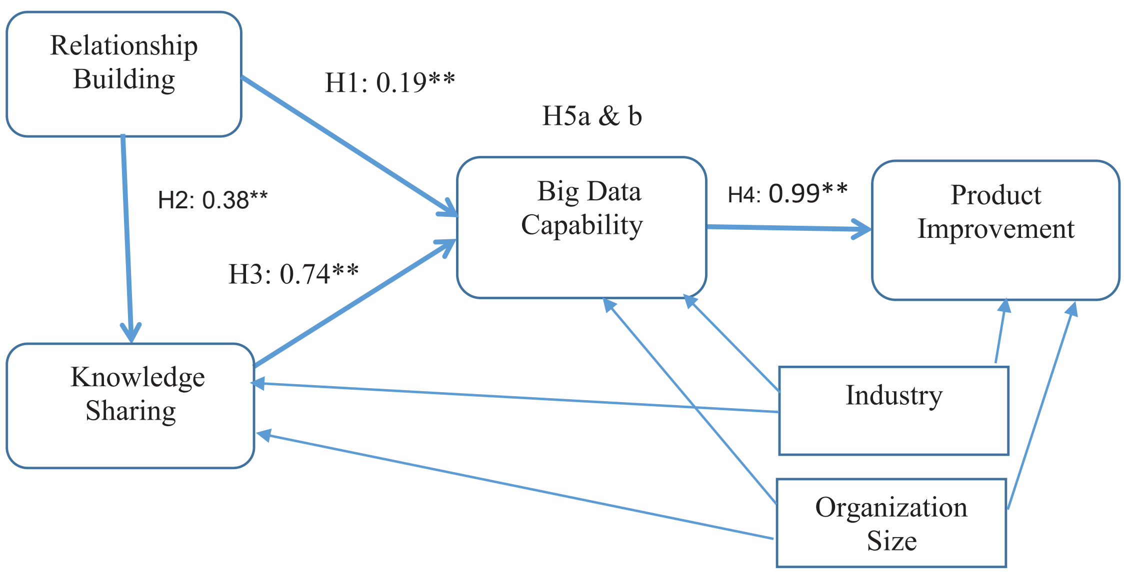
Table 4: Results of Bootstrapped CI Tests for Full and Partial Mediation Hypotheses:

			Mediation Test (*ab*)			Full/Partial Mediation Test (*c*)			
X variable	Mediator	Y variable	Lower bound	Upper bound	Zero included?	Lower bound	Upper bound	Zero included?	Type of mediation
RB	BDC	PI	0.06	0.22	No	0.21	0.49	No	Full
KS	BDC	PI	0.22	0.40	No	0.11	0.17	No	Partial

DISCUSSION AND CONCLUSIONS

In the emerging area of big data research, on the basis of numerous cases reporting how big data helps to enhance specific business functions including supply chains (for a review, see Ylijoki & Porras, 2016), academic research has pointed to the importance of conceptualizing big data initiatives and practices as dynamic capabilities and empirically verified the positive relationship between such capabilities and firms' market and financial performance (see Braganza et al., 2017; Chen, Preston, & Swink, 2015; Gupta & George, 2016; Wamba et al., 2017). To extend current big data research, this study draws on multiple theoretical perspectives to propose a framework where Big Data capability contributes to firm product improvement performance but its own development benefits from supply chain relationship

Figure 2: Adjusted Conceptual Model



building and knowledge sharing. Then it tests the validity and feasibility of this proposed framework with survey data collected from U.S. companies. We now present a discussion of the findings of this study with respect to theory and management practice.

First, the hypothesis testing results of this study show that Big Data capability directly enhances product improvement. Consistent with those from prior studies, this finding supports the view that only when Big Data initiatives and practices develop into Big Data capabilities can firms remain competitive in its performance (Braganza et al., 2017; Chen, Preston, & Swink, 2015; Gupta & George, 2016; Wamba et al., 2017), in this case, product improvement. Additionally, this finding contributes to the literature by offering us a more complete understanding of why or how big data enhances firm performance. It is that big data capabilities enhance product improvement, which then translates into market performance, and finally financial performance. The contribution lies in that it brings to light this step that connects big data capabilities and market and financial performance, which is missing in prior research.

Second, this study shows that supply chain relationship building and knowledge sharing are positively related to the development of Big Data capability. These findings suggest that supply chains supply nutrition to Big Data capability in its development. Additionally, they show that even though Big Data capability in nature is firm specific and path-dependent based on the dynamic capabilities perspective (Teece et al., 1997; Teece, 2007), its growth benefits from external resources, in this case, those available in supply chains. Further, these findings indicate that access to and even acquisition of such supply chain resources including knowledge is facilitated by firms' ability of building relationships with their supply chain partners. With these findings, this study contributes to the literature by identifying supply chain management activities, more specifically, supply chain relationship building and knowledge sharing, as antecedents to Big Data capability. This helps to extend this emergent enterprise of big data research to a new area and points to new directions for future research. Moreover, while previous research showed that big data contributes to supply chain performance, this study offers a new insight into supply chains that they constitute a site for nourishing the growth of big data dynamic capabilities with partners making efforts to build relationships and share knowledge with one another.

Third, the findings of this study indicate that Big Data capability strongly mediates the relationships between the two supply chain management activities, namely, relationship building and knowledge sharing, and product improvement. Our up-to-date knowledge about big data rests on the research finding that big data capabilities enhance firm performance. Thus, another contribution of this study is that it extends the literature by showing that big data capabilities not only contribute to firm performance but also facilitate and even aggrandize the effects of other contributors to product improvement, here in this case, supply chain relationship building and knowledge sharing.

In addition to the theoretical implications, this study offers several implications for managerial practices. First, the finding that Big Data capability enhances product improvement suggests that firms should integrate Big Data resources and competencies in alignment with their strategic directions to develop a Big Data capability. Second, the findings of this study inform firms that to increase Big Data capability's impact on product improvement performance, firms should learn to build positive relationships with their supply chain partners. These positive relationships also facilitate knowledge sharing among the partners. Knowledge from supply chain partners is a rewarding input to the development of Big Data capability.

Despite these contributions to both theory and practice, this study has some limitations that can be addressed in future research. First, in supply chain management, there are multiple factors, besides relationship building, that serve as conditioners to knowledge sharing's effect on product performance. We only examined the role of relationship building in this area. Future research can extend the examination to other factors such as supply chain integration, and organizational culture. Second, even though we took measures to ensure data validity and reliability in the data collection process, the endogeneity problem may still exist. Future research should consider different research designs, such as experiments, or use secondary data to establish the causal relationship between Big Data capability and product improvement.

REFERENCES

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.

Argote, L., McEvily, B., & Reagans, R. (2003). Managing knowledge in organizations: An integrative framework and review of emerging themes. *Management Science, 49(4)*, 571-582.

Atuahene-Gima, K., & Murray, J. Y. (2007). Exploratory and exploitative learning in new product development: A social capital perspective on new technology ventures in China. *Journal of International Marketing, 15(2)*, 1–29.

Autry, C. W., & Golicic, S. L. (2010). Evaluating buyer-supplier relationship-performance spirals: A longitudinal study. *Journal of Operations Management, 28*, 87-100.

Autry, Ch. W., & Griffis, S. (2008). Supply chain capital: the impact of structural and relational linkages on firm execution and innovation. *Journal of Business Logistics, 29 (1)*, 157-173.

Azadegan, A., & Dooley, K. J. (2010). Supplier innovativeness, organizational learning styles and manufacturing performance: An empirical assessment. *Journal of Operations Management, 28*, 488-505.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17(1)*, 99–120.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology,* 51(6): 1173-1182.

Barringer, B., R., & Harrison, J. S. (2000). Walking a tightrope: Creating value through interorganizational relationships. *Journal of Management, 26(3)*, 367-403.

Barton, D., & Court, D. (2012). Making advanced analytics work for you: A practical guide to capitalizing on big data. *Harvard Business Review*, October, 79-83.

Bercovitz, J. E. L., & Feldman, M. P. (2007). Fishing upstream: Firm innovation strategy and university research alliances. *Research Policy, 36(7)*, 930-948.

Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research, 70*, 328-337.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.

Chang, R. M., Kauffman, R. J., & Kwon, Y. O. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems, 63*, 67-80.

Chen, D. D., Preston, D. S., & Swink, M. (2015). 'How the use of Big Data analytics affects value creation in supply chain management'. *Journal of Management Information Systems*, *32(4)*, 4-39.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36(4),* 1165-1188.

Cheung, M-S., Myers, M. B., & Mentzer, J. T. (2010). Does relationship learning lead to relationship value? A cross-national supply chain investigation. *Journal of Operations Management, 28*, 472-487.

Christensen, C. M. (1997). *The innovator's dilemma*. Boston: Harvard Business School Press.

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16(1)*, 64-73.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: The Belknap Press of Harvard University Press.

Collins, D. J. (1994). How valuable are organizational capabilities? *Strategic Management Journal, 15,* 143-152.

Corsten, D., Gruen, T., & Peyinghaus, M. (2011). The effects of supplier-to-buyer identification on operational performance – An empirical investigation of inter-organizational identification in automotive relationships. *Journal of Operations Management, 29*, 549-560.

Cousins, P. D., & Menguc, B. (2006). The implications of socialization and integration in supply chain management. *Journal of Operations Management, 24(5)*, 604–620.

Cousins, P. D., Handfield, R. B., Lawson, B., & Petersen, K. J. (2006). Creating supply chain relational capital: The impact of formal and informal socialization processes. *Journal of Operations Management, 24(6)*, 851–863.

Dyer, J. H., & Nobeoka, K. (2000). Creating and managing a high-performance knowledge shaing network: The Toyota case. *Strategic Management Journal, 21*, 345-367.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal, 21*, 1105-1121.

Eisenhardt, K. M., & Schoonhoven, C. B. (1996). Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. *Organization Science, 7(2),* 136-150.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39-50.

Gaskin, J. (2012). Common method bias using common latent factor. Retrieved on September 8, 2018. https://www.youtube.com/watch?v=Y7Le5Vb7_jg.

Gefen, D., Straub, D.W., & Boudreau, M.C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the AIS*, *41*, 1–78.

Gibbons, R., & Henderson, R. (2012). Relational contracts and organizational capabilities. *Organization Science, 23(5)*, 1350-1364.

Girod, S. J. G., & Whittington, R. (2017). Reconfiguration, restructuring and firm performance: Dynamic capabilities and environmental dynamism. *Strategic Management Journal, 38*, 1121-1133.

Grant, R. M. (1991). The resource-based theory of competitive advantage: Implications for strategy formulation. *California Management Review, 33*, 114-135.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, Winter Special Issue, *17*, 109–122.

Griffith, D. A., Harvey, M. G., & Lusch, R. F. (2006). Social exchange in supply chain relationships: The resulting benefits of procedural and distributive justice. *Journal of Operations Management 24(2)*, 85–98.

Gulati, R., & Nickerson, J. A. (2008). Interorganizational trust, governance choice, and exchange performance. *Organization Science, 19(5)*, 688-708.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management, 53*, 1049-1064.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 422-445.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43(1)*, 115-135.

Ho, H., & Ganesan, S. (2013). Does knowledge base compatibility help or hurt knowledge sharing between suppliers in coopetition? The role of customer participation. *Journal of Marketing, 77,* 91-107.

Holahan, P. J., Sullivan, Z. Z., & Markham, S. K. (2014). Product development as core competence: How formal product development practices differ for radical, more innovative, and incremental product innovations. *Journal of Product Innovation Management, 31(2)*, 329-345.

Inkpen, A. C., & Tsang, E. W. K. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review, 30(1)*, 146–165.

Ireland, R. D., Hitt, M. A., & Vaidyanath, D. (2002). Alliance management as a source of competitive advantage. *Journal of Management, 28(3)*, 413-446.

Johnston, D. A., McCutcheon, D. M., Stuart, F. I., & Kerwood, H. (2004). Effects of supplier trust on performance of cooperative supply relationships. *Journal of Operations Management, 22*, 23-38.

Karna, A., Richter, A., & Riesenkampff, E. (2015). Revisiting the role of the environment in the capabilities-financial performance relationship: A meta-analysis. *Strategic Management Journal, 37*, 1154-1173.

Kemper, J., Schilke, O., & Brettel, M. (2013). Social capital as a microlevel origin of organizational capabilities. *Journal of Product Innovation Management, 30(3)*, 589-603.

Kirshner, S. N., Levin, Y., & Nediak, M. (2017). Product upgrades with stochastic technology advancement, product failure, and brand commitment. *Production and Operations Management, 26(4)*, 742-756.

Kor, Y. Y., & Mesko, A. (2013). Dynamic managerial capabilities: Configuration and orchestration of top executives' capabilities and the firm's dominant logic. *Strategic Management Journal, 34*, 233-244.

Krause, D. R., Handfield, R. B., & Tyler, B. B. (2007). The relationships between supplier development, commitment, social capital accumulation and performance improvement. *Journal of Operations Management, 25(2)*, 528–545.

Lawson, B., Tyler, B. B., & Cousins, P. D. (2008). Antecedents and consequences of social capital on buyer performance improvement. *Journal of Operations Management, 26(3)*, 446–460.

Lemon, J. R. (1991) A structured approach to concurrent product/process development. *Manufacturing Systems, 9(6)*, 32-39.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal, 13 (Special issue)*, 111-125.

Lopez-Cabrales, A., Valle, R., & Herrero, I. (2006). The contribution of core employees to organizational capabilities and efficiency. *Human Resource Management, 45(1)*, 81-109.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7(1)*, 83-104.

Matthews, J. (2003). Strategizing by firms in the presence of markets for resources. *Industrial and Corporate Change, 12*, 1157-1193.

McAfee, A, & Brynjolfsson, E. (2012). Big Data: The management revolution. *Harvard Business Review,* October issue, 60-68.

McGinnis, M. A., & Mele Vallopra, R. (1998). *Purchasing and supplier involvement: New product development and production/operations process development and improvement.* Tempe, AZ: Center for Advanced Purchasing Studies.

Min, K., Stephen, K., & Chen, H. (2008). Developing social identity and social capital for supply chain management. *Journal of Business Logistics, 29(1)*, 283–304.

Mitchell, W., & Singh, K. (1996). Survival of businesses using collaborative relationships to commercialize complex goods. *Strategic Management Journal, 17*, 169-195.

Monczka, R. M., Petersen, K. J., Handfield, R. B., & Ragatz, G. L. (1998). Success factors in strategic supplier alliances: The buying company perspective. *Decision Sciences, 29(3)*, 553–577.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review* 23(2): 242–266.

Nunnally, J. C. (1978). *Psychometric theory*. NY: McGraw-Hill.

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management, 16(4),* 387–405.

Paulraj, A., Lado, A. A., & Chen, I. J. (2008). Inter-organizational communication as a relational competency: Antecedents and performance outcomes in collaborative buyer-supplier relationship. *Journal of Operations Management, 26*, 45-64.

Peppard, J., & Ward, J. (2004). 'Beyond strategic information systems: Toward an IS capability'. *Journal of Strategic Information Systems, 13*, 167-194.

Peteraf, M., Di Stefano, G., &Verona, G. (2013). 'The elephant in the room of dynamic capabilities: Bringing two diverging conversations together'. *Strategic Management Journal, 34*, 1389-1410.

Phillips, L. W., & Bagozzi, R. P. (1986). On measuring the organizational properties of distribution channels: methodological issues in the use of key informants. In: Sheth, J. N. (Ed.), *Research in Marketing: a Research Annual: Distribution Channels and Institutions*. JAI Press, Greenwich, pp. 313–369.

Pigni, F., Piccoli, G., & Watson, R. (2016). Digital data streams: Creating value from the real-time flow of big data. *California Management Review, 58(3)*, 5-25.

Pisano, G. P., & Wheelwright, S. C. (1995). The new logic of high-tech R&D. *Harvard Business Review, 73(5)*, 93-105.

Podsakoff, P. M., &Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12(4),* 531–44.

Podsakoff, P. M., MacKenzie, S. B., & Lee, J. (2003). Common method bias in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903.

Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, *22*(4), 375–385.

Salomon, R., & Martin, X. (2008). Learning, knowledge transfer, and technology implementation performance: A study of time-to-build in the global semiconductor industry. *Management Science, 54(7)*, 1266-1280.

Schilke, O. (2014). On the contingent value of dynamic capabilities for competitive advantage: The nonlinear moderating effect of environmental dynamism. *Strategic Management Journal, 35*, 179-203.

Schoenherr, T., Ellram, L. M., & Tate, W. L. (2015). A note on the use of survey research firms to enable empirical data collection. *Journal of Business Logistics, 36(3)*, 288-300.

Schwarz, A., Rizzuto, T., Carraher-Wolverton, C., Roldan, J. L., & Barrera-Barrera, R. (2017). Examining the impact and detection of the “Urban Legend” of common method bias. *The DATA BASE for Advances in Information Systems, 48(1)*, 93-119.

Scott, J. (1987). *Organizations*. Englewoods Cliffs, NJ: Simon & Schuster.

Shrout, P., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7(4)*, 422-445.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal, 28*, 1319-1350.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18(7),* 509-533.

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal, 21*, 1147-1161.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal, 41(4)*, 464–476.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly, 31*, 439-465.

Tushman, M. L., & O’Reilly III, C. A. (1996). Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review, 38(4),* 8-30.

Villena, V. H., Revilla, E., & Choi, T. Y. (2011). The dark side of buyer-supplier relationships: A social capital perspective. *Journal of Operations Management, 29*, 561-576.

Waller, M. A., & Fawcett, S. E. (2013). Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. *Journal of Business Logistics, 34(4)*, 249-252.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356-365.

Wang, J., & Zhang, J. (2016). Big data analytics for forecasting cycle time in semiconductor wafer fabrication system. *International Journal of Production Research, 54(23)*, 7231-7244.

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique. *Organizational Research Methods, 13(3)*, 477-514.

Winter, S. G. (2003). Understanding dynamic capabilities. *Strategic Management Journal, 24*, 991-995.

Ylijoki, O., & Porras, J. (2016). Conceptualizing big data: Analysis of case studies. *Intelligent Systems in Accounting, Finance and Management, 23*, 295-310.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37(2)*, 197-206.

Appendix

Codes	*Questionnaire Items*
	Big Data Capability (BDC)
BDC1	We are able to identify sources of big data that meet our needs.
BDC2	We are able to collect big data that meet our needs.
BDC3	We are able to store large volumes of data.
BDC4	We are able to process big data with a fast speed.
BDC5	We adopt state of the art technologies to process big data.
BDC6	We constantly update our computing equipment to process big data.
BDC7	We constantly update our IT architecture to process big data.
BDC8	We constantly update our IT infrastructure to process big data.
BDC9	We are good at data analytics which is mainly data mining and statistical analysis.
BDC10	We are good at text analytics that deals with unstructured textual format data.
BDC11	We are good at web analytics that deals with web sites.
BDC12	We are good at mobile analytics that deals with mobile computing.
BDC13	We rely on Big Data to identify new business opportunities.
BDC14	We rely on Big Data to develop new products.
BDC15	We rely on Big Data to enhance our innovativeness.
BDC16	We rely on Big Data to formulate our business strategy.
	Relationship Building (RB)
RB1	Trust level in the relationship with your buyers.
RB2	Commitment level in the relationship with your buyers.
RB3	Collaboration level in the relationship with your buyers.
RB4	Long-term orientation level in the relationship with your buyers.
RB5	Trust level in the relationship with your suppliers.
RB6	Commitment level in the relationship with your suppliers.
RB7	Collaboration level in the relationship with your suppliers.
RB8	Long-term orientation level in the relationship with your suppliers.
	Knowledge Sharing (KS)
KS1	My firm and our supply chain partners have shared a significant amount of knowledge with each other.
KS2	My firm has exchanged many ideas with our supply chain partners about how

	to improve each other's capabilities (in manufacturing, research and development, logistics, etc.).
	In general, the skills and knowledge that have been shared between my firm and our supply chain partners are
KS3	a. substantial amount.
KS4	b. advanced.
KS5	c. of significant value to both parties.
KS6	d. of significant use.
	Product Improvement (PI)
PI1	Big Data has enabled us to add new features that meet customers' needs to our current products.
PI2	Big Data has enabled us to add new functionalities that reflect customers' needs to our current products.
PI3	Big Data has made our current products more competitive in the market.

DECISION SCIENCES INSTITUTE

Systematic Review: Paradoxical Elements in Supply Chain Management

Jiayuan Zhang
The University of Rhode Island
Email: jiayuanzhang@uri.edu

Mehmet G. Yalcin
The University of Rhode Island
Email: mgyalcin@uri.edu

Douglas N.Hales
The University of Rhode Island
Email: dhales@uri.edu

ABSTRACT

Grounded with paradox theory, we conduct a systematic literature review within the supply chain management context to extract interdependent elements which possess persistent contradictions. We embark on this study by casting a wide net and searching through four different databases over the last 20 years of peer-reviewed literature between 1998 and 2019. We found 63 articles as the basis of our literature review. Our findings shed light on the extent of usage for the paradox theory as well as elements perceived as paradoxical by supply chain management scholars in the extant literature.

Keywords: Paradox theory, Systematic literature review, Paradoxical elements, Supply chain management

INTRODUCTION

In the past 20 years, paradox perspective has evolved from ground of zero to a theory (Lewis, 2000; Smith and Lewis, 2011; Lewis and Smith, 2014; Smith et al., 2017). According to paradox theory (PT), paradox is defined as "persistent contradictions between interdependent elements" (Schad et al., 2016). Paradox is suggested as an effective "theoretical lens to understand and to lead contemporary organizations" (Smith and Lewis 2011, pp.398) where fierce competition in a globalized marketplace that is fueled by innovation and sustainability pressures creates a setting that is full of uncertainties. Paradox theorists offer future research agendas (Schad et al., 2016; Smith et al., 2017, Schad et al., 2019) that point to the potential of paradox lens across diverse streams of management science. In this study, we follow their lead and embark on an introductory search to examine how the PT relates to the extant supply chain management (SCM) literature. We conduct our investigation by extending an exploration for the use of PT and those elements that are perceived as paradoxical by scholars within SCM literature. Our study conducts a systematic literature review (SLR) for paradoxical elements (PEs) by implementing the methodology that Durach et al. (2017) propose in SCM. Prior studies have not delved into the systematical review of PEs, especially in the domain of supply chain. This study is a forerunner of reviewing these PEs in-depth and group them based on the paradox theory (Smith and Lewis, 2011; Smith et al., 2017).

Scanning the SCM literature, Adler et al. (1999) investigate the paradoxical dynamic between flexibility and efficiency. A decade later, Adler et al. (2009) further developed the discussion of productivity dilemma that Abernathy (1978) discovered in the automobile industry. Innate trade-off approach of SCM literature is also rich in plenty of other instances of paradoxical elements, though usually presented in dualities such as lean and agile, efficiency and flexibility, etc. (e.g. Naylor et al. 1999). Moreover, for instance, interactions created between an academic and a practitioner (Bartunek and Rynes, 2014) can be perceived as paradoxical at times. By investigating PEs in the SCM domain, our study makes several contributions.

We structured our paper as follows. We briefly introduce the PT in the following section. Then, we explain the SLR approach in the methodology section and continue with reporting the study findings in the results section. In the end, we reported the findings of the study and future discussions along with potential future research avenues were presented.

PARADOX THEORY (PT)

Lewis (2000) explores paradoxes and provides insights into the increasing organizational complexity and ambiguity. Lewis proposes a paradox framework that clarifies the nature of paradoxical tensions, reinforcing cycles, and their management by initially providing three paradox types – paradox of learning, paradox of organizing, and paradox of belonging. However, the study does not include the performing paradox which is added much later. Lewis' work serves as a lens for examining surprising findings and seemingly absurd aspects of organizational life. In Smith and Lewis (2011), the authors review the paradox literature ten years after the paradox framework is introduced. Our results reveal that the performing paradox has the largest number of PEs among the four paradoxes. Highlighting the debates in paradox literature, the article extends the paradox categories mentioned in Lewis (2000) and includes a fourth paradox category – performing paradox. In addition to the paradox categories, the two main challenges to paradox theory are also discussed. The first challenge of the paradox research stems from the lack of clarity of what paradox is, and the second challenge stems from the ontological debate that differentiates paradoxical tensions either as an inherent feature of a system or as social constructions. In their study, Smith and Lewis propose a model that seeks conceptual clarity, describes both the inherent and socially constructed features of organizational tensions and integrates management strategies of acceptance and resolution.

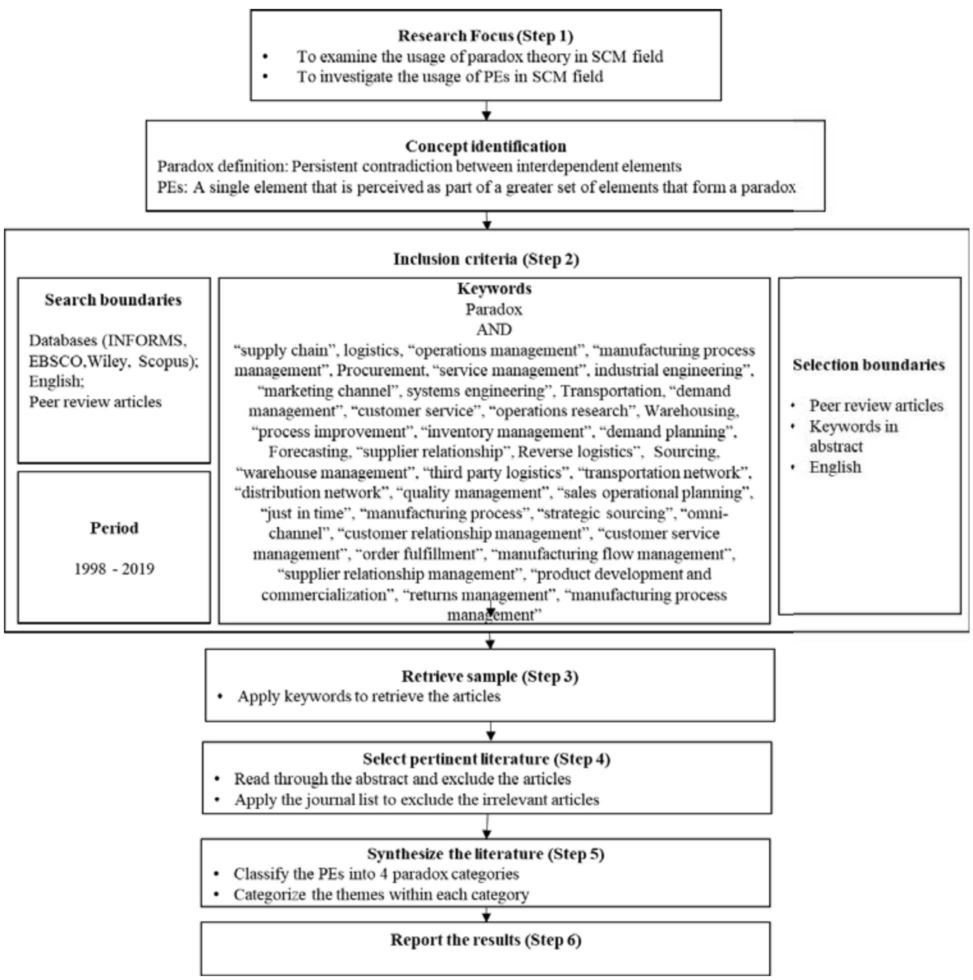
Smith et al. (2017) arguably pin phenomenon of paradox to ancient philosophy while citing that paradox research increased by ten percent per year between 1990 and 2014 after building on early research conducted in psychoanalysis, communications, and macro sociology disciplines. The authors also add a collection of studies that advance the PT that is rich in the explorative nature of paradox and posit that interdependent contradictions are inherent in human nature, its environment as well as the constructs that we humans build (e.g. life-death, knowledge-ignorance, self-other, expansion-constriction, independence-dependence, time-space, particles-waves, stability and change, empowerment and alienation, flexibility and control, diversity and inclusion, exploration and exploitation, social and commercial, competition and collaboration, learning and performing, profits and purposes, novelty and usefulness). In this literature review, we follow the appeal of the paradox scholars and intend to classify and summarize the PEs in the SCM domain.

Based on our SLR, we extract the PEs found in the SCM related articles and classify them into 10 paradox categories (Smiths and Lewis, 2011) and then group them under themes within each category. In Lewis (2000), the themes are specified within each paradox category except the performing paradox. By classifying the PEs from the SCM domain in the paradox categories and grouping them within different themes, we provide a foundation to import the PT into the SCM domain. In doing so, we hope to help spur more research on introducing the theory to the SCM field.

SYSTEMATIC LITERATURE REVIEW (SLR) METHODOLOGY

Durach et al. (2017) argue that while SLR has helped substantially to developing knowledge in fields such as medicine, it has made limited contributions to developing knowledge in the SCM domain. They suggest that an SLR needs to be adjusted to unique characteristics of SCM and thus further recommend six steps to achieve this. We adhered to these six steps of the SLR in this study (see Figure 1): step 1 define the research question, step 2 determine the required characteristics of the study, step 3 retrieve a sample of potentially relevant literature, step 4 select the pertinent literature, step 5 synthesize the literature, and step 6 report the results. Each step needs to be addressed carefully. Otherwise, several biases associated with each step will occur. The discussion of addressing the biases associated with each step will also be described in the following sections.

Figure 1: Systematic Literature Review (SLR) Process



Defining the Research Questions (Step 1)

This SLR intends to explore the PT in the SCM domain. Particularly, we want to explore the level that PT is employed in the SCM literature. Therefore, the first research question in the SLR is “What is the scale of usage of PT in the SCM literature.” Besides, the SLR also intends to extract the PEs used in the SCM domain where the PEs are defined as those elements that are *persistently contradictory and interdependent with one another.* Therefore, the second research question in the SLR is “What are the PEs used in the extant SCM literature.”

Determining the Required Characteristics of the Study (Step 2)

In this study, we included the previous studies that range from January 1998 to December 2019. We believe that the duration of two decades is long enough to keep track of the usage of the paradox related studies in the SCM field. This time frame is also in line with the emergence of the idea of organizational paradox. We include articles with both qualitative and/or quantitative methodologies because both methods benefit the SLR in SCM (Durach et al., 2017). The four databases that we employed are ABI/INFORM Complete (ProQuest), Academic Search Complete (EBSCO host), Scopus (Elsevier), and Wiley Online Library. ABI/INFORM Complete (ProQuest) is the largest database including full-text of scholarly and trade journal articles in business, management, and trade. Wiley Online Library covers a large number of supply chain management journals. Academic Search Complete (EBSCO) and Scopus are the interdisciplinary databases that also cover business and management journals. Initially, we searched by applying keywords “paradox” and “supply chain” to search the articles in the entire article texts. However, we found that including the entire text did not yield any added value to the search. Instead, it provided a large number of search results and increased time to read through all the articles and identify any relevant PEs. We assumed that if the PE that is mentioned in the article is a significant one, then the authors would mention it in the abstract. Therefore, we only included previous studies whose abstract includes synonyms of “supply chain” and “paradox” together. We limited the search process within abstracts of the selected peer-reviewed journal articles in the above databases and set the language of the articles in English only. Upon collecting all the articles, we created a finalized SCM journal list to exclude articles that are not tightly related to the SCM domain. We first selected three SLR papers (Fabbe-Costes and Jahre, 2008; Colicchia and Strozzi, 2012; Seuring and Gold, 2012) in SCM. Based on these three-literature review papers, we used the impact factor and citation centrality criteria to select the relevant journals in the SCM journal list. We also assumed that if a journal publishes a literature review in supply chain management, the journal can be considered related to the SCM domain. Therefore, we also search the databases for journals that have published SLR in the SCM domain.

We only focused on those articles that mention the word paradox and we excluded keywords such as ambidexterity, trade-off, and dilemma. The reason that we exclude these keywords is that even though they are similar to paradox, they do not fully satisfy the definition of paradox which suggests interdependence and persistent contradiction among the PEs.

Retrieving a Sample of Potentially Relevant Literature (Step 3)

The use of multiple databases can help us address the retrieval bias because it reduces the chance to miss any relevant articles. In addition to the multiple databases, we also finalized a list of 60 SCM journals to retain the articles (see Appendix). The use of the list of 60 SCM journals in lieu of any subjective preferred list of journals is instrumental in addressing the publication bias because we not only included the leading SCM journals but also all of the relevant SCM journals.

Durach et al. (2017) suggest that the search applies a combination of search strings, which are based on research purpose, research questions, and inclusion/exclusion criteria. We subjected the selection of our search strings to a series of processes. In the first step, we created a SCM domain keyword table which includes the phrases “Supply Chain”, “Logistics”, “Operations Management”, “Manufacturing Process Management”, “Service Management”, “Production Management”, “Industrial Engineering”, “Systems Engineering”, “Procurement”, “Marketing channel”.

In the second step, we formed an expert panel to help us refine the keywords list. We first thought that SCM scholars would have more knowledge on the supply chain synonyms, so we first consulted with 5 SCM scholars to refine our SCM keyword table. We also consulted with two SCM practitioners who have many years of industry experience in SCM practice. They helped us to include more relevant keywords that were not captured previously. In the end, we also consulted with subject librarians on the selection of keywords. After repeated discussions with the expert panel, we finally came up with a list of 40 synonyms for the key phrase “supply chain”. Table 1 lists the synonyms for the keyword “supply chain”. The key phrases are ranked according to the frequency mentioned. We opted to include all keywords/phrases in this study to cast a wide net for our search.

Table 1: Search keywords developed by the expert panel

	Supply chain management
10	Logistics
9	Operation Management
8	Manufacturing Process Management
8	Procurement
7	Service Management
6	Industrial Engineering
6	Production Management
5	Marketing channel
4	Systems engineering
2	Transportation
2	Demand Management
2	Customer Service
1	Operation Research
1	Warehousing
1	Process Improvement
1	Inventory Management
1	Demand Planning
1	Forecasting
1	Supplier Reltaionship
1	Reverse Logistics
1	Sourcing
1	Warehouse Management
1	Third party logistics
1	Transportation Network
1	Distribution Network
1	Quality Management
1	Sales Operational Planning
1	Just In Time/Lean Manufacturing
1	Manufacturing Processes
1	Strategic Soucring
1	Onni-Channel
1	Customer relationship management
1	Customer service management
1	Order fullfillment
1	Manufacturing flow management
1	Supplier relationship management
1	Product development and commercialization
1	Returns management
1	Manufactruring Process Management

With the combination of each of the 40 SCM field related keywords/phrases and the keyword “Paradox”, we performed the search within the abstracts of the target literature. The initial number of articles resulted in the databases were 540 in the ABI/INFORM Complete (ProQuest), 258 for in the Academic Search Complete (EBSCO host), 320 in the Scopus (Elsevier) and 699 in the Wiley Online Library is 740. After the initial search, the combined total number of articles in the four databases was 1,817.

Selecting the Pertinent Literature (Step 4)

In step 4, the inclusion and exclusion criteria are applied to select the articles. The authors first developed the inclusion and exclusion criteria independently and then compared with each other. To develop the inclusion and exclusion criteria individually help to address the inclusion bias, which might have led to the emergence of the incorrect results. The inclusion criteria are 4 databases, “1998-2019”, “English”, “Peer review”, “search strings”. As we mentioned earlier, we do not include the articles that focus dilemma, dialectic, and trade-off.

In each database, the authors read through the abstract of all the articles. If an article mentioned both the SCM field keywords and “paradox” in the abstract, the authors kept the article. Any articles that did not include both search strings in abstract were excluded. After this scanning process, the number of studies kept was 168 articles from ABI/INFORM Complete (ProQuest), 46 articles from Academic Search Complete (EBSCO host), 90 articles from Scopus (Elsevier), and 102 articles from Wiley Online Library. At this stage, the total number of articles was 423. In the final step, we applied the list of 60 journals to exclude the articles that are not published in the SCM related journals. Total of 63 articles are included in the literature review. The articles that were eliminated in this round were kept for gaining any possible insights from journals that are discussing paradoxes within the SCM context.

Synthesizing the Literature (Step 5)

In the beginning, two of the researchers in this study extracted the PEs individually. Having multiple researchers to extract the PEs individually can address the within-study bias, which would provide an incorrect outcome of review (Durach et al, 2017). Both coders further discussed and then agreed on correctly identifying less than ten percent of the PEs which ultimately yielded to full intercoder agreement rate. A total of 74 sets of PEs were coded through these two independent coders.

During the synthesis stage, the PE sets were coded based on the four core paradox categories (Learning Paradox, Belonging Paradox, Organizing Paradox, and Performing Paradox) and six combined paradox categories (Learning::Belonging, Learning::Organizing, Belonging::Organizing, Learning::Performing, Performing::Belonging, Performing::Organizing) proposed by Sandberg (2017) that is grounded in Smith and Lewis (2011). The authors strictly followed the category criteria in Sandberg (2017), and if there was disagreement on putting the sets of PEs into a category, the authors discussed until an agreement was reached. There were no stalemates that necessitated the involvement of the third researcher in this study. The PE sets were then grouped into 10 categories which are Learning Paradox, Belonging Paradox, Organizing Paradox, Performing Paradox, Learning-Organizing Paradox, Learning-Performing Paradox, Belonging-Organizing Paradox, Performing-Organizing Paradox, Belonging-Performing Paradox. We do not find any group to paradox category Learning-Belonging.

Within each category, we color-coded the themes of the PEs. The authors coded the themes after a detailed discussion and only classified the PEs into a theme until a full agreement was made. In the learning paradox category, the authors classified PEs under three themes: old, new, and present. In the organizing paradox category, the authors classified PEs under five themes: collaboration, efficiency and control, exploration, competition, and others. In the belonging paradox category, the authors classified PEs under two themes: single entity and multiple entities. And lastly, in the performing paradox category, the authors classified PEs under seven themes: specific, service, cost, forecast, sustainability, overall, and investment.

Reporting the Results (Step 6)

Figure 2 presents a histogram for the number of PE sets in each year. The usage of the PEs in SCM has increased from 1998 to 2019. More PE sets are perceived as paradoxical in recent years than two decades ago. This increase in numbers has also been steady over the years.

Figure 2: PE sets over the years of 1998 – 2019

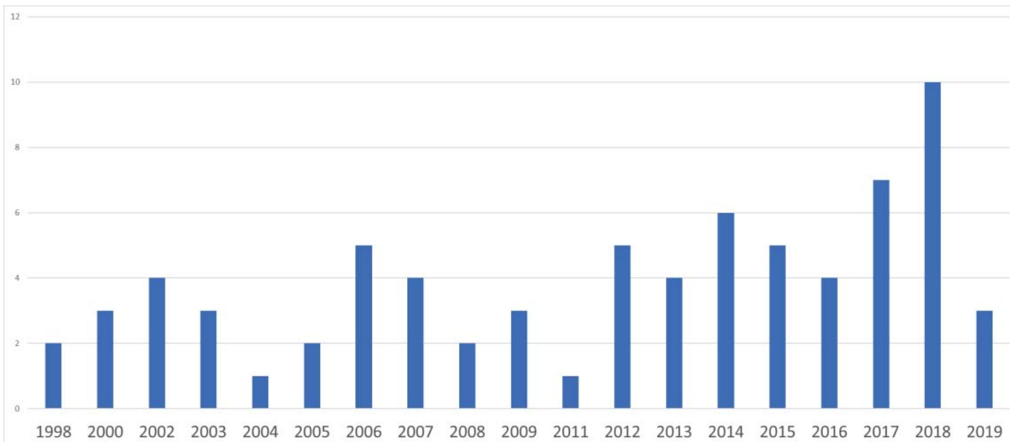


Table 2 shows the usage of PE sets in SCM journals. The Decision Science Journal has the greatest number of PE sets at 7. The second-ranking journals are the International Journal of Operations and Production Management, Journal of Supply Chain Management, and Production and Operations Management with 6 PE sets. Many journals seem to have published at least one or two articles that contain a PE set.

Table 2: The usage of PE sets in the SCM journals

Number of Journals	**Journal Name**	**Number of PE sets in each Journal**
9	Benchmarking Computers & Operation Research European Journal of Operational Research European Journal of Purchasing and Supply Management Journal of cleaner production Journal of the Operational Research Society Management Science Supply Chain Management Technological Forecasting & Social Change	1
7	International Journal of Logistics Management International Journal of Logistics Systems and Management International Journal of Management Reviews International journal of production economics	2

	International transactions in operational research Journal of Business and Industrial Marketing Sustainability	
5	Corporate Social Responsibility and Environmental Management International Journal of Physical Distribution & Logistics Management Journal of Business Logistics Transportation Research Part E	3
1	International Journal of Logistics Research and Applications	4
2	Journal of Operations Management Omega: International Journal of Management Science	5
3	International Journal of Operations & Production Management Journal of Supply Chain Management Production and Operations Management	6
1	Decision Science	7

Figure 3 shows the number of PE sets under each paradox category. The performing paradox category has the largest number of PE sets which is followed by the intersection of performing/organizing paradox category, and then by the organizing paradox category.

Figure 3: Number of PE sets under each paradox category

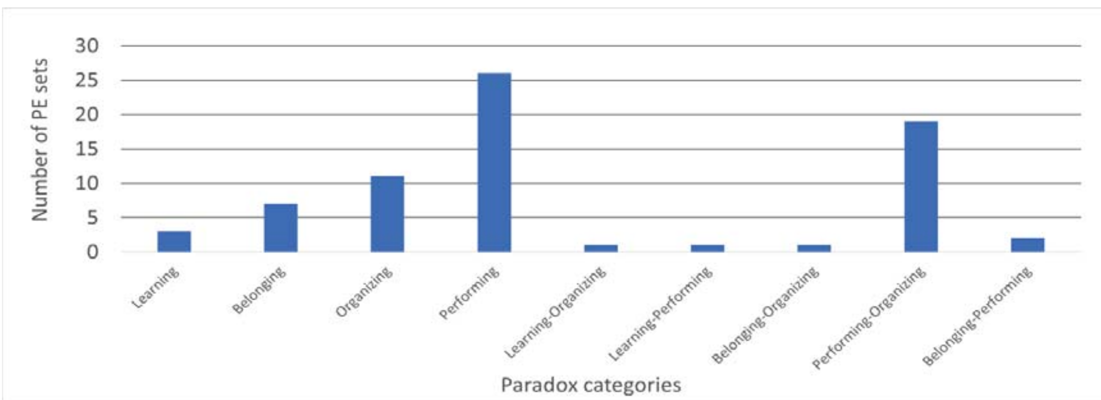


Table 3 summarizes the number of themes in each paradox category. The learning Paradox has three themes. The organizing paradox has 5 themes. The belonging paradox has 2 themes. The performing paradox has 7 themes.

Table 3: The number of themes under each paradox category

Learning Paradox (3 themes) Old; New; Present	Organizing Paradox (5 themes) Collaboration; efficiency and control; exploration; competition; and others.
Belonging Paradox (2 themes) Single entity; Multiple entities	Performing Paradox (7 themes) specific, service, cost, forecast, sustainability, overall, and investment.

Table 4 shows the paradox categories in each article. Overall, most of the articles only identify one paradox. Sandberg (2017) identifies four paradox categories. Looney et al (2016) and Kolk (2012) identify three paradox categories. Wacker and Lummus (2002) identify two paradox categories.

Table 4: Usage of PEs in the selected articles

No.	Author	Learning Paradox	Belonging Paradox	Organizing Paradox	Performaning paradox	Learning-Organizing Paradox	Learning-Performing Paradox	Belonging-Organizing Paradox	Performing-Organizing Paradox	Belonging-Performing Paradox
1.	Shmueli, Galit, and Inbal Yahav (2018)							X		
2.	Li et al. (2005)				X					
3.	Agarwal and Prasad (1997)				X					
4.	Storer et al. (2014)				X					
5.	Khazanchi et al. (2006)			X						
6.	Stefan and Bengtsson (2017)								X	
7.	Kull et al. (2013)			X						
8.	Tyworth and Saldanha (2014)			X						
9.	Tyworth (2018)			X						
10.	Kim et al. (2005)			X						
11.	Rindova (2011)	X								
12.	Sandberg (2017)	X	X	X	X					
13.	Barros and Hilmola (2007) *									
14.	Basso et al. (2018)								X	
15.	Cerruti et al. (2016)			X						
16.	Duray et al. 2000								X	
17.	Upton and Mcafee (2009)								X	
18.	Looney et al. (2006)				XX		X		X	
19.	Tazelaa and Snijders (2013)									X
20.	Sarkis and Semple (2007)								X	
21.	Voordijk et al. (2000)			X						
22.	Graetz and Smith (2008)			X						
23.	Mellat-Parast and Digman (2008)			X						
24.	Terziovski and Guerrero(2014)	X								
25.	Choi and Eboch (1998)								X	
26.	Wacker and Lummus (2002)				XX				X	
27.	Batiz-Lazo (2004)			X						
28.	Koufteros et al. (2003)				X					
29.	Sousa and Voss (2009)				X					
30.	Steinbach, Wallenburg and Selviaridis (2018)		X							

No.	Author	Learning Paradox	Belonging Paradox	Organizing Paradox	Performaning paradox	Learning-Organizing Paradox	Learning-Performing Paradox	Belonging-Organizing Paradox	Performing-Organizing Paradox	Belonging-Performing Paradox
31.	Pereira (2003)				X					
32.	Kastalli and Van Looy (2013)								X	
33.	Brockett et al. (2003)				X					
34.	Wilhelm and Sydow (2018)			X						
35.	Chiadamrong and Wajcharapornjinda (2012)								X	
36.	"Henry" Jin, Fawcett, and Fawcett (2013)								X	
37.	Vedel (2016)									X
38.	Kolk (2012)			X	X				X	
39.	Kronborg Jensen (2012)		X							
40.	Zhu et al. (2018)					X				
41.	Waller, Fawcett and Johnson (2015)				X					
42.	Busse,Kach and Bode (2016)		X							
43.	Rizzi, Frey, Testa and Appolloni (2014)		X							
44.	Matthews, Power, Touboulic and Marques (2015)		XX							
45.	Schmidt, Foerstl, and Schaltenbrand (2017)								X	
46.	Xiao, Wilhelm; van der Vaart and van Donk (2018)				X					
47.	Murali, Lim and Petruzzi (2015)				X					
48.	New (2015)			X						
49.	Klumpp (2016)				X					
50.	Brix-Asala et al. (2018)				X					
51.	Shalley and Gilson (2017)			X						
52.	Yang and Chen (2009)								X	
53.	Masuda and Whang (2002)								X	
54.	Adlakha and Kowalski (1998)				x					
55.	Adlakha and Kowalski (2000)				x					
56.	Adlakha, Kowalski, Vemuganti, and Lev (2007)				x					
57.	Rapoport, Gisches, and Mak (2014)								X	
58.	Mak et al. (2018)								X	
59.	Zhao, Fu and Wang (2014)								X	
60.	Pal, Harper, and Vellesalu (2018) *									
61	Niesten and Stefan (2019)								X	
62	D'Ambrosio, Gentili, and Cerulli (2019)								X	
63	Yin and Zhang (2019)								X	

* There is no clear paradoxical elements in the paper.

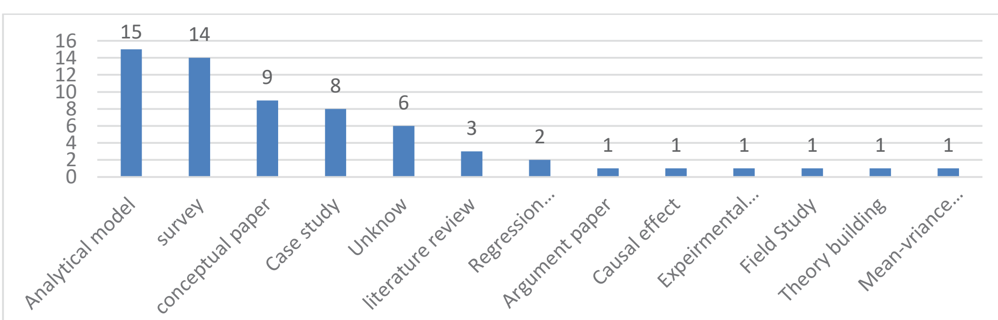
Table 5 shows the usage of theory in articles. Most of the articles in the review don't use theory. We also find that paradox theory is the most used theory in the literature review. It also shows the increased usage of the paradox perspective.

Table 5: The usage of theory in articles

Institutional Theory	1
Adaptation Theory	1
Agency theory	1
Behavioral theory	1
Componential Theory of Creativity	1
Decision Theory	1
Downs-Thomson paradox	1
forecast theory	1
Information Theory	1
institutional theory	1
Interactionist Model of Creativity	1
Inventory management theory	1
Jevons Paradox	1
Knowledge Based View	1
Mass Customization	1
Natural resource based view	1
Organizational information processing theory	1
Prospect theory	1
radical innovation	1
Relational view theory	1
Servitization Paradox	1
Simpson's paradox	1
Social Cognitive Theory	1
Sound Behavioral Principle	1
sustainable supply chain management theory	1
The theory of planned behavior	1
Utility Maximization Theory	1
Stakeholder theory	2
Resource Based View	2
Transaction Cost Economics	2
More-for-less	4
Braess Paradox	5
Paradox Perspective	6
Unknow	6
None	16

Figure 4 shows the usage of method in articles. The analytical model is the mostly used method in the literature. The survey method is the second mostly used method. Conceptual paper and case study are the third and fourth mostly used methods in the article.

Figure 4: The usage of method in articles



Our SLR also came across two studies that offered unique PE sets. PE set in Choi and Eboch (1998)'s Journal of Operations Management paper was formed via 3 PEs and this TQM study's results explicated paradoxical relations among TQM practices, plant performance, and customer satisfaction. On the other hand, a Journal of Supply Chain Management paper, Matthews et al. (2016) presented 4 PEs within a PE set by looking into paradoxical tensions among different levels of analysis: encompassing individuals, the organization, inter-organizational networks, and macroenvironmental levels.

DISCUSSION

Based on the results, we find that the trend of the occurrence of paradox sets is increasing. The increase of the use of PEs may be related to ever-increasing business complexity, globalized organization in the supply chain network. Especially as firms increasingly establish global and complex network structures, more tensions appear, and more PEs surface.

Addressing the first challenge raised by paradox theorists regarding what paradox is, we find that performing paradox is the dominant paradox category, where we have identified 7 themes. The seven themes in performing paradox are *specific, service, cost, forecast, sustainability, overall, and investment.* The *specific* theme is related to specific supply chain performance. The *service* theme is related to service elements in the supply chain, such as lead time variability and optimal safety inventory. The *cost* theme is related to the cost elements in the supply chain. These supply chain cost elements are summarized as the total cost of logistics and the amount of costs to ship. The *forecast* theme is related to forecast accuracy and error. The *sustainability* theme is related to sustainability issues in the supply chain, such as green supply chain management performance and sustainability. The *overall* theme is related to overall supply chain performance, such as plant performance. The *investment* theme is related to investment performance, such as IT investment. We thus provide new insights for the performing paradox. These 7 themes in the performing paradox reflect the characteristics of supply chain management, which focus on cost, performance, and efficiency. Some of the paradox sets under the performing category are closely related to supply chain outcomes, such as traffic flow performance, amount of goods shipped, and optimal safety inventory.

The five themes in organizing paradox are *collaboration, efficiency and control, exploration, competition, and others.* The *collaboration* and *competition* themes are similar to that in Smith and Lewis (2011), while the *efficiency and control*, *exploration*, and *others* themes are newly added. The *collaboration* theme consists of elements related to supply chain collaboration, such as supply chain coordination, supply chain integration, and global integration. These elements focus on the collaboration within the supply chain. The *efficiency and control* theme consist of elements such as TQM practices, standardization, and level of automation. The exploration theme investigates the exploratory activities in the supply chain. The elements consist of openness, supply chain transparency, and alternation to a road network to improve traffic flow, etc. The *competition* theme focuses on the competition element. The *others* theme consists of elements such as network distance to consumers and unsatisfied objectives.

We slightly modify the individual and collective tensions in belonging paradox (Smith and Lewis, 2011) and classify them into our newly created categories, the *single entity* and *multiple entities*. The themes reflect level of supply chain entities. The PE elements that are related to single level are grouped into single entity, while elements related to multiple levels are grouped into multiple entities.

The learning paradox has three themes, old, present, and new. The PE elements in the old theme are extant literature and historical experiences and know-how. The PE elements in the present theme are investor competence and level of analytics capabilities. The PE elements in the new theme are theory development, new ideas and knowledge, process innovation, and product innovation.

According to Smith (2000), the second challenge in the paradox literature stems from an ontological debate that differentiates paradoxical tensions either as an inherent feature of a system or as social constructions. Based on our findings, especially in the performing paradox, we find that the paradoxical elements can be classified into either as inherent features of the system or as social constructions. Take the widely used Braess's paradox as an example where the traffic performance decreases when an additional route is added to the network. When we talk about the traffic performance it can be classified as a social construction, and the additional route can be classified under the feature of the system. Our findings show that the paradoxical elements in supply chain management reflect this debate. Even though our findings do not help solve the debate, they provide a foundation to introduce the PT to the supply chain domain.

We also find that four methods, which are analytical paper, survey, conceptual model, and cast study, dominate the overall range of method used in the articles. It is not surprising that the analytical model is the most used method in the SCM domain (Wacker,1998). Therefore, even though our focus is on paradox research, the analytical approach still dominates the overall usage of methods. As for the qualitative approach, case study and conceptual model dominate. Because the PT is in its development stage, case study and conceptual model contribute to the development of the theory. It is interesting to notice that the 4 articles that apply paradox theory use either case study or conceptual model. Sandberg (2017), Xiao et al (2018), and Brix-Asala (2018), and Miriam and Jörg (2018) use case study. Matthews et al (2016) apply the conceptual model. Niesten and Stefan (2019) apply the paradox framework in the literature review.

In two of articles that were identified as containing more than two PEs within a PE set, Choi and Eboch (1998) and Matthews et al. (2016) are able to discuss paradoxical tensions 'among' the PEs. This is an interesting finding which positions SCM PEs outside the definition of paradox. The dualities, dilemmas, and dialectics which Smith and Lewis (2011, 387) suggest help us distinguish among organizational tensions, however, SCM PEs seem to protrude beyond, and perhaps due to the pressure emanating from the mindset that reside within performance paradox category.

From the results, we noticed repeatedly the perceived paradox between sustainability and business performance/economic goals. SCM scholars seemed to have combined the environmental and social issues under one PE while designating a standalone PE for the economic aspect. There are similarities between this perception and the one that is prevalent among layman which is tied to the expression 'time is money'. Whereas time is not money and it is well understood, for instance, the conventional management practices force many practitioners to fold time onto cost so a trade-off can be set up between cost and scope. Following a similar logic, when sustainability components are being considered, even in scholarly works, social and environmental dimensions seem to be lumped together so a duality is formed along with the use of economic dimension. This action might ultimately be reducing the complexity of decision-making process and therefore managers can perhaps constrain the cognitive workload under two PEs and then pursue trade-offs and perform optimization tasks.

A plausible argument can be made regarding the trade-offs. The analytical tools may not allow SCM scholars to obtain desired results when more assumptions are relaxed such that a strict trade-off is not warranted in a problem. Paradox theorists demonstrate over and over again that even if desired results may not be possible, meaningful results can be obtained and put into action by dealing with elements that form the trade-offs in PE sets. So, is it the PT that needs to navigate and find inroads into SCM domain, or should the SCM scholars entertain exploring the not-so-frequently explored territories beyond trade-offs?

CONCLUSIONS

This study summarizes and classifies the PEs in SCM via the help of SLR methodology. With the time span of 20 years and 4 databases, we cover a comprehensive list of PEs in SCM domain. According to the paradox categories (Smith and Lewis, 2011), we group the paradoxical elements into these categories and classify the themes within each category. By providing the paradox categories and groupings, we provide an inventory of paradox elements and themes to scholars and practitioners to the inventory of the paradox categories and themes. With the hope of providing paradox categories groups to the literature, we also provide insights to help address the debate in paradox research. The need for a SLR of PEs in SCM exists due to several reasons. First, there is a lack of a comprehensive list for PEs in the SCM domain. Based on PT, we identify and summarize the PEs in SCM in a list so an initial assessment of the state of PE usage can be offered for use by scholars and practitioners. The list also offers a path to further extend PT into SCM while informing back the PT for its refinement. In other words, we hope to contribute to both, centripetal forces that “buffer the existing boundaries” and the centrifugal forces that “foster boundary spanning” in PT’s development (Schad et al., 2019).

Second, recognizing PEs is the first step to handle these elements. Handling PEs simultaneously and successfully through the creation of “cyclical responses… enable sustainability – peak performance in the present that enables success in the future” (Smith and Lewis, 2011).

Paradox theory explores how organizations can attend to competing demands simultaneously (Smith and Lewis, 2011). Organizations who can handle the competing demands simultaneously can be more efficient and more successful. However, many researchers and practitioners claim that they find something contradicting but they did not notice that the findings can be analyzed under the paradoxical lens. With the use of PT, we can identify these elements and guide the practitioners to recognize and handle these elements simultaneously and effectively. This study applies PT to summarize and classify PEs in the SCM domain. The study intends to provide a list of paradoxical elements to SCM researchers and practitioners. The findings can help practitioners recognize and thus handle the PEs simultaneously.

The purpose of our research study was to search, collect, and then present the PEs in SCM literature. However, our SLR study is not without its limitations. First, we believe that the articles that were eliminated in Step 4 of our SLR that are ‘talking about paradoxes and supply chains’ will need to be included for gaining any possible supplementary insights in lieu of assuming them on the fringes. It will be interesting to compare and contrast the similarities and difference between PEs inside and outside SCM field. Second, even though mostly captured under the core elements of ‘exploitation and exploration’, including ambidextrous relationships that are examined in SCM research could further enrich our study. Third, the tradeoffs in SCM could be searched, collected, and then presented along with our findings in another study to draw a broader picture of the above arguments. Finally, we also introduce the paradox debate in the SCM domain, but our work is a first step toward bringing the debate to the SCM domain. We need more paradox studies in SCM.

The future study of paradox in the SCM filed is promising. Our results show that performing paradox and performing-organizing paradox has the most dominant paradox themes. Future studies can focus to investigate these two paradox categories where further insights can be obtained from these paradox categories. Our study also provides different themes of paradoxes within different categories which can be investigated for fitness in their contexts. This study also provides evidence for the ontological debate of paradox. We hope to spur interest and encouragement for the researchers to work on providing further evidence to inform the debate.

APPENDIX

List of the 60 SCM journals

Number	Journal Name	Impact Factor (Fabbe-Costes and Jahre,2008)	Citation Centrality (Colicchia and Strozzi,2012)	SCM literatrue Review (Seuring and Gold, 2012)	Updated SCM literature Review (2000-2019)	SCI Impact Factor 2018
1	Journal of Operations Management	x	x	x		6.48
2	Journal of Supply Chain Management				x	6.44
3	Management Science		x			6.08
4	Omega: International Journal of Management Science				x	3.29
5	Production and Operations Management		x		x	3.28
6	International Journal of Management Reviews			x	x	2.9
7	Journal of Business Logistics	x				2.49
8	International Journal of Production Economics		x	x		2.48
9	International Journal of Physical Distribution and Logistics Management	x		x	x	2.41
10	Surveys in Operations Research and Management Science				x	2.24
11	European Journal of Operational Research		x			2.21
12	Transport Reviews				x	2.14
13	International Journal of Operations and Production Management	x		x	x	2.1
14	Supply Chain Management: An International Journal	x	x	x		2.1
15	Transportation Research Part E: Logistics and Transportation Review					1.97
16	Computers and Operations Research				x	1.86
17	Corporate Social Responsibility and Environmental Management				x	1.67
18	Journal of Cleaner Production			x	x	1.62
19	International Journal of Production Research			x	x	1.59
20	Production Planning and Control				x	1.43
21	Technological Forecasting and Social Change				x	1.42
22	Journal of Intelligent Manufacturing				x	1.39
23	Decision Science					1.33
24	Expert Systems with Applications				x	1.19
25	International Journal of Logistics: Research and Applications	x				1.05
26	International Journal of Logistics Research and Applications				x	1.05
27	Annals of Operations Research				x	1.03
28	IMA Journal of Management Mathematics				x	1.02
29	International Journal of Advanced Manufacturing Technology				x	0.99
30	International Journal of Logistics Management	x		x	x	0.871

Number	Journal Name	Impact Factor (Fabbe-Costes and Jahre,2008)	Citation Centrality (Colicchia and Strozzi,2012)	SCM literatrue Review (Seuring and Gold, 2012)	Updated SCM literature Review (2000-2019)	SCI Impact Factor 2018
31	Journal of Humanitarian Logistics and Supply Chain Management				x	0.84
32	International Transactions in Operational Research				x	0.83
33	Journal of the Operational Research Society		x			0.82
34	Transportation Journal	x				0.81
35	Journal of Enterprise Information Management				x	0.69
36	International Journal of Productivity and Performance Management			x		0.64
37	Benchmarking				x	0.59
38	Sustainability				x	0.55
39	Logistics Research				x	0.51
40	IIMB Management Review				x	0.41
41	Journal of Industrial Engineering and Management				x	0.35
42	International Journal of Logistics Systems and Management				x	0.33
43	Risk Management				x	0.29
44	Management Review Quarterly				x	0.25
45	International Journal of Information Systems and Supply Chain Management				x	0.25
46	Quality - Access to Success				x	0.23
47	International Journal of Business Science and Applied Management				x	0.21
48	International Journal of Supply Chain Management				x	0.2
49	International Journal of Operations and Quantitative Management				x	0.13
50	Advanced Science Letters				x	0.12
51	International Journal of Applied Business and Economic Research				x	0.11
52	European Journal of Purchasing and Supply Management			x		/
53	Journal of Operations and Supply Chain Management				x	/
54	International Journal of Networking and Virtual Organizations				x	/
55	Journal of Advances in Management Research				x	/
56	The Journal of Business & Industrial Marketing				x	/
57	IUP Journal of Supply Chain Management				x	/
58	Journal of Supply Chain Management Systems				x	/
59	International Journal of Sustainable Strategic Management				x	/
60	Journal of Developing Areas				x	/

REFERENCES

Abernathy, W. J. 1978. *The productivity dilemma: Roadblock to innovation in the industry*.: Baltimore: Johns Hopkins University Press.

Adlakha, V., & Kowalski, K. 1998. A quick sufficient solution to the more-for-less paradox in the transportation problem. *Omega*, 26(4): 541-547.

Adlakha, V., & Kowalski, K. 2000. A note on the procedure MFL for a more-for-less solution in transportation problems. *Omega*, 28(4): 481-483.

Adlakha, V., Kowalski, K., Vemuganti, R. R., & Lev, B. 2007. More-for-less algorithm for fixed charge transportation problems. *Omega*, 35(1): 116-127.

Adler, P. S., Goldoftas, B., & Levine, D. I. (1999). Flexibility versus efficiency? A case study of model changeovers in the Toyota production system. Organization science, 10(1), 43-68.

Adler, P. S., Benner, M., Brunner, D. J., MacDuffie, J. P., Osono, E., Staats, B. R., Takeuchi, H., Tushman, M., & Winter, S. G. 2009. Perspectives on the productivity dilemma. *Journal of Operations Management*, 27(2): 99-113.

Agarwal, R., & Prasad, J. 1997. The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision sciences*, 28(3): 557-582.

Andriopoulos, C., & Lewis, M. W. 2009. Exploitation-exploration tensions and organizational ambidexterity: Managing paradoxes of innovation. *Organization Science*, 20(4): 696-717.

Barros, L., & Hilmola, O. 2007. Quantifying and modelling logistics at business and macro levels. *International Journal of Logistics Systems and Management*, 3(4): 382-394.

Bartunek, J.M. & Rynes, S.L., 2014. Academics and practitioners are alike and unlike: The paradoxes of academic–practitioner relationships. *Journal of Management,* 40(5), 1181–1201

Basso, F., D'Amours, S., Rönnqvist, M., & Weintraub, A. 2019. A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research*.

Bátiz-Lazo, B. 2004. Benchmarking financial services and online innovations. *Benchmarking: An International Journal*, 11(5): 431-446.

Birkinshaw, J., C. Gibson. 2004. Building ambidexterity into an organization. *MIT Sloan Management Review,* 45 47-55.

Birkinshaw, J., & Gupta, K. 2013. Clarifying the distinctive contribution of ambidexterity to the field of organization studies. *Academy of Management Perspectives*, 27(4): 287-298.

Blome, C., Schoenherr, T., & Kaesser, M. 2013. Ambidextrous governance in supply chains: The impact on innovation and cost performance. *Journal of Supply Chain Management*, 49(4): 59 80.

Brix-Asala, C., Geisbüsch, A., Sauer, P. C., Schöpflin, P., & Zehendner, A. 2018. Sustainability tensions in supply chains: A case study of paradoxes and their management. *Sustainability*, 10(2): 424.

Brockett, P. L., Cooper, W. W., Kwon, K. H., & Ruefli, T. W. 2003. Commentary on "a review of research on the negative accounting relationship between risk and return: Bowman's paradox by MN nickel and MC rodriguez". *Omega*, 31(5): 409-412.

Busse, C., Kach, A. P., & Bode, C. 2016. Sustainability and the false sense of legitimacy: How institutional distance augments risk in global supply chains. *Journal of Business Logistics*, 37(4): 312-328.

Cerruti, C., Mena, C., Skipworth, H., & Tavoletti, E. 2016. Characterizing agile supply partnerships in the fashion industry. *International Journal of Operations & Production Management*, 36(8): 923-947.

Chiadamrong, N., & Wajcharapornjinda, P. 2012. Developing an economic cost model for quantifying supply chain costs. *International Journal of Logistics Systems and Management*, 13(4): 540-571.

Choi, T. Y., & Eboch, K. 1998. The TQM paradox: Relations among TQM practices, plant performance, and customer satisfaction. *Journal of Operations Management*, 17(1): 59-75.

Colicchia, C., & Strozzi, F. (2012). Supply chain risk management: a new methodology for a systematic literature review. *Supply Chain Management: An International Journal*.

D'Ambrosio, C., Gentili, M., & Cerulli, R. (2019). The optimal value range problem for the interval (immune) transportation problem. *Omega*.

Duncan, R. B. 1976. The ambidextrous organization: Designing dual structures for innovation. *The Management of Organization*, 1: 167-188.

Durach, C. F., Kembro, J., & Wieland, A. 2017. A new paradigm for systematic literature reviews in supply chain management. *Journal of Supply Chain Management*, 53(4): 67-85.

Duray, R., Ward, P. T., Milligan, G. W., & Berry, W. L. 2000. Approaches to mass customization: Configurations and empirical validation. *Journal of Operations Management*, 18(6): 605-625.

Fabbe-Costes, N., & Jahre, M. (2008). Supply chain integration and performance: a review of the evidence. *The International Journal of Logistics Management*, *19*(2), 130-154.

Ghoshal, S., & Bartlett, C. A. 1994. Linking organizational context and managerial action: The dimensions of quality of management. *Strategic Management Journal*, 15(S2): 91-112.

Gibson, C. B., & Birkinshaw, J. 2004. The antecedents, consequences, and mediating role of organizational ambidexterity. *Academy of management Journal*, 47(2): 209-226.

Graetz, F., & Smith, A. C. 2008. The role of dualities in arbitrating continuity and change in forms of organizing. *International Journal of Management Reviews*, 10(3): 265-280.

Gualandris, J., Legenvre, H., & Kalchschmidt, M. 2018. Exploration and exploitation within supply networks: Examining purchasing ambidexterity and its multiple performance implications. *International Journal of Operations & Production Management*, 38(3): 667-689.

"Henry" Jin, Y., Fawcett, A. M., & Fawcett, S. E. (2013). Awareness is not enough: Commitment and performance implications of supply chain integration. *International Journal of Physical Distribution & Logistics Management*, *43*(3), 205-230.

Hult, G. T. M. (2011). Toward a theory of the boundary-spanning marketing organization and insights from 31 organization theories. *Journal of the Academy of Marketing Science*, *39*(4), 509 536.

Im, G., & Rai, A. 2008. Knowledge sharing ambidexterity in long-term interorganizational relationships. *Management Science*, 54(7): 1281-1296.

Kannothra, C. G., Manning, S., & Haigh, N. 2018. How hybrids manage growth and social business tensions in global supply chains: The case of impact sourcing. *Journal of Business Ethics*, 148(2): 271-290.

Kastalli, I. V., & Van Looy, B. 2013. Servitization: Disentangling the impact of service business model innovation on manufacturing firm performance. *Journal of Operations Management*, 31(4): 169-180.

Keller, J., & Sadler‐Smith, E. (2019). Paradoxes and Dual Processes: A Review and Synthesis. *International Journal of Management Reviews*, *21*(2), 162-184.

Khazanchi, S., Lewis, M. W., & Boyer, K. K. 2007. Innovation-supportive culture: The impact of organizational values on process innovation. *Journal of Operations Management*, 25(4): 871 884.

Kim, D., Tamer Cavusgil, S., & Calantone, R. J. 2005. The role of information technology in supply-chain relationships: Does partner criticality matter? *Journal of Business & Industrial Marketing*, 20(4/5): 169-178.

Klumpp, M. 2016. To green or not to green: A political, economic and social analysis for the past failure of green logistics. *Sustainability*, 8(5): 441.

Kolk, A. 2012. Towards a sustainable coffee market: Paradoxes faced by a multinational company*Corporate Social Responsibility and Environmental Management*, 19(2): 79-89.

Koufteros, X., Droge, C., Heim, G., Massad, N., & Vickery, S. K. 2014. Encounter satisfaction in e-tailing: Are the relationships of order fulfillment service quality with its antecedents and consequences moderated by historical satisfaction? *Decision Sciences*, 45(1): 5-48.

Kristal, M. M., Huang, X., & Roth, A. V. 2010. The effect of an ambidextrous supply chain strategy on combinative competitive capabilities and business performance. *Journal of Operations Management*, 28(5): 415-429.

Kronborg Jensen, J. 2012. Product carbon footprint developments and gaps. *International Journal of Physical Distribution & Logistics Management*, 42(4): 338-354.

Kull, T. J., Barratt, M., Sodero, A. C., & Rabinovich, E. 2013. Investigating the effects of daily inventory record inaccuracy in multichannel retailing. *Journal of Business Logistics*, 34(3): 189 208.

Lavie, D., Stettner, U., & Tushman, M. L. 2010. Exploration and exploitation within and across organizations. *The Academy of Management Annals*, 4(1): 109-155.

Lee, S. M., & Rha, J. S. 2016. Ambidextrous supply chain as a dynamic capability: Building a resilient supply chain. *Management Decision*, 54(1): 2-23.

Levinthal, D. A., & March, J. G. 1993. The myopia of learning. *Strategic Management Journal*, 14(S2): 95-112.

Lewis, M. W. (2000). Exploring paradox: Toward a more comprehensive guide. Academy of Management review, 25(4), 760-776.

Lewis, M. W., & Smith, W. K. (2014). Paradox as a metatheoretical perspective: Sharpening the focus and widening the scope. *The Journal of Applied Behavioral Science*, *50*(2), 127-149.

Li, G., Wang, S., Yan, H., & Yu, G. 2005. Information transformation in a supply chain: A simulation study. *Computers & Operations Research*, 32(3): 707-725.

Looney, C. A., Valacich, J. S., Todd, P. A., & Morris, M. G. 2006. Paradoxes of online investing: Testing the influence of technology on user expectancies. *Decision Sciences*, 37(2): 205-246.

Mak, V., Seale, D. A., Gisches, E. J., Yang, R., Cheng, M., Moon, M., & Rapoport, A. 2018. The braess paradox and coordination failure in directed networks with mixed externalities. *Production and Operations Management*, 27(4): 717-733.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2(1): 71-87.

Masuda, Y., & Whang, S. 2002. Capacity management in decentralized networks. *Management Science*, 48(12): 1628-1634.

Matthews, L., Power, D., Touboulic, A., & Marques, L. 2016. Building bridges: Toward alternative theory of sustainable supply chain management. *Journal of Supply Chain Management*, 52(1): 82-94.

Mellat-Parast, M., & Digman, L. A. 2008. Learning: The interface of quality management and strategic alliances. *International Journal of Production Economics*, 114(2): 820-829.

Murali, K., Lim, M. K., & Petruzzi, N. C. 2015. Municipal groundwater management: Optimal allocation and control of a renewable natural resource. *Production and Operations Management*, 24(9): 1453-1472.

Narasimhan, R., & Narayanan, S. 2013. Perspectives on supply network–enabled innovations. *Journal of Supply Chain Management*, 49(4): 27-42.

Naylor, J.B., Naim, M.M. and Berry, D., 1999. Leagility: Integrating the lean and agile manufacturing paradigms in the total supply chain. *International Journal of Production Economics,* 62(1-2), pp.107-118.

New, S. J. 2015. Modern slavery and the supply chain: The limits of corporate social responsibility? Supply Chain Management: An International Journal, 20(6): 697-707.

Niesten, E., & Stefan, I. (2019). Embracing the Paradox of Interorganizational Value Co-creation Value Capture: A Literature Review towards Paradox Resolution. *International Journal of Management Reviews*, *21*(2), 231-255.

O'Reilly III, C. A., & Tushman, M. L. 2013. Organizational ambidexterity: Past, present, and future. *Academy of Management Perspectives*, 27(4): 324-338.

Pal, R., Harper, S., & Vellesalu, A. 2018. Competitive manufacturing for reshoring textile and clothing supply chains to high-cost environment: A delphi study. *The International Journal of Logistics Management*, 29(4): 1147-1170.

Pereira, M. J. 2004. Impacts of information systems and technology on productivity and competitiveness of the portuguese banking sector: An empirical study. *International Transactions in Operational Research*, 11(1): 43-62.

Raisch, S., & Birkinshaw, J. 2008. Organizational ambidexterity: Antecedents, outcomes, and moderators. *Journal of Management*, 34(3): 375-409.

Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. 2009. Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organization Science*, 20(4): 685-695.

Rapoport, A., Gisches, E. J., & Mak, V. 2014. Distributed decisions in networks: Laboratory study of routing splittable flow. *Production and Operations Management*, 23(2): 314-331.

Rindova, V. 2011. Moving from ideas to a theoretical contribution: Comments on the process of developing theory in organizational research. *Journal of Supply Chain Management*, 47(2): 19 21.

Rizzi, F., Frey, M., Testa, F., & Appolloni, A. 2014. Environmental value chain in green SME networks: The threat of the abilene paradox. *Journal of Cleaner Production*, 85: 265-275.

Rojo, A., Llorens-Montes, J., & Perez-Arostegui, M. N. 2016. The impact of ambidexterity on supply chain flexibility fit. *Supply Chain Management: An International Journal*, 21(4): 433 452.

Sandberg, E. 2017. Introducing the paradox theory in logistics and SCM research–examples from a global sourcing context. *International Journal of Logistics Research and Applications*, 20(5): 459-474.

Sarkis, J., & Semple, J. H. 1999. Vendor selection with bundling: A comment. *Decision Sciences*, 30(1): 265-271.

Schad, J., Lewis, M. W., Raisch, S., & Smith, W. K. 2016. Paradox research in management science: Looking back to move forward. *The Academy of Management Annals*, 10(1): 5-64.

Schad, J., Lewis, M. W., & Smith, W. K. 2019. Quo vadis, paradox? Centripetal and centrifugal forces in theory development. *Strategic Organization*, 17(1): 107-119.

Schmidt, C. G., Foerstl, K., & Schaltenbrand, B. 2017. The supply chain position paradox: Green practices and firm performance. *Journal of Supply Chain Management*, 53(1): 3-25.

Seuring, S., & Gold, S. (2012). Conducting content - analysis based literature reviews in supply chain management. *Supply Chain Management: An International Journal*.

Shalley, C. E., & Gilson, L. L. 2017. Creativity and the management of technology: Balancing creativity and standardization. *Production and Operations Management*, 26(4): 605-616.

Shmueli, G., & Yahav, I. 2018. The forest or the trees? tackling simpson's paradox with classification trees. *Production and Operations Management*, 27(4): 696-716.

Simsek, Z. 2009. Organizational ambidexterity: Towards a multilevel understanding. *Journal of Management Studies*, 46(4): 597-624.

Smith, W. K., & Lewis, M. W. 2011. Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of Management Review*, 36(2): 381-403.

Smith, W. K., Lewis, M. W., Jarzabkowski, P., & Langley, A. 2017. *The Oxford Handbook of Organizational Paradox* Oxford University Press.

Sousa, R., & Voss, C. A. 2009. The effects of service failures and recovery on customer loyalty in e-services: An empirical investigation. *International Journal of Operations & Production Management*, 29(8): 834-864.

Stefan, I., & Bengtsson, L. 2017. Unravelling appropriability mechanisms and openness depth effects on firm performance across stages in the innovation process. *Technological Forecasting and Social Change*, 120: 252-260.

Steinbach, T., Wallenburg, C. M., & Selviaridis, K. 2018. Me, myself and I: Non-collaborative customer behavior in service outsourcing–the key role of outcome orientation and outcome attributability. *International Journal of Operations & Production Management*.

Storer, M., Hyland, P., Ferrer, M., Santa, R., & Griffiths, A. 2014. Strategic supply chain management factors influencing agribusiness innovation utilization. *The International Journal of Logistics Management*, 25(3): 487-521.

Tazelaar, F., & Snijders, C. 2013. Operational risk assessments by supply chain professionals: Process and performance. *Journal of Operations Management*, 31(1-2): 37-51.

Terziovski, M., & Guerrero, J. 2014. ISO 9000 quality system certification and its impact on product and process innovation performance. *International Journal of Production Economics*, 158: 197-207.

Tokman, M., Richey, R. G., Marino, L. D., & Weaver, K. M. 2007. Exploration, exploitation and satisfaction in supply chain portfolio strategy. *Journal of Business Logistics*, 28(1): 25-56.

Turner, N., Swart, J., & Maylor, H. 2013. Mechanisms for managing ambidexterity: A review and research agenda. *International Journal of Management Reviews*, 15(3): 317-332.

Tushman, M. L., & O'Reilly III, C. A. 1996. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review*, 38(4): 8-29.

Tyworth, J. E. 2018. A note on lead-time paradoxes and a tale of competing prescriptions. *Transportation Research Part E: Logistics and Transportation Review*, 109: 139-150.

Tyworth, J. E., & Saldanha, J. 2014. The lead-time reliability paradox and inconsistent value-of reliability estimates. *Transportation Research Part E: Logistics and Transportation Review*, 70: 76-85.

Upton, D. M., & McAfee, A. P. 1998. Computer integration and catastrophic process failure in flexible production: An empirical investigation. *Production and Operations Management*, 7(3): 265-281.

Vedel, M. 2016. The triad value function–theorizing the value potential of connected relationships.

Journal of Business & Industrial Marketing, 31(7): 849-860.

Voordijk, H., De Haan, J., & Joosten, G. 2000. Changing governance of supply chains in the building industry: A multiple case study. *European Journal of Purchasing & Supply Management*, 6(3-4): 217-225.

Wacker, J. G. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. Journal of operations management, 16(4), 361 385.

Wacker, J. G., & Lummus, R. R. 2002. Sales forecasting for strategic resource planning. *International Journal of Operations & Production Management*, 22(9): 1014-1031.

Waller, M. A., Fawcett, S. E., & Johnson, J. L. 2015. The luxury paradox: How systems thinking and supply chain collaboration can bring sustainability into mainstream practice. *Journal of Business Logistics*, 36(4): 303-305.

Wilhelm, M., & Sydow, J. 2018. Managing coopetition in supplier Networks–A paradox perspective. *Journal of Supply Chain Management*.

Wong, C. W., Wong, C. Y., & Boon-itt, S. 2013. The combined effects of internal and external supply chain integration on product innovation. *International Journal of Production Economics*, 146(2): 566-574.

Xiao, C., Wilhelm, M., van der Vaart, T., & van Donk, D. P. 2019. Inside the buying firm: Exploring responses to paradoxical tensions in sustainable supply chain management. *Journal of Supply Chain Management*, 55(1): 3-20.

Yang, C., & Chen, A. 2009. Sensitivity analysis of the combined travel demand model with applications. *European Journal of Operational Research*, 198(3): 909-921.

Yin, X., & George Zhang, Z. (2019). On Downs–Thomson paradox in two-tier service systems with a fast pass and revenue-based capacity investment. *Journal of the Operational Research Society*, 1-14.

Zhao, C., Fu, B., & Wang, T. 2014. Braess paradox and robustness of traffic networks under stochastic user equilibrium. *Transportation Research Part E: Logistics and Transportation Review*, 61: 135-141.

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. 2018. How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. *International Journal of Physical Distribution & Logistics Management*, 48(1): 47-68.

DECISION SCIENCES INSTITUTE

An Empirical Study of Buyer-Supplier Collaboration

Osam Sato
Tokyo Keizai University
Email: osamsato@tku.ac.jp

Yoshiki Matsui
Open University of Japan & Yokohama National University
Email: ymatsui@ynu.ac.jp

ABSTRACT

Buyer-supplier collaboration is an important approach for successful supply chain management. Although it has been studied for a long time and some anecdotal evidences were introduced, study on typical causal patterns of supply chain orientation (SCO) between buyers and suppliers is not enough. Sandberg (2007) wrote “more research, especially empirical, is needed in order to verify existing literature.” We, based on the relational view, propose four hypotheses that explain buyer-supplier collaboration performance with a set of SCO aspects. We conducted an empirical study with data collected from manufacturing companies using SEM. Our study tries to narrow the gap and show a pattern of causal relationship.

KEYWORDS: SCM, Buyer-supplier collaboration, Information sharing, Relational view, and SEM

INTRODUCTION

Nowadays, any organizations do not operate alone. They are networked to many other partners and a member of one or more supply chains (Lofti et al., 2013). Christopher & Peck (2004) define “supply chain as the network of organizations that are involved, through upstream and downstream linkages, in the different processes and activities that produce value in the form of products and services in the hands of the ultimate consumer.” Buyer-supplier collaboration, which is also called as supplier relationship management (SRM) (Dubey et al., 2019), is claimed as an important approach for successful supply chain management (SCM). Rubber & Plastics News *(*2019) wrote “there's one general truth about the relationships between auto makers and their suppliers that doesn't change: The better the two sides work together, the more likely both sides will be successful.” Although it has been said for a long time and some anecdotal evidences were introduced, study on popular or typical causal patterns of supply chain orientation (SCO) between buyers and suppliers that are endorsed empirically is not still enough. Sandberg (2007) wrote “more research, especially empirical, is needed in order to verify existing literature.” Buyer-supplier collaboration “is an increasingly important unit of analysis and, therefore, deserves more study” (Dyer & Singh, 1998). Our study tries to narrow the gap with an empirical study and show a pattern of antecedent factors for successful buyer-supplier collaboration.

LITERATURE REVIEW

Collaboration

Collaboration is "a process where two or more companies formally undertake to share responsibility for sharing information, planning, and managing process execution to enable the formation of synergies between all supply chain members, thereby achieving greater mutual benefit than the firms would achieve individually" (Botes & Kotzé, 2017). Collaboration theory recommended cooperative transaction or inter-firm relation as a critical factor for corporate success (Frydlinger et al., 2013). According to the collaboration theory, successful collaboration should be based on embedded ties (Frydlinger et al., 2013). The embedded ties are "characterized by trust, fine-tuned information transfer, and joint problem-solving arrangements" (Uzzi, 1996). The collaboration theory claims that successful collaboration is a source of economic performance of organizations (Shekarian, 2018).

Relational View

The relational view is a stem of collaboration theory and claimed that idiosyncratic inter-firm linkage "is an important source of competitive advantage" (Dyer & Singh, 1998; Soosay & Hyland, 2015). The relationship is an economic rent. It reduces a variety of transaction costs through reduction of uncertainty, bounded rationality, and opportunistic behavior, and increases mutual learning and pooling of resources (Dyer & Singh, 1998；Frydlinger et al., 2013). Therefore, buyer and supplier need "relationship-specific investments" (Wagner & Krause, 2009) to realize these benefits.

The relationship is often based on self-enforcing mechanisms. It is "more effective than third-party enforcement mechanisms at minimizing transaction costs and maximizing value-creation initiatives both. Transaction costs are lower under self-enforcing agreements for four primary reasons" (Dyer & Singh, 1998).

Sandberg (2007) conducted a survey study of SCM logistics in Sweden, and found that supply chain orientation (SCO) between buyers and suppliers reduces cost and improves services not only of buyers and suppliers but also of buyers' customers. It concluded:

- There is a strong relationship among the process orientation, the degree of joint operational planning, and intensity of information sharing.
- The intensity of these aspects are positively related to the experienced effects of the collaboration.

However, Sandberg (2007) could not "make a strictly causal order of the aspects."

Transaction Cost Reduction

Dyer and Singh (1998) wrote "growing body of literature on transaction value is emphasizing the influence of governance on the value-creation initiatives of alliance partners. Effective governance can generate relational rents by either (1) lowering transaction costs or (2) providing incentives for value-creation initiatives, such as investing in relation-specific assets, sharing knowledge, or combining complementary strategic resources."

Transaction cost theory predicts that trust-based inter-firm linkage reduces transaction cost and expedites manufacturing process (Williamson, 1994).

HYPOTHESES DEVELOPMENT

In modern economy, companies play a role as a member in a supply chain network. Therefore, buyer-supplier relationship, and especially, buyer-supplier collaboration are important based on the collaboration theory and relational view. Supply chain collaboration (SCC) is critical to manage risk and realize supply chain resilience (SCR). Botes & Kotzé (2017) wrote “SCC aims to leverage the capabilities and knowledge of key suppliers under market uncertainty”, and to “improve operational flexibility by facilitating cooperation between (the) various supply chain members.” Rubber & Plastics News *(*2019) introduced that “the Toyota group vice president of purchasing cites the ‘three Cs’ as key. That's as in communication, collaboration and consistency.”

Resilience is defined as “the ability of a system to return to its original state or move to a new, more desirable state after being disturbed” (Christopher and Peck, 2004). Christopher and Peck (2004) distinguish it from robustness. Collaborative buyer-supplier relationship or creative working is a critical construct of SCR. SCR increases “supply chain’s ability to react to and absorb supply disruptions and return to the operational status quo” (Botes & Kotzé, 2017). The management philosophy is supply chain orientation (SCO), which includes soft factors such as trust, organizational compatibility, commitment, vision, and information sharing (Mentzer et al., 2001). They are requested to improve buyer-supplier collaboration performance. Companies cannot optimize their “supply chain operations until the constraints of suppliers and demands of customers are understood” (Botes & Kotzé, 2017).

Botes et al. (2017) wrote that one of three fundamental elements of vertical SCC was information sharing. Information sharing is to distribute “useful information for systems, people, or organizational units,” which “serves as an essential approach for the survival of enterprises and enabler of supply chain integration” (Lofti et al., 2013). Information sharing includes ‘knowledge sharing’ and ‘information integration’. Shared knowledge includes both explicit and implicit knowledge. Lofti et al. (2013) listed 14 benefits of supply chain information sharing based on the literature review. Supply chain information sharing by suppliers such as forecast (Sandberg, 2007) is critical to improve buyer-supplier collaboration performance.

Of course, supply chain information sharing by buyers is also critical. However, because this study views the problem not from supplier’s but from buyer’s point of view, supply chain information sharing by buyers is not considered. Cooperation and coordination of activities between buyers and suppliers include information sharing, joint planning, joint demand management, joint inventory management, and sharing of goals and objectives. They are critical for the creation of supply chain-wide velocity, visibility, and flexibility (Christopher & Peck, 2004; Botes et al., 2017).

Among some factors that affect buyer-supplier collaboration performance, we pick up information sharing, agreement on vision and goal, and SCM planning because they are popular activities as buyer-supplier collaboration. Sandberg (2007) wrote “typical collaborative logistics activities are joint planning and information sharing.” Christopher & Peck (2004) suggested collaborative planning and supply chain intelligence were two important sources for supply chain collaboration. Both buyer and supplier invest in relation-specific resources that include information sharing to realize the relational rent (Wagner & Krause, 2009). Therefore,

H1: Supply chain information sharing by suppliers affects buyer-supplier collaboration performance.

In reality, there are cooperative processes between information sharing and performance evaluation (Sandberg, 2007). However, as this study employs cross-sectional data analysis, the process view is neglected in our study (Figure 1). Agreement on supply chain vision and goals should be a base of supply chain information sharing by suppliers. Wagner & Krause (2009) wrote "When making investments in the buyer-supplier relationship, the buying firm should determine what goals it has for the relationship." Therefore,

H2: Agreement on supply chain vision and goals affects supply chain information sharing by suppliers.

Supply base reduction (SBR) is an important supplier management approach. De Toni and Nassimbeni (1999) found that SBR policies were positively related to the exclusive buyer-supplier relational exchange practice. Wagner & Krause (2009) suggests "the relationship between the goal to improve a supplier's capabilities and knowledge transfer from the buyer to the supplier firm is moderated by the degree of human interaction." However, supply base reduction is necessary in advance to manage human interaction and agree on supply chain vision and goals. SBR is a prerequisite to conduct "training and co-location of buyer and supplier employees to leverage the knowledge transfer to the supplier firm" (Wagner & Krause, 2009). Therefore,

H3: Supply base reduction affects agreement on supply chain vision and goals with suppliers.

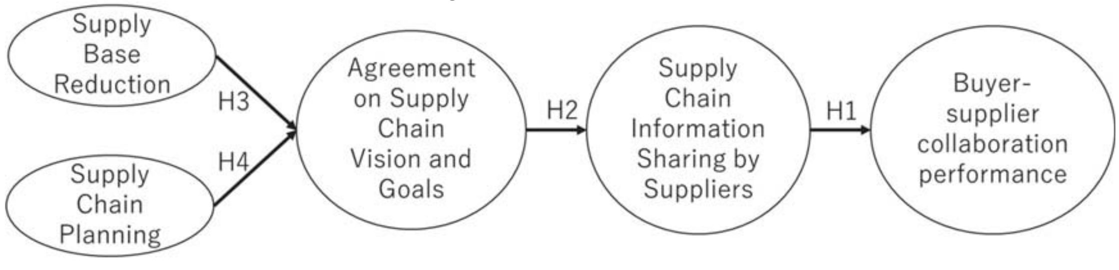
Visibility, velocity, and flexibility are typical SCR antecedents (Botes et al., 2017). Among them, upstream visibility requires collaborative planning (Christopher & Peck, 2004). Botes et al., (2017) wrote that collaborative supply chain planning "aims to balance supply and demand within a supply chain network through demand-driven processes." This balance is enabled through demand and supply information sharing, coordinated sales forecasts, efficient inventory management, effective materials management for production, and better performance management.

Supply chain planning is necessary as a base or antecedent for agreement on supply chain vision and goals with suppliers because it can improve visibility. Also, supply chain planning improves agreement on supply chain vision and goal by reducing uncertainty. Therefore,

H4: Supply chain planning affects agreement on supply chain vision and goals with suppliers.

Summing up, Figure 1 illustrates our research model.

Figure 1: Research model



RESEARCH DATA AND METHOD

Data

We employ the fourth round dataset of High Performance Manufacturing (HPM) project (Phan, 2014) because it is the latest version available. The fourth round of HPM project started in 2012-2013 and the data were collected in 2014-2015. Table 1 shows the distribution of sample size by country. The total sample size amounts to 306.

Table 1: Sample size by country

Country and Region	Count	%
Brazil	24	7.8
China	30	9.8
Spain	25	8.2
Finland	17	5.6
Germany	28	9.2
Israel	26	8.5
Italy	29	9.5
Japan	22	7.2
Korea	26	8.5
Sweden	9	2.9
Taiwan	30	9.8
United Kingdom	13	4.2
Vietnam	25	8.2
Missing	2	0.7
Total	306	100.0

Table 2 shows the sample size by industry. The data were collected from factories in electric & electronics, machinery, and transportation equipment. Transportation equipment factory includes automobile assemblers and their parts manufacturers.

Table 2: Data distribution by industry

Industry	Count	%
Electronics	114	37.3
Machinery	110	35.9
Transportation	80	26.1
Missing	2	0.7
Total	306	100.0

We picked up some relevant constructs in the dataset, set the research model in Figure 1 and hypotheses based on the above discussion, and tested it empirically.

Method

We employed Structural Equation Modeling (SEM, AMOS v.25) to study our research model in Figure 1. Each construct consisted of some question items. Table 3 shows these constructs and their questions items.

Table 3: Constructs and Question Items

Constructs and variables	Questions
RED	Supply base reduction
RED1	We rely on a small number of high quality suppliers. (Is this an item from an HPM scale?)
RED2	We maintain a close relationship with a limited pool of suppliers.
RED3	We try to keep our supply base small.
PLAN	Supply chain planning
PLAN1	We actively plan supply chain activities.
PLAN2	We consider our customers' forecasts in our supply chain planning.
PLAN3	We strive to manage each of our supply chains as a whole.
PLAN4	We monitor the performance of members of our supply chains, in order to adjust supply chain plans.
PLAN5	We gather indicators of supply chain performance.
VIS	Agreement on supply chain vision & goals
VIS1	Our supply chain members understand our goals for supply chain management.
VIS2	Our supply chain members understand that we expect them to continuously improve their supply chain practices and operations.
VIS3	Our supply chain members have clearly defined goals within our supply chain.
VIS4	We all know which supply chain members are responsible for particular goals with our supply chains.
BYS	Supply chain information sharing by suppliers: Our plant has access to the following information about our key suppliers
BYS1	Delivery information
BYS2	Demand change information
BYS3	Demand forecast information
BYS4	Inventory information
EVAL	Buyer-supplier collaboration performance: We are satisfied with the performance of our key suppliers on the following criteria:
EVAL1	On-time delivery
EVAL2	Product liability
EVAL3	Service level
EVAL4	Willingness to adapt processes to meet your changing needs

First, before employing SEM, we did explorative factor analysis (EFA, SPSS version 24) with a promax rotation based on principal axis factor analysis. Table 4 shows the result. We confirmed five constructs as expected. We believe it is a satisfactory result that allows us to use the data.

Table 4: EFA Pattern Matrix

RED1	0.028	-0.046	-0.020	-0.071	0.750
RED2	-0.003	-0.031	-0.016	-0.012	0.639
RED3	0.016	0.088	0.048	0.044	0.586
PLAN1	-0.004	0.808	-0.002	-0.065	-0.076
PLAN2	0.138	0.679	-0.021	-0.154	-0.005
PLAN3	-0.016	0.842	0.100	-0.062	-0.027
PLAN4	-0.034	0.753	-0.035	0.177	0.036
PLAN5	-0.072	0.678	-0.071	0.118	0.071
VIS1	0.020	-0.020	0.838	0.039	0.033
VIS2	-0.117	0.000	0.653	0.049	0.217
VIS3	0.086	0.043	0.741	0.034	-0.100
VIS4	0.018	-0.030	0.812	-0.074	-0.051
BYS1	0.657	-0.022	0.103	0.137	-0.135
BYS2	0.965	0.022	-0.009	-0.020	0.016
BYS3	0.976	0.006	0.009	-0.073	0.000
BYS4	0.711	0.007	-0.073	0.085	0.143
EVAL1	0.100	-0.086	0.007	0.700	0.029
EVAL2	-0.023	-0.001	-0.049	0.880	-0.124
EVAL3	0.029	0.014	0.070	0.673	0.018
EVAL4	0.036	0.047	0.016	0.510	0.116

Second, we conducted a confirmatory factor analysis (CFA) to confirm unidimensionality of each construct. “Unidimensionality is the existence of one latent trait or construct underlying a set of measures” (Anderson, et. al, 1988). “Assessing unidimensionality means determining whether indicators reflect one, as opposed to more than one construct” (Li et al., 2009). Table 3 shows the item factor loadings with average variance extracted (AVE), square root of AVE, composite reliability (CR), and Cronbach’s alpha. All item factor loadings are larger than 0.5 ($p<0.01$), above the suggested minimum of 0.30 (Schoenherr & Swink, 2012). AVE is used to indicate convergent validity. AVE should be greater than 0.50 (Dinger et al., 2015). All AVE are upper than the cut-off rate. All composite Reliability (CRs) exceed 0.7 (Gefen, Straub, & Boudreau, 2000). Cronbach’s alpha suggests reliability in terms of internal consistency. Its value for REDUC is less than 0.7. If it were 0.7 or higher, it could be considered acceptable, but with 0.6 being acceptable for new scales (Shin, 2000). Bhatnagar et al. (2016) wrote that Cronbach’s alpha between 0.6 and 0.7 was acceptable.

We checked common method bias (CMB) with Harman’s single-factor method (Dubey et al., 2019). We fixed the number of factors equal to one, and computed the un-rotated factor solution. Then, a single factor was obtained which explained 31.198 percent of the variance. It is less than 50 percent. Therefore, we can find that CMB is not a problem for this study and continue our analysis.

Table 5: CFA, Convergence Validity, and Reliability

Variable	Loading	AVE	CR	Sqrt (AVE)	Cronbach’s Alpha
RED1	0.654	0.528	0.770	0.726	0.672
RED2	0.588				
RED3	0.667				
PLAN1	0.687	0.567	0.930	0.753	0.852

PLAN2	0.614				
PLAN3	0.863				
PLAN4	0.820				
PLAN5	0.688				
VIS1	0.887	0.761	0.959	0.872	0.853
VIS2	0.708				
VIS3	0.773				
VIS4	0.724				
BYS1	0.740	0.838	0.975	0.916	0.906
BYS2	0.972				
BYS3	0.941				
BYS4	0.720				
EVAL1	0.748	0.672	0.940	0.820	0.814
EVAL2	0.723				
EVAL3	0.767				
EVAL4	0.652				

Table 6 shows the standardized correlations among constructs.

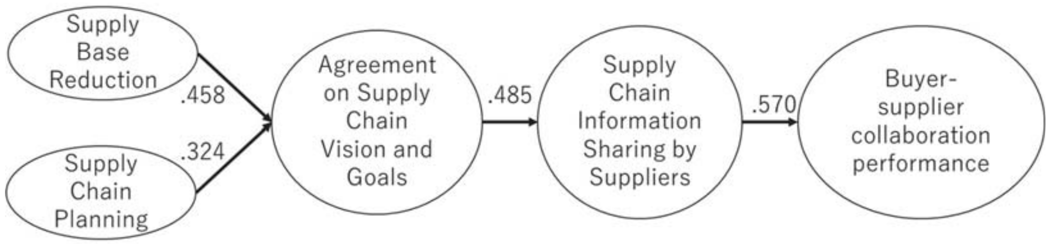
Table 6: Standardized Correlations Among Constructs			
PLAN	<-->	VIS	0.387
BYS	<-->	RED	0.212
PLAN	<-->	BYS	0.179
PLAN	<-->	EVAL	0.311
PLAN	<-->	REDUCN	0.137
EVAL	<-->	REDUCN	0.464
VIS	<-->	EVALSX	0.685
VIS	<-->	REDUCN	0.507
BYS	<-->	EVALSX	0.557
VIS	<-->	BYS	0.475

The minimum value of square-roots of AVE in Table 5 (0.726) is greater than the maximum of inter-construct correlations (0.685), which confirms discriminant validity (Fornell & Larcker, 1981). This suggests that the constructs of our model possess the discriminant validity.

RESEARCH RESULT

We apply SEM to the model in Figure 1 now. Figure 2 shows the result. All standardized coefficients are significant strongly ($p<0.000$). Model-fit is not perfect but satisfactory as below. CMIN/DF=2.424, NFI=.870, IFI=.919, CFI=.918, FMIN=1.311, RMSEA=.068.

Figure 2: SEM result



DISCUSSION AND CONCLUSIONS
\
Discussion

We proposed a research model of buyer-supplier collaboration and derived four hypotheses, which were tested for the dataset of 306 manufacturing companies. The performance (satisfaction level) of buyer-supplier collaboration from buyers' point of view was affected by supply chain information sharing by suppliers, which was in turn affected by agreement on supply chain vision and goals between suppliers and buyers. Buyers set purchasing policy to supply base reduction and supply chain planning in advance of the agreement on supply chain vision and goals with suppliers. Supply base reduction is necessary to establish intimate relationship with suppliers and agree on vision and goals. Then, on the base of agreement, suppliers share their information with buyers, which finally improves supply chain performance assessment by buyers.

Above network of cause-effect relationship is within natural inference and not new. However, our empirical study result is unique. Firstly, because there are many possible path from policy-setting to performance, there is no past research that is same as our research model as far as we know. Secondly, although there are some past studies that suggest above relationship, they are fragmented and suggesting a part of our model, and there is no research that tested our long cause-effect network empirically.

Our results are congruent with Sandberg (2007), who wrote

- There is a strong relationship between the process orientations, the degree of joint operational planning, and intensity of information sharing.
- The intensity of these aspects is positively related to the experienced effects of the collaboration.

Our research result suggests that the conclusion by Sandberg (2007) is applicable to industry not only in Swedish but also in other advanced countries in our dataset.

Although Sandberg (2007) was descriptive and not an article of hypotheses-testing type, it tried to find causal effects of related factors with correlation and ANOVA analysis. Correlation analysis is not suitable or enough to find a cause-effect relationship. Our study result not only endorses the relationship among above aspects but also suggests a causal sequence among those five constructs with SEM that Sandberg (2007) could not.

Limitations and Future Research

There are some research limitations in our study. First, although sample size in our dataset is enough for SEM analysis (Wang & Wang, 2012), our study rely solely on the fourth round HPM dataset. We need to repeat this study with different and hopefully more data to check the ubiquity. Second, we included some SCO factors such as vision, goals, and information sharing, but there are more factors that must affect performance evaluation by buyers such as trust, organizational compatibility, and top management support (Dubey et al., 2019). Although we have these data in our dataset, we did not include them for model simplicity. Future studies need to include these factors. Third, we studied only buyer-supplier relationship from buyers' point of view, and did not include seller-customer relationship. The relationship of these two factors that Sandberg (2007) suggested may need to be included in future study. Fourth, Wagner & Krause (2009) wrote "various goals for supplier development efforts might affect the relationship between the buying firm and the supplier firm" differently, but we neglected these differences for simplicity. These are left for future research.

REFERENCES

Anderson, J. C., Gerbing, D. W., & Hunter, J. E. (1987). On the Assessment of Unidimensional Measurement - Internal and External Consistency, and Overall Consistency Criteria. Journal of Marketing Research, 24(4), 432-437.

Bhatnagar, R., Kim, J., & Many, J.E. (2014). Candidate Surveys on Program Evaluation: Examining Instrument Reliability, Validity and Program Effectiveness. American Journal of Educational Research, 2(8), 683-690.

Botes, A. W., & Kotzé, T. (2017). Buyer-Supplier Collaboration and Supply Chain Resilience: A Case Study in the Petrochemical Industry, South African Journal of Industrial Engineering, 28(4), 183-199.

Christopher, M., & Peck, H. (2004). Building the Resilient Supply Chain. International Journal of Logistics Management, 15(2), 1-13.

De Toni, A., & Nassimbeni, G., (1999). Buyer–supplier operational practices, sourcing policies and plant performance: result of an empirical research. International Journal of Production Research, 37(3), 597–619.

Dinger, M., Thatcher, J.B., Treadway, D., Stepina, L., & Breland, J. (2015). Does Professionalism Matter in the IT Workforce? An Empirical Examination of IT Professionals. Journal of the Association for Information Systems, 16(4), 281-313.

Dubey, R., Gunasekaran, A., Childe, S. J., Papadopoulos, T., & Helo, P. (2019). Supplier Relationship Management for Circular Economy. Management Decision 57(4), 767-790.

Dyer, J., & Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Strategy. Academy of Management Review, 23(4), 660-679.

Fornell, C. R., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(3), August, 39-50.

Frydlinger, D., Nyden, J., & Vitasek, K. (2013). Unpacking Collaboration Theory. (http://www.jnyden.com/wp-content/uploads/2013/08/Unpacking_Collaborative_Negotiations_.pdf)

Gefen, D., Straub, D., & Boudreau, M. (2000). Structural Equation Modeling Techniques and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, 7(7), 1-78.

Li, G., Yang, H., Sun, L. and Sohal, A.S. (2009). The impact of IT implementation on supply chain integration and performance. *International Journal of Production Economics*, 120(1), 125-138.

Lofti, Z., Mukhtar, M., Sahran, S., & Zadeh, A.T. (2013). Information Sharing in Supply Chain Management. *The 4th International Conference on Electrical Engineering and Informatics (ICEEI 2013)*, Procedia Technology 11, 298-304.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Soonhoong, M., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining Supply Chain Management. *Journal of Business Logistics*, 22(2), 1-25.

Phan, A. P. (2014). High Performance Manufacturing (HPM) Project Summary. http://dl.ueb.edu.vn/bitstream/1247/7579/2/Mr%20Chi%20Anh_En.pdf

Rubber & Plastics News (2019). Auto Makers Must Be Partners with Their Supply Base. *Rubber & Plastics News* 48(19), 8-8.

Sandberg, E. (2007). Logistics Collaboration in Supply Chains: Practice Vs. Theory. *International Journal of Logistics Management*, 18(2), 274-293.

Schoenherr, T & Swink, M. (2012). Revisiting the arcs of integration: Cross-validations and extensions. *Journal of Operations Management*, 30(1-2), 99-115.

Shekarian (2018). A literature Review on the Impact of Antecedents of Supply Chain Resilience on Mitigating Supply Chain Disruption Risks. *Proceedings of DSI annual Meeting*, November.

Shin, H. (2000). Supply management orientation and supplier/buyer performance. *Journal of Operations Management,* 18(3), 317-333.

Soosay, C. A., & Hyland, P. (2015). A Decade of Supply Chain Collaboration and Directions for Future Research. *Supply Chain Management*, 20(6), 613-630.

Uzzi, B. (1996). The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review,* 61(4), 674-698.

Wagner, S.M. & Krause, D.R. (2009). Supplier development: communication approaches, activities and goals. *International Journal of Production Research*, 47(12), 3161-3177.

Wang, X. & Wang, J. (2012). *Structural Equation Modeling: Applications Using Mplus*, John Wiley & Sons.

Williamson, O. E. (1994). Transaction Cost Economics and Organization Theory, in: *The Handbook of Economic Sociology B2 - The Handbook of Economic Sociology*, N.J. Smelser & R. Swedberg (eds.). Princeton, NJ: Princeton University Press, 77-107.

Managerial Decisions in Your Firm

DECISION SCIENCES INSTITUTE

Selling a Small Business After Tax Reform: New Benefits for Buyers; Sellers, Not So Much

John Cooper, Department of Accounting, College of Business and Economics,
California State University Los Angeles, 5151 State University Drive,
Los Angeles, CA 90032, jcooper3@calstatela.edu, (323) 343-5256

Nathan Oestreich, Charles W. Lamden School of Accountancy,
Fowler College of Business, San Diego State University, 5500 Campanile Drive,
San Diego, CA 92182-8221, drno@sdsu.edu, (858) 337-7815

ABSTRACT

Business owners operate their enterprises for profit and someday may harvest their accumulated capital. While many will transition their ownership through gifts and bequests, others will sell to unrelated, or potentially related, parties. This study attempts to explore the salient significant tax considerations to buyers and sellers in such transactions.

KEYWORDS: Sale of Business, Tax Reform, Tax Cuts and Jobs Act of 2017

INTRODUCTION

It is estimated that 9,919 sales of businesses closed in 2017, compared to 7,842 in 2016 (Small Biz). These numbers are from a survey of about 5,000 business brokers; however, considering that many sales occur without brokers, the actual number of business sales is likely much larger. Bob House, President of BizBuySell.com, reported that the small business market has been on the rise for years. Reflecting on 2017, House opined "(l)ast year, however, seems to represent a new tier of activity that brings a lot of momentum into 2018 . . . . It will be interesting to see how the recently passed GOP tax reform affects activity." This article examines the tax factors related to a sale of a business including the portions of the *Tax Cuts and Jobs Act of 2017* (TCJA 2017) that affect sales of business assets.

AN EXAMPLE FOR CONTEXT

Judith Pierce and her son, Paul Pierce, started a business that rents out heavy duty pick-ups to tradesmen who travel to Southern California. The business was an early entry into this growing market. In 2016, Big Trucks, LLC, purchased 24 Sierra 2500HD crew cabs from McClellan Buick-Pontiac-GMC for $800,000 to be used as rentals. The business operates through referrals from existing rental agencies at the airport and elsewhere. When a rental company has a renter for one of the units, it is rented by Big Trucks to the agency, which, in turn, rents it to the customers. This is a strong business model since it is unnecessary overhead for each rental agency to keep its own stock of large pickup trucks.

The Pierces have decided to go a different direction with a new business plan and have offered all the assets of Big Trucks for sale to a national equipment leasing company, LeBron Leasing Specialists, Inc. in 2019. The company balance sheet currently reflects assets having tax bases and estimated market values in Exhibit 1.

The tentative cash price for the assets is $850,000. Notice the value of the business exceeds the sum of the values of the identified assets by $165,000. This represents goodwill and other intangibles not identified in the balance sheet. The sellers are also paid $100,000 under a separate enforceable non-compete covenant.

Exhibit 1

Asset	Basis	Estimated Market Value
Equipment (Cost = $800,000)	$0	$475,000
Lease (the ground lease is transferable and the lease payments are below current market)	0	25,000
Leasehold improvements (building on the leased land, Cost = $90,000)	85,585	100,000
Solar panels (Cost $40,000, purchased 35 months before the sale)	0	35,000
Contracts	0	50,000
Total	$85,585	$685,000

STRUCTURE OF A SMALL BUSINESS SALE

Small business sales may be structured as a sale of the assets (including goodwill or going concern value), or as a sale of the stock or ownership interest. According a leading small business sale transaction database (DealStats), over 80% of small business sales are structured as a sale of assets. There are many pros and cons to each structure, but, from the buyer's standpoint, an asset sale is clearly advantageous. Not only does the buyer avoid liability exposure from the seller's operations, but the buyer gets to increase the tax bases of the acquired assets resulting in increased depreciation and amortization benefits.

ASSET VALUATION

In an asset sale, the identifiable assets of the business must be appraised to determine fair market value. U.S. Treasury Regulation §25.2512-1 provides the definition of fair market value for tax purposes: "The value of the property is the price at which such property would change hands between a willing buyer and a willing seller, neither being under any compulsion to buy or to sell, and both having reasonable knowledge of relevant facts." A discussion of items to be appraised and reasons for appraisals is available at firstbusiness.com. Internal Revenue Code § 1060 (IRC) provides guidance regarding allocation of purchase price to the acquired assets:

> **(a) General rule.** In the case of any applicable asset acquisition, for purposes of determining both—
> (1) the transferee's basis in such assets, and
> (2) the gain or loss of the transferor with respect to such acquisition, the consideration received for such assets shall be allocated among such assets acquired in such acquisition in the same manner as amounts are allocated to assets under section 338(b)(5). If in connection with an applicable asset acquisition, the transferee and transferor agree in writing as to the allocation of any consideration, or as to the fair market value of any of the assets, such agreement shall be binding

> on both the transferee and transferor unless the Secretary determines that such allocation (or fair market value) is not appropriate.

In most instances, the buyer and seller of a business must file Form 8594 (https://www.irs.gov/pub/irs-pdf/f8594.pdf), which includes details as to the values of various assets. The instructions to the form indicate that, in general, ". . . both the purchaser and seller must file Form 8594 and attach it to their income tax returns . . . when there is a transfer of a group of assets that make up a trade or business . . . and the purchaser's basis in such assets is determined wholly by the amount paid for the assets. This applies whether the group of assets constitutes a trade or business in the hands of the seller, the purchaser, or both" (https://www.irs.gov/pub/irs-pdf/i8594.pdf).

Valuations are based on market value, and the value assigned to an asset cannot exceed its fair market value. If a return is examined, the ". . . Internal Revenue Service may challenge the taxpayer's determination of the fair market value of any asset by any appropriate method and take into account all factors, including any lack of adverse tax interests between the parties" (https://www.law.cornell.edu/cfr/text/26/1.338-6).

The significance of the last part of this statement, and the reason why the IRS generally accepts the agreed to values, is that in almost every instance what is favorable for the buyer is unfavorable for the seller. For example, if more value is assigned to inventory and supplies, the seller has more ordinary income and the buyer enjoys a more immediate ordinary deduction.

BUSINESS VALUATION

Valuation of a business can be arrived at using any number of methods including industry related multipliers based on gross receipts, net cash flow analyses based on some definition of net income, and asset appraisals if there are no significant unidentified intangibles (First Business). One example of a multiplier deals with sales of CPA firms where the market value of the firm is typically around 1 to 1.35 times annual revenue (West). Since customer loyalty and sustaining goodwill are always an uncertainty, payouts may be based on sustained billing, i.e., a percentage based on actual future billings, or some uncertain future indicator (Sinkin).

Business brokers are available for owners who are not prepared to appraise the assets, the business as a whole, or who are not prepared to execute the transaction (Accounting Broker).

TAX CONSIDERATIONS: NEW AND OLD

The following paragraphs address the tax factors, particularly those affected by the Tax Cuts and Jobs Act of 2017 (TCJA 2017), involved in the sale and purchase. The treatment of various assets common to many business sales, including the sale of hypothetical Big Trucks, LLC, are addressed.

Tax Rates

The Tax Cuts and Jobs Act of 2017 (TCJA 2017) significantly altered tax rates and calculations. These individual rates, and most other non-corporate changes, are scheduled to be temporary for calendar years 2018–2025. The corporate tax rates are scheduled to be permanent.

Corporate Tax Rates

The corporate rate was reduced for years beginning in 2018 and beyond from a progressive structure, with rates from 15 to 35 percent, with even higher marginal rates over some ranges, to a flat 21 percent (IRC, § 11).

Individual and Estates and Trusts Rates

The rates for other entities, i.e., individuals, estates and trusts, were cut from a high of 39.6 to 37 percent. The brackets were altered somewhat, and the rates were lowered within most of the brackets (IRC § 1). Exhibit 2 summarizes the changes over approximate ranges:

Exhibit 2
Ordinary Tax Rates Before and After Tax Reform

Single Taxpayers	Before	After
Approximately $-0- to $9,525	10%	10%
Approximately $9,525 to $38,700	15%	12%
Approximately $38,700 - $82,500	25%	22%
Approximately $82,500 - $157,500 (beginning at $92,000 in 2017)	28%	24%
Approximately $157,500 - $200,000 (beginning at $192,000 in 2017)	33%	32%
Approximately $200,000 - $500,000 (beginning at $416,000 in 2017)	35%	35%
Over $500,000 (beginning at $418,000 in 2017)	39.6%	37%

Married Taxpayers	Before	After
Approximately $-0- to $19,050	10%	10%
Approximately $19,050 to $77,400	15%	12%
Approximately $77,400 - $165,000	25%	22%
Approximately $165,000 - $315,000 (beginning at $153,000 in 2017)	28%	24%
Approximately $315,000 - $400,000 (beginning at $233,000 in 2017)	33%	32%
Approximately $400,000 - $600,000 (beginning at $417,000 in 2017)	35%	35%
Over $600,000 (beginning at $471,000 in 2017)	39.6%	37%

The “ordinary” rates are presented here since not all sales of businesses result solely in capital gains and not all ordinary income gains are taxed at the highest rates. The brackets reflected are not precise but reflect the general positions. The shifts in the tax rates from 2017 to 2018 tend to represent an increased benefit to sellers of businesses where some of the gain is taxed at ordinary income rates.

Other structural changes for individual taxpayers that affected tax burdens were numerous. Large shifts in burden were caused by the repeal of the personal and dependency exemption deductions (still in existence but were set at $-0-), the doubling of the child tax credit, the approximately doubling of the standard deduction, and the cap on the itemized deduction for personal state and local taxes at $10,000. These structural items, other than the rate changes, affect tax burden, but they do not significantly affect the analyses in this study, so they will not be discussed further.

Capital Gains Rates for Individuals, Estates and Trusts

Capital gains rates remained essentially the same, with rates of 0 percent for lower income taxpayers, 20 percent for very high-income taxpayers, and 15 percent for most others. There are two additional capital gains rates. The maximum rate on dispositions of collectibles and § 1202 stock (IRC), which are not applicable in this study, remains at 28 percent. The maximum rate on “unrecaptured §1250 gain” (IRC), which could apply to sales of businesses that own real estate, is 25 percent. “Unrecaptured § 1250 gain,” is gain on the disposition of depreciable real property to the extent of depreciation allowed or allowable. Gain in excess of the 25 percent gain is subject to the 0 percent for lower income taxpayers, 20 percent for very high-income taxpayers, and 15 percent for most others.

Qualified Business Income Deduction

Congress created a large disparity between the tax rates for corporations and other taxpayers, creating a disadvantage for pass-through entities operating businesses, i.e. sole proprietorships, partnerships, and electing S corporations. Rather than causing a near mandatory shift for pass-through entities to switch to corporate status, Congress implemented what is called the qualified business income (QBI) deduction under new Code § 199A (IRC). This is an artificial deduction, not requiring any outlay of resources.

A detailed explanation of the QBI deduction is beyond the scope of this study, but since it could impact decision making, QBI is summarized here. This deduction cannot exceed 20 percent of QBI, or if smaller, 20 percent of taxable income before the QBI deduction. For taxpayers below established thresholds, i.e. taxable income before the QBI deduction of $321,400 for married filing jointly and $160,725 for others (2019 amounts), the deduction is simply 20 percent of QBI, or if smaller, 20 percent of taxable income before the QBI deduction.

Taxpayers whose taxable incomes exceed these thresholds are subject to a QBI deduction cutback over the next $100,000 for married and $50,000 for others. For example, married taxpayers with taxable income of $421,400 or more would get no QBI deduction because their deduction is fully cut back.

This cutback can obviously eliminate the benefit of the QBI deduction. Higher income taxpayers, other than taxpayers in service businesses, will enjoy a QBI deduction to the extent of a floor equal to 50 percent of W-2 wages or 25 percent of W-2 wages plus 2.5 percent of investments in depreciable assets, whichever is larger. If the married couple in the example above paid W-2 wages of $120,000, their QBI deduction would be $60,000, not to exceed 20 percent of QBI or 20 percent of taxable income before the QBI deduction.

QBI is generally just what it sounds like. It excludes investment type income, including capital gains and losses, and salaries and wages. For purposes of this study, a question arises as to what gains and losses upon the sale of a business for the seller and what deductions for the buyer, affect QBI. For example, gain on the sale of inventory and supplies should be QBI for the seller and reduce QBI for the buyer, assuming the buyer is not a regular corporation since corporations do not benefit from the QBI deduction. Capital gains are not QBI for the seller, but depreciation and amortization deductions for the buyer reduce QBI, again, assuming the buyer is not a regular corporation. These, and other determinations may change over time, as this deduction is new, and uncertainties still exist.

Inventory and Supplies

Any gain or loss from the sale of inventory and supplies is ordinary income or loss. A new twist was created by the Tax Cuts and Jobs Act of 2017(TCJA 2017), in that § 448 of the Internal Revenue Code (IRC, also see Revenue Procedure 2018-40) was amended to allow small businesses with gross receipts of $25 million or less to use the cash method of accounting and discontinue accounting for inventory as well as for the burdensome uniform capitalization procedure (IRC § 263A). Eligible businesses choosing to no longer account for inventory can simply treat purchases of goods similar to the treatment of supplies.

This really doesn't change the treatment of the sale of a business, but the gain attributable to the inventory to the seller will be greater since the basis will be zero. The buyer may be able to immediately expense the inventory if the buyer is eligible.

Equipment

Any gain attributable to equipment in a business sale will likely involve depreciation recapture under § 1245 (IRC). This recapture, of course, is ordinary income. On the buyer side, the taxpayer will be able to expense the entire cost since TCJA (TCJA 2017) allows 100 percent bonus depreciation for purchased property, including used property (IRC, § 168(k)). Prior to TCJA, used property was not subject to bonus depreciation.

Solar Panels

If the solar panels are sold along with the other assets of the business, part or all of the cost recovery deductions and the investment tax credit (ITC) will need to be recaptured. Specifically, the seller would have $35,000 of depreciation recapture, based on the $35,000 fair market value, and be required to pay back 60 percent of the ITC as recapture, resulting in a cash outflow of $7,200. ITC must be recaptured at a rate of 20% per year for each part or full year short of five years (i.e., three years in this case). The recapture of $7,200 is $40,000 x 30 percent credit x 60 percent.

The purchaser would not be entitled to a solar tax credit since the credit is only available to the original owner of the panels, so the panels could simply be expensed under bonus depreciation with no energy credit.

It is clear the panels could be removed from the structure and installed somewhere else. As a result, the seller should be able to sell the building and retain title to the panels. The seller could then retain title and sell energy to the buyer. After the credit is fully earned at the end of five years, a decision could be made as to whether to sell the panels, or not.

Real Estate Including Improvements on Ground Leases

Tax reform did not significantly affect the tax treatment for sale of real property.

Sale of Real Property

Depreciable property and land used in a trade or business that is held for more than one year is § 1231 property. Depreciable § 1231 property that is real property is not technically subject to depreciation recapture. But, as described previously, net gain is 25 percent capital gain to the extent of "unrecaptured § 1250 gain." Any remaining gain on improvements and land is 15 percent gain, or 0 percent for very low income taxpayers and 20 percent for very high income taxpayers.

Acquisition of Real Property

For a buyer, real property is classified as residential rental property or nonresidential real estate as before and depreciated using straight line and lives of 27.5 and 39 years, respectively (IRC, § 168). Real estate improvements constructed on rented land are depreciated the same as those on land that is owned by the taxpayer.

Leases and Contracts

If value is assigned to leases or similar contracts in which the owner has no basis, capital gain should be allowed. For example, if the seller entered into a favorable lease and the rent was lower than the prevailing market, the lease has value. The leases and contracts can be amortized over the life of the assets, however, they are not §197 assets (IRC).

Other Intangible Assets

Most other intangible assets will be §197 assets (IRC). Internally generated assets like goodwill generally qualify for capital gain treatment to the seller. Those that were purchased and amortized are §1231 assets (IRC) and any amortization deductions are subject to recapture under §1245 (IRC).

Section 197 property (IRC) can be amortized over a period of 15 years.

Covenants Not-to-Compete

A separate agreement not to compete adds another element to a business sale. The overall price is certainly affected by the potential future competition from the seller. One explanation, by Steven Thompson, follows:

> "When a business is sold, it is not uncommon for a portion of its sale price to be attributable to a noncompete agreement between the seller and the buyer. For tax purposes, a covenant not to compete is recognized when it is severable from goodwill, the agreement is separately bargained for, and the covenant can be shown to have economic substance. When this is the case, the portion of the sale price attributable to goodwill is generally treated as a capital asset, and the payment received for the noncompete agreement is taxable as ordinary income under Revenue Ruling 69-643, 1969-2 CB 10" (Thompson).

The buyer is entitled to an ordinary deduction as a business expense for periodic payments to the seller for the agreement not to compete.

EFFECT OF NEGOTIATIONS

For the seller, as has been indicated, sales of some assets result in capital gain or loss; others, ordinary income or loss. For the buyer, purchases of some assets result in immediate expensing; others, deferred deductions, or in the case of land, no deduction at all until the land is sold. Since § 1060 (IRC), and other authorities, respect allocations of value between the various assets, the buyer and seller can negotiate, not only the price for the business, but the reasonable fair value of the underlying assets.

Generally, in such negotiations, what is good for the buyer is not good for the seller, and vice versa. For example, with inventory, the seller has ordinary income (bad, but it might be QBI, which is good) and the seller has an ordinary deduction (good, but this might reduce QBI, which is bad). So, in negotiations, the buyer would favor a relatively high value for inventory, while the seller would prefer a lower value.

Another example is goodwill. For internally generated goodwill, the seller has capital gain and the buyer has a § 197 (IRC) asset that can be amortized over 15 years. In this case, the seller would prefer a higher value for goodwill, while the buyer would prefer a lower value since the buyer typically seeks current deductions.

As indicated, if it would be reasonable based on facts and circumstances, the buyer would negotiate for a higher value for inventory and a lower value for the goodwill. The seller has the opposite inclinations. These negotiations could, due to the give and take and varying tax rates between the buyer and seller (perhaps 21 percent for a corporation and 37 percent for an individual), result in renegotiating the overall price upward or downward. This is referred to as “tax arbitrage” (Investopedia).

It is important to realize that tax arbitrage could affect these negotiations. A corporate buyer might be willing to give up some ordinary deductions, rather than deferred deductions, for a slight increase in the overall price. A non-corporate seller might accommodate the change to get more capital gain.

STOCK OR OWNERSHIP INTEREST SALES

Our discussion so far has dealt with asset sales. If the business is operated as a partnership, including an LLC reporting as a partnership, limited liability company, or a corporation, including an S corporation, there is the option to sell the partnership interest, member interest, or stock. An elaborate discussion of the issues here will be left for another day. However, it is unlikely that the buyer of a partnership or an S corporation would gain from acquiring an ownership interest or stock over purchasing the assets.

<u>Sale of Corporate Stock</u>

The sale of stock in a regular corporation seems appealing, but an informed buyer would not pay as much for the stock of a corporation with appreciated assets, as they would for the appreciated assets because they would keep the low basis in the underlying assets resulting in less deductions. This is compared to the fair market value, as negotiated, basis in purchased assets that result in larger deductions.

Sale of Partnership Interest or S Corporation Stock

The sale of an interest in a partnership or S corporation might seem tempting, but in an asset sale, the gain or loss flows through to the owners and retains its character as capital gain or ordinary income. Since the owner's basis in the entity goes up by any gains and down by any losses, and since the buyer gets a basis equal to cost, there is no benefit to acquiring a partnership interest or stock in a corporation.

There are complex rules for the sale of a partnership interest that further complicate things. Even though a partnership interest is generally treated as a capital asset (IRC, § 741), ordinary income that would have been recognized had the underlying ordinary assets been sold will be recharacterized as ordinary income (IRC § 751(a)).

CONCLUSION

It is imperative that business buyers and sellers, and their advisers, be well versed in the advantages and disadvantages in structuring the transfer of a business. Understanding the effects of various elements allows a party involved in the transaction to seek advantages and avoid the risks of overlooking those advantages. A taxpayer can suffer economic consequences if the other party in the transaction is better informed and prepared. These taxpayers must involve their professional tax advisers and may benefit from engaging business brokers.

REFERENCES

Accounting Broker:
https://accountingbroker.com/areyouselling/?gclid=CjwKCAjwwtTmBRBqEiwA-b6c_7ZJuQyPqapSHqTnvv8792CSWVbtJwMMcd7Ehv-82vwQFq5CX1STDhoCbQ8QAvD_BwE

"DealStats (Formerly Pratt's Stats): Business Valuation Resources." *DealStats (Formerly Pratt's Stats). Business Valuation Resources*, 2019, www.bvresources.com/products/dealstats.

First Business: https://firstbusiness.com/resource-center/determining-the-value-of-your-business/

Internal Revenue Code (IRC) of 1986, Title 26, United States Code.

Investopedia: Tax Arbitrage, https://www.investopedia.com/terms/t/tax-arbitrage.asp.

Nitti, Tony, CPA. "Understanding the new Sec. 199A business income deduction," Tax Adviser, April 1, 2018. ttps://www.thetaxadviser.com/issues/2018/apr/understanding-sec-199A-business-income-deduction.html.

Sinkin, Joel and Terrence Putney, CPA**.** **"**How to value a CPA firm for sale**,"** Journal of Accountancy**,** October 31, 2013**.**
https://www.journalofaccountancy.com/issues/2013/nov/20138232.html

Small Biz: https://smallbiztrends.com/2018/01/bizbuysell-2017-insight-report.html

TCJA 2017, Tax Cuts and Jobs Act of 2017, Public Law No: 115-97 (12/22/2017).

Thompson, Steven C**.** "Noncompete Agreement Payment Wasn't Sale of Personal Goodwill," April 30, 2009.
https://www.journalofaccountancy.com/issues/2009/may/noncompeteagreement.html

West, T. 2019. *2019 Business Reference Guide, The Essential Guide to Pricing a Business*. 29th edition. Business Brokerage Press.

DECISION SCIENCES INSTITUTE

An Undergraduate Business Statistics Course

Micheline Al Harrack
Marymount University
2807 North Glebe Road, Arlington, VA 22207
malharra@marymount.edu
703-522-5600

Charles B. Tichenor
Marymount University
2807 North Glebe Road, Arlington, VA 22207
charley_tichenor@hotmail.com
703-901-3033

ABSTRACT

Marymount University is upgrading its analytics program for undergraduate business students, as the US workplace demands workers with stronger business analytics skills. Part of this is the implementation of an Advanced Business Statistics course. This course is based on several pedagogical “axioms,” uses real world cases, and uses a new textbook written by the professors. This permits updates of material in real time, provides fresh-graded material each semester, and is downloaded by the students without charge. This paper describes the “axioms,” the corresponding new textbook structure, the student literature review exercise, and the results achieved after the first semester’s presentation.

KEYWORDS: statistics, education, management science, undergraduate

INTRODUCTION

Marymount University is a comprehensive Catholic university offering a wide array of undergraduate and graduate degrees through the Schools of Design, Arts, and Humanities, Business and Technology, Sciences, Mathematics, and Education, and the Malek School of Health Professions. It is located in the Washington, DC metropolitan area in Arlington, VA, with a total enrollment of about 3400 graduate and undergraduate students.

The School of Business and Technology provides its Advanced Business Statistics courses in modern computer labs with maximum class sizes of 24 students per course section. Typical class sizes are about 20-22. Professors teach all of these courses.

The university upgraded its undergraduate business analytics (also called “operations research” or “O.R.”, or “Management Science”) program by the introduction of an Advanced Business Statistics course. This course is designed to help business students (BBA) become more competitive in the analytics marketplace upon graduation. According to the trade association “Institute for Operations Research and the Management Sciences” (INFORMS),

> O.R. and Analytics drive performance and change in organizations of all types – large and small, private and public, for-profit and not-for-profit.

> O.R. and Analytics are used in incredible ways, to inform high-level strategy, enhance day-to-day operations, design better public policies, and more.
>
> Using techniques such as mathematical modeling to analyze complex situations, O.R. and Analytics enable more effective decisions and more productive systems based on robust data, the fuller consideration of available options, and careful predictions of outcomes and estimates of risk. (INFORMS, 2018)

The following analytics-based jobs are considered among the best business jobs for 2017 by US News and World Report, October 6, 2017, Statistician (#1) and Mathematician (#3), and Operations Research Analyst (#5) (US News and World Report, 2017)

Our students previously took a basic statistics course in the business school from a Management Science professor. Now, these students take their first statistics course through the Mathematics department, and then take this Advanced Business Statistics course from Management Science professors. All business undergraduate students, therefore, take both the basic and advanced courses.

LITERATURE REVIEW

There are different approaches to setting goals for statistics courses. In the paper “Goals of Your Introductory Statistics Course,” Donald B. Macnaughton discusses “topic-based” course approaches, with goals of covering a set of topics such as probability, point and interval estimation, etc. He goes on further to state “Unfortunately, topic-based goals have a serious drawback: by emphasizing statistical topics, the goals usually *fail* to emphasize what is essential, which I believe is to help students appreciate the vital role of the field of statistics in empirical research.” (Macnaughton, 1997)

Jeffrey A Witmer, Oberln College, wrote the paper “A Second Course in Statistics.” In this paper he discusses his experience reviewing requests to transfer Statistics I courses to Oberlin from other institutions; he indicates that there is wide commonality in the US for this basic course. However, his experience is that there is not such commonality for Statistics II courses. He writes, “On some campuses Stat2 was, and still is, a regression course. For others it is an ANOVA course. Some statisticians prefer to follow Stat1 with a course on nonparametric methods, while others teach design of experiments.” “Our course is focused on how statisticians function as they work with clients and analyze data. … The Stat2 course that we have developed takes modeling as its theme.” (Witmer, n. d.)

Another approach is from Thomas E. Love. In his paper from the “Journal of Statistics Education,” he writes “In a single semester, students explore data using tools from EDA, multiple regression, analysis of variance, time series analysis, and categorical data analysis. The focus is on understanding and forecasting in a variety of data settings, learning how to summarize relationships and measure how well these relationships fit data, and how to make meaningful statistical inferences when the usual assumptions do not hold. The course emphasizes what the statistical process is all about: how to conduct studies, what the results mean, and what can be inferred about the whole from pieces of evidence.” (Love, 1998)

Our Advanced Business Statistics course goals were designed to show students how to use statistics in a business context to improve an organization’s return on investment, how to conduct business research, and how to evaluate business research done by others. It is designed to provide students with an environment very similar to that experienced by an entry-level analytics professional; the professor is both an "instructor" and a "supervisor." About 10% of Marymount students taking the analytics-based courses express strong interest in analytics as an eventual career field. This number is expected to grow as the workforce continues to grow for analytics. This course will give them first-hand experience in a professionally safe environment. The other students will get experience with what those analytics people will do, what to expect from them, and how to appraise analytics products at work. In particular:

- Students solve case-like exercises, not "homework problems," and explain their results in plain English. Their solutions must be sold to a "supervisor" or other "stakeholder/ customer" -- those roles played by the professor. Students are taught that an unsold analytics study is practically worthless – even if they achieved the correct answer. This is the primary component of the course grading.
- This focus includes using technology – tools that they will find in the workplace – primarily Microsoft Excel.
- The course also requires students to evaluate scholarly business research. Analytics professionals need to conduct literature reviews as part of their normal work and base their work on the findings of others.
- There is also a quiz, two midterms, and a final exam to access students' comprehension of the text material.

The course is basically conducted in three generic modules.

- Module 1: Excel Enrichment
    - One class period (75 minutes) of Excel graphics.
    - One class period of Advanced Excel (sorting, filtering, anchoring, pivot tables, logic formulas).
- Module 2: Advanced analytics.
- Module 3: Scholarly literature review.

The course was developed (and is currently being further developed) based on several "axioms." These are opinions of the professors from which the course is based. They certainly are arguable but they seem to work well for us – at least as of now. As this paper proceeds, those "axioms" will be introduced by italic font.

Axiom. Many collegiate advanced business statistics courses can be improved by formulating them as real world laboratories, where students take the roles of entry-level analytics professionals and are given a series of real-world case studies to solve which are from the experience of the professor. The professor is both their "instructor" and "supervisor." Because each case is from professor experience, the professor specifically knows why the case's business solution was chosen. This contrasts with the pre-published cases approach used by professors who may have had no personal involvement in the corresponding business solution. Students gain first-hand experience of what that profession is like, rather than reading textbooks and solving exercises in the back of the chapters.

Axiom. Business statistics textbooks are far too expensive. Costs can easily run about $250 for a traditional hardcover edition purchase and over $100 for softcover editions. Although many students will not keep the text as they are taking a “required” course and will choose to instead rent the book, about 10% of the students express sincere interest in analytics as a career field and may want to keep the text.

We wrote our own book. Students download the .pdf version from the university network.

- This presents a $0 cost to the students for the text, regardless of whether they decide to keep the text or not.
- This allows Marymount professors to update the gradable materials each semester with new materials to minimize cheating.
- This allows Marymount professors to change the text material as needed between semesters (or in real time if necessary).
- This is one less book for the student to carry if they select to review online but can be printed if student prefers.
- The book is also used in other analytics/management science courses at no charge to those students.

Axiom. Many of the current field of undergraduate business statistics books are far too lengthy. Most of the pages will go either unassigned or unread in a typical semester-length course. This is an undue burden to the student.

Today, many statistics books are full-sized and can be well over 850 pages. We place ourselves into the position of the general student who must buy a textbook of formulas, graphs, tables, new symbols, and new vocabulary. This can be quite daunting. It almost initiates the question “What will we ever use this for?” Also, a large number of those pages will not be in the curriculum and are therefore unneeded for the course. Our book has 292 pages, all used in the course.

Axiom. Some statistics courses are either taught or learned from the viewpoint of the “formalist” mathematical philosophical approach. This course is based on the “structuralist” approach.

In our experience, many statistics courses are taught from the approach of teaching formulas (“Here’s the formula.”) Many weaker students learn the formulas and manipulate them to achieve the solution. Examples of these are the numerous probability formulas and the z normal distribution formula. This student approach is a form of the philosophy of “formalism.” In this viewpoint, “The various philosophers that go by the name ‘formalism’ pursue a claim that the *essence* of mathematics is the manipulation of characters. A list of the characters and the allowed rules all but exhausts what there is to say about a given branch of mathematics.” (Shapiro, 2000)

Our book uses the philosophy from Stewart Shapiro that “mathematics is the study of patterns” (structuralism approach) (Shapiro, 2000).

“Mathematical structuralism is the view that, in some sense, mathematics is about structures and their relations, rather than about objects and their properties.” (Landry, 2016). For example, in this approach, the number “4010” refers to the room number location of a computer lab in the

Business School building. The number “5” refers to a location on the number line. The probability of rolling a die and getting a total of “4” is obtained by noticing the pattern that there are six possible outcomes, and one of those is a “4,” so a 6 is placed in the denominator and a 1 in the numerator: the pattern is “1/6.” Probability patterns showing all basic probability rules can be shown by using dice and playing a popular board game such as Monopoly®, as opposed to mathematical derivations.

By teaching the underlying patterns, students who learn them do not need to memorize or learn formulas. They merely learn how to quantify the patterns; the entire probability chapter can be taught without naming or algebraically deriving particular formulas per se.

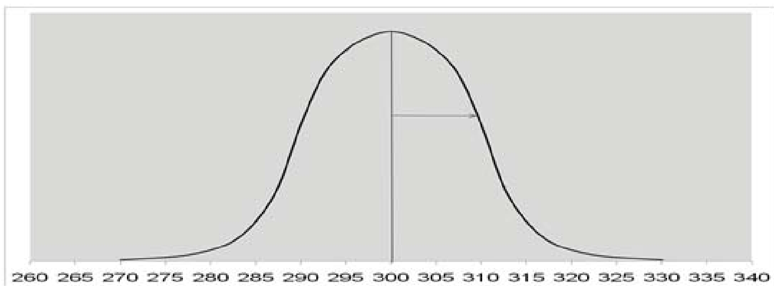


Figure 4-15

Figure 1

Figure 1 is from our book (book Figure 4-15) representing a hypothetical population of events with values normally distributed with a mean of 300, and a standard deviation of 10.

I was taught the relationships between z, x, μ, and σ by memorizing the following formula.

$$z = \frac{x - \mu}{\sigma} \qquad (1)$$

For example, Figure 2 is an extract from a presentation from a University of Wisconsin Statistics course (Lindstrom, 2004) which shows this approach. This approach consists of, basically, three steps: formulate the problem, convert the problem into a standard normal curve problem, and then solve the problem.

Standardization

If Y is a normally distributed Random Variable $Y \sim \mathcal{N}(\mu, \sigma^2)$

And we define

$$Z = \frac{Y - \mu}{\sigma}$$

Then Z is a standard normal random variable

$$Z \sim \mathcal{N}(0, 1)$$

Every problem that asks for an area under a normal curve is solved by first finding an equivalent problem for the standard normal curve.

Figure 2

For me, it was difficult to intuitively understand why normalizing the distance between x and μ by dividing by σ produced the value of z; or, why is it necessary to add a logical step to first convert the problem to a standard normal curve problem, and then solve. Instead, formulate the problem using this approach (i.e., one is allowed to multiply each side of Figure 2's formula by σ).

$$z\,\sigma = x - \mu \tag{2}$$

Then solve, eliminating the step of the standardized normal curve. Teach that the distance between x and μ is simply the corresponding number of σs. For example, if σ = 10, then distance between 310 and 300 is one σ (z = 1). The distance between 320 and 300 is two σs (z = 2). This pattern can now be taught and understood intuitively by most students in a few minutes without having to memorize the formula. Since it works for any shape normal curve, the additional step's discussion about the standard normal curve can be avoided.

Axiom. In the real world, analysts present the solutions to their work to their superiors, and often to groups of managers and analysts in business meetings. Critical thinkers in those meetings will appraise the work in terms of its accuracy, and also in terms of its usefulness. For the analyst who did the work, getting the right answer to a statistics problem is important, but it is no more than 49% of the requirement. One must sell the answer to the decision maker, and the sales part is worth at least 51% of the final product. Basically, an unsold correct answer is almost worthless.

Many statistics books and courses end an exercise upon "getting the correct answer." In the real world, there is an additional step: answers reached by analysts must be "sold" to decision makers before the decision makers will approve any subsequent action. Answers to the cases must be mathematically sound; but a manager, senior to the analyst, who may not be qualified in statistics, must also understand the process and solution. Case study answers therefore must be statistically/mathematically sound, and must be presented to the instructor in plain English who simulates a decision maker who may not be qualified in statistics.

THEORETICAL DEVELOPMENT/MODEL

Textbook Structure

Figure 3 shows part of the Table of Contents. For most chapters, there is a “chapter number,” and a “chapter number R.” The “R” denotes “review.” For example, Chapter 4 is the Normal Distribution case study with new material, and Chapter 4R is the basic normal distribution text material as a review from a typical introductory statistics course. This review chapter is optional for those who need it, and includes solved end of chapter problems; this provides guidance to students who cannot recall information from the introductory course.

Table of Contents

Figure 3

Axiom. In the real world, analysts present the solutions to their work to their superiors, and often to groups of managers and analysts in business meetings. Critical thinkers in those meetings will appraise the work in terms of its accuracy, and also in terms of its usefulness. For the analyst who did the work, getting the right answer to a statistics problem is important, but it is no more than 49% of the requirement. One must sell the answer to the decision maker, and the sales part is worth at least 51% of the final product. Basically, an unsold correct answer is almost worthless.

Many statistics books and courses end an exercise upon “getting the correct answer.” In the real world, there is an additional step: answers reached by analysts must be “sold” to decision makers before the decision makers will approve any subsequent action. Answers to the cases must be mathematically sound; but a manager, senior to the analyst, who may not be qualified in statistics, must also understand the process and solution. Case study answers therefore must be statistically/mathematically sound, and must be presented to the instructor in plain English who simulates a decision maker who may not be qualified in statistics.

Student Literature Review Assignment

Many times entry-level (and higher level) analysts will be tasked to somehow develop an analytic procedure that has never been done at their organization. We believe that students should be taught to perform a literature search early in their effort to be exposed to the lessons learned when other tried similar work (if they did). This part of the course simulates a serious literature review. This module of the course requires students to perform a scholarly review of an academic journal paper in the business field that contains a statistical analysis. During our first semester, some students indicated that this was the first time they performed such a serious literature review on a business paper. Figure 4 is an extract from the syllabus describing the literature review exercise.

> Choose any business-related paper from the recent (less than 5 years old) general business journal literature (accounting, sales, marketing, operations research/analytics, business medical services, business software development, IT, etc.) Suitable journals include Harvard Business Review, CrossTalk, Interfaces, University of Chicago Journal of Business, etc. Not acceptable publications are popular magazine articles, trade association magazine articles, or newspaper articles. Meet with the instructor if there are any questions about the suitability of the paper and source.

Figure 4

Figure 5 is the grading rubric. This is also from the syllabus.

_____ 1. Was the selected journal paper business-oriented. 5 pts.

_____ 2. Is there a copy of the selected journal paper attached to the back of the critique. Did it include a page showing the Cover and Table of Contents (or masthead "screenshot" if from an electronic format) from the journal chosen. 4 pts. (Zero points if not from an academic journal.)

______ 3. Was the journal paper a recap of the author's research that made a new contribution to the body of knowledge, as opposed to a "report" which accumulated already existing knowledge. The body of knowledge contribution may be either substantial or small. 8 pts.

_____ 4. Did the student's critique contain a statement and explanation of the author's null and alternative hypotheses in plain English and in the student's own words. 8 pts.

_____ 5. Was the student's critique of the author's research methods free from statistical jargon and symbols, and written in plain English easily understandable by a middle manager who does not have an analytics background. 5 pts.

_______6. Was the student's critique of the author's conclusions free from statistical jargon and symbols, and written in plain English easily understandable by a middle manager who does not have an analytics background. 5 pts.

_______7. Was the student's critique of the author's conclusions logical and based on data or other facts, and not based on opinions such as often found in editorials, social media postings, popular discussions, etc. Were the sources of the data/other facts references footnoted. 5 pts.

Figure 5

This rubric, including the weighting factors, reflects our professional experiences both performing such literature reviews and supervising junior analysts in their literature reviews.

RESULTS

Figure 6 shows the final grade distribution for two sections of the first 22 students taking the course. Minus and plus grades are included with their base grades (e.g., all B+, B, and B- grades are shown here as a "B").

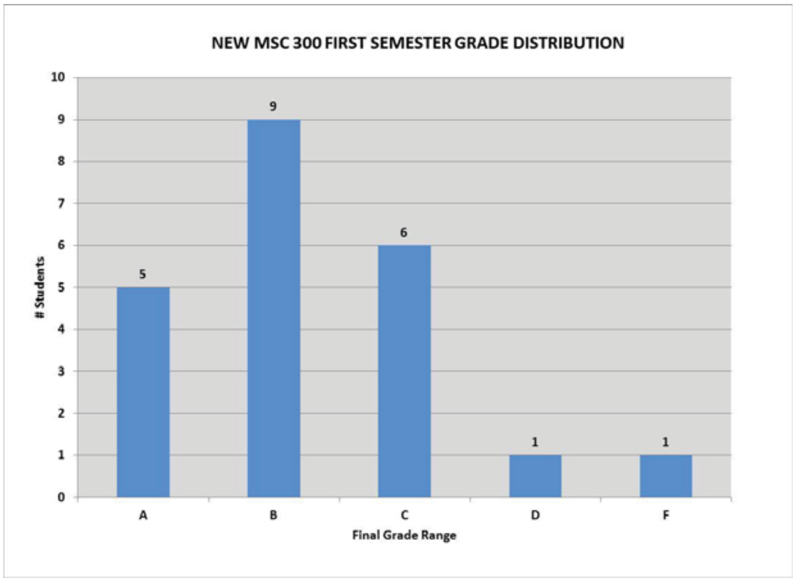


Figure 6

This shows the center of mass in the “B” range, and the overall distribution was consistent for both sections of the course. The course is designed so that most “A” and “B” range students should be fully academically qualified for an entry-level statistical analytics position, based on the professional experience of the instructors.

Figure 7 shows the student feedback scores for the first two sections of 22 students.

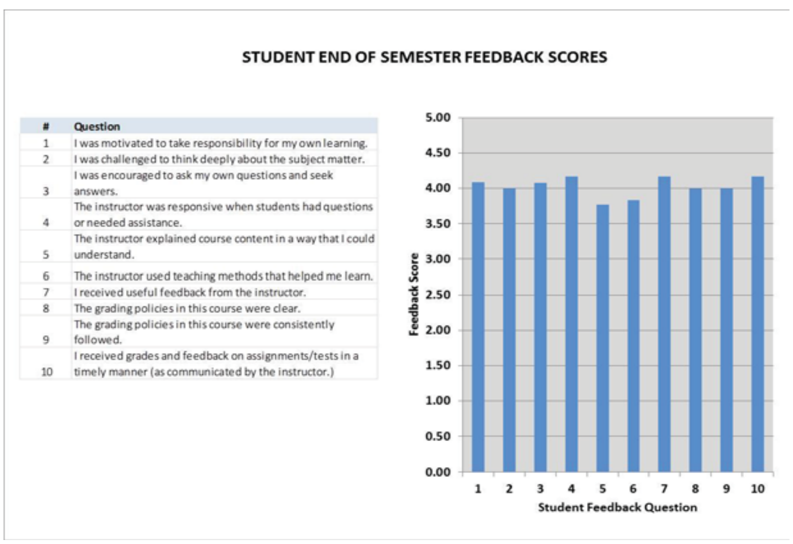


Figure 7

The center of mass on the student comment sheets was about a “4” out of “5” possible. These ratings were also consistent between the two sections. These scores were statistical averages; in the future, it might also be good to report medians to remove effects of possible insignificant outliers.

Finally, the 22 students avoided the list or rental prices of our previous textbook, representing a significant cost avoidance for them.

DISCUSSION AND CONCLUSIONS

One area for future development is to continue to solicit input from students for course improvements. This will include end of course student comment surveys, but will also include meetings with individual students for their confidential feedback after the course is over and their grades are turned in.

We will monitor how much time is needed for students to complete the case studies. Since the cases will be different, but similar, each semester, we will want to modify the time allotted for them so that the workload is cross-leveled more evenly in future semesters.

This course will be probably offered in two or three sections each semester. These professors will collaborate throughout the semester, and after the semester, to seek course improvement.

Another area for course improvement is to meet with professional analytics job recruiters. These meetings will result in a comparison of what they are seeking for their clients and what we are presenting in the course. This will help us keep the course calibrated for the job market.

Our results, both formally obtained through student feedback forms and by student interaction, indicate that we should continue with this approach as it reflects the real world as much as we can in the classroom and sparks many students to consider analytics as a career field.

REFERENCES

INFORMS. Retrieved from landing page of https://www.informs.org/Explore/Operations-Research-Analytics, August 24, 2018

Landry, Elaine (May 26, 2016). *Mathematical Structuralism*. Oxford Bibliographies, Retrieved from http://www.oxfordbibliographies.com/view/document/obo-9780195396577/obo-9780195396577-0305.xml, July 31, 2018.

Lindstrom, Mary (February 11, 2004). *The Normal Distribution.* Retrieved from http://www.stat.wisc.edu/courses/st371-lindstro/notes.mirror/normal.pdf July 6, 2018.

Love, Thomas E. (1998). *A Project-Driven Second Course*. Journal of Statistics Education, 6:1, DOI: 10.1080/10691898.1998.11910605.

Macnaughton, Donald (1997). *Goals of Your Introductory Statistics Course*. Retrieved from https://www.matstat.com/teach/p0002.htm, July 31, 2018.

Shapiro, Stewart (2000). *Thinking About Mathematics: The Philosophy of Mathematics*, page 140, Oxford University Press, Inc. New York.

Shapiro, Stewart (2000). *Thinking About Mathematics: The Philosophy of Mathematics*, page 250, Oxford University Press, Inc. New York.

US News and World Report (October 6, 2017). Retrieved from https://money.usnews.com/careers/best-jobs/rankings/best-business-jobs.

Witmer, Jeffrey A., (n.d.). *A Second Course in Statistics*. Retrieved from https://www.statistics.gov.hk/wsc/IPS032-P4-S.pdf, July 31, 2018.

DECISION SCIENCES INSTITUTE
Analyzing and Predicting Student Retention and Graduation Rates

Mohan Rao
Texas A&M University-Corpus Christi
Email: mohan.rao@tamucc.edu

Ashraf El-Houbi
Lamar University
Email: aelhoubi@lamar.edu

ABSTRACT

Low college retention is a universal problem. This study lists the factors and causes identified by prior research on this issue and other non-academic factors that include family college history, family support, student self-knowledge, student preparation, a school's educational and health provisions, and faculty's role in students' development, among others. Based on the data collected from a university over a six-year period, this study attempts to predict retention as well as graduation rates and the categories of students at risk. This study uses statistical techniques of logistic regression and the binary choice decision model.

KEYWORDS: Data analytics, enrollment management, logistic regression, student retention rate, student graduation rate

INTRODUCTION

Low college retention is a national and international problem. The problem of low retention may suggest a declining trend at several colleges. Several studies have identified the factors and causes of this problem. This study will identify academic and non-academic factors that influence the success of college students. There should be an early identification and intervention strategy for at-risk first year undergraduate students. This study looks at various factors aiming to predict what affects retention rate and which categories of students are at risk. It uses 2008 – 2014 cumulative data from a university in southern United States. It includes the following variables: age, gender, ethnicity, start time, full/part-time status, admission status, classification upon entering, SAT and ACT scores, high school, high school rank, if a student considered to be a high achiever, first generation (first person to attend university in the family), college, major, athletic affiliation, and cumulative GPA. Several statistical techniques such as logistic regression, the binary choice decision model, and discriminate analysis will be used to study the relationship between the response variable and several independent variables.

Logistic regression is a very powerful statistical technique that is used to model the relationship between a categorical outcome, the dependent variable, and a set of independent variables, such as education level, age, gender (Hosmer and Lemeshow, 2000). It models the probability associated with each level of the independent variable by presenting a linear relationship between predictor variables and a link function of these probabilities. On the other hand, linear regression predicts the actual values of the dependent variable. Different link functions in the linear equation show different levels of goodness-of-fit for the given data. The logistic regression technique will be used to find a best fitting and most reasonable model that can justify the relationship between response and independent variables and other explanatory variables. Since the response variable was ordinal, the authors used ordinal logistic regression to analyze and ascertain if there were significant relationships between the response variable and the predictor variables.

The rest of the paper is organized as follows: Literature review, discussion of analysis and results, summary and conclusions.

LITERATURE REVIEW

The literature review presented here is very limited due to space limitations.

DISCUSSION OF ANALYSIS AND RESULTS

The results discussed in this section are based on the 2008-2014 data collected from a university in southern United States. Logistic regression is a very powerful method and used to analyze and determine if there are significant relationships between the response variables and set of independent variables such as gender, ethnicity, high school rank, financial aid, GPA, SAT, and age In contrast with linear regression that predicts the actual values of the response variables, logistic regression models the probability associated with each level of the response variable by finding a linear relationship between predictor variables and a link function of these probabilities. Different link functions offer different levels of “goodness of fit” for the data. The aim of the logistic regression model is to find the best fitting and most sound model to describe the relationship between the response variable and set of the explanatory variables. The multiple logistic regression model can be written as:

$$\ln(\text{odds ratio}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i \qquad (1)$$

In equation (1) *k* represents number of independent variables in the model, and ε_i represents the random error in observation *i*. A logistic regression is based on the odds ratio, which represents the probability of a success compared with the probability of failure. The maximum likelihood technique is used to obtain the estimates of the parameters in the logistic regression equation (1). The deviance can be used as a statistic to assess the model's goodness of fit. The model is generally considered reasonable if the deviance is not significantly large.

Table 1: Logistic Regression Results (2008)

*** NOTE * 915 cases were used**

*** NOTE * 35 cases contained missing values**

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (1/2)							
Constant	-0.405984	1.01205	-0.40	0.688			
Gender	0.887462	0.349091	2.54	0.011	2.43	1.23	4.81
Ethnicity	-0.322429	0.236495	-1.36	0.173	0.72	0.46	1.15
High school rank	-0.388284	0.311660	-1.25	0.213	0.68	0.37	1.25
Financial Aid	0.242677	0.236461	1.03	0.305	1.27	0.80	2.03
SAT	-0.361496	0.230118	-1.57	0.116	0.70	0.44	1.09
Logit 2: (0/2)							
Constant	2.93549	0.427806	6.86	0.000			
Gender	0.329301	0.150211	2.19	0.028	1.39	1.04	1.87
Ethnicity	-0.0486634	0.0833814	-0.58	0.559	0.95	0.81	1.12
High school rank	-0.540953	0.132032	-4.10	0.000	0.58	0.45	0.75
Financial Aid	-0.138814	0.101940	-1.36	0.173	0.87	0.71	1.06
SAT	-0.364825	0.0948634	-3.85	0.000	0.69	0.58	0.84

Log-Likelihood = -728.229

Test of All Slopes Equal to Zero

DF	G	P-Value
10	56.124	0.000

Table 2: Logistic Regression Results (2009)

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (1/2)							
Constant	2.60351	4.11797	0.63	0.527			
Gender	0.533793	0.238341	2.24	0.025	1.71	1.07	2.72
High school rank	-0.169143	0.219580	-0.77	0.441	0.84	0.55	1.30
SAT	-0.374823	0.166542	-2.25	0.024	0.69	0.50	0.95
Financial Aid	0.0443422	0.161879	0.27	0.784	1.05	0.76	1.44
Ethnicity	-0.130747	0.149548	-0.87	0.382	0.88	0.65	1.18
Age	-0.123474	0.223902	-0.55	0.581	0.88	0.57	1.37
Logit 2: (0/2)							
Constant	7.00502	2.56105	2.74	0.006			
Gender	0.257387	0.152835	1.68	0.092	1.29	0.96	1.75
High school rank	-0.614193	0.139653	-4.40	0.000	0.54	0.41	0.71
SAT	-0.419710	0.105402	-3.98	0.000	0.66	0.53	0.81
Financial Aid	-0.0801022	0.102492	-0.78	0.434	0.92	0.76	1.13
Ethnicity	-0.0516719	0.0885980	-0.58	0.560	0.95	0.80	1.13
Age	-0.199771	0.138791	-1.44	0.150	0.82	0.62	1.07

Log-Likelihood = -873.743

Test of All Slopes Equal to Zero

DF	G	P-Value
12	58.341	0.000

Table 3: Logistic Regression Results (2010)

* NOTE * 60 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (1/2)							
Constant	4.42911	0.643685	6.88	0.000			
Gender	0.480821	0.221126	2.17	0.030	1.62	1.05	2.49
High school rank	-0.807288	0.199151	-4.05	0.000	0.45	0.30	0.66
SAT	-0.557582	0.140216	-3.98	0.000	0.57	0.44	0.75
Financial Aid	-0.387749	0.143113	-2.71	0.007	0.68	0.51	0.90
Ethnicity	-0.121382	0.101047	-1.20	0.230	0.89	0.73	1.08
Logit 2: (0/2)							
Constant	6.04141	0.584712	10.33	0.000			
Gender	0.624895	0.197548	3.16	0.002	1.87	1.27	2.75
High school rank	-1.06597	0.177624	-6.00	0.000	0.34	0.24	0.49
SAT	-0.623162	0.125014	-4.98	0.000	0.54	0.42	0.69
Financial Aid	-0.415262	0.126682	-3.28	0.001	0.66	0.52	0.85
Ethnicity	-0.131574	0.0894560	-1.47	0.141	0.88	0.74	1.04

Log-Likelihood = -907.392

Test of All Slopes Equal to Zero

DF	G	P-Value
10	112.340	0.000

Based on the logistic regression results, we notice that there is a significant relationship between student retention and the following independent variables: high school rank, SAT, Gender, and financial aid.

P-value =0.000 for the G-statistic which indicates that there is sufficient evidence to conclude that at least one of the estimated coefficients for the independent variable is different from zero, and this helps the authors to conclude that the model with all variables is significant at α = 0.05.

Table 4: Logistic Regression Results (2011)

* NOTE * 1047 cases were used

* NOTE * 9 cases contained missing values

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (1/2)							
Constant	4.82208	1.48261	3.25	0.001			
Gender	1.15102	0.581842	1.98	0.048	3.16	1.01	9.89
Ethnicity	0.254735	0.291717	0.87	0.383	1.29	0.73	2.29
High school rank	-0.0915281	0.468801	-0.20	0.845	0.91	0.36	2.29
Financial Aid	-0.496598	0.317442	-1.56	0.118	0.61	0.33	1.13
SAT	-0.458729	0.326937	-1.40	0.161	0.63	0.33	1.20
Logit 2: (0/2)							
Constant	7.61519	1.48045	5.14	0.000			
Gender	1.11959	0.581208	1.93	0.054	3.06	0.98	9.57
Ethnicity	0.177682	0.291741	0.61	0.542	1.19	0.67	2.12
High school rank	-0.841892	0.467678	-1.80	0.072	0.43	0.17	1.08
Financial Aid	-0.604496	0.316829	-1.91	0.056	0.55	0.29	1.02
SAT	-0.636071	0.327039	-1.94	0.052	0.53	0.28	1.00

Log-Likelihood = -768.934

Test of All Slopes Equal to Zero

DF	G	P-Value
10	62.676	0.000

Based on the logistic regression results, we conclude that there is a significant relationship between student retention and the following independent variables: high school rank, SAT, Gender, and financial aid.

SUMMARY AND CONCLUSIONS

Low college retention is a national and international problem. This paper listed the factors and causes identified by several authors for this problem. It also identified academic and non-academic factors that influence the success of college students.

Based on the logistic regression results, we notice that there is a significant relationship between student retention and the following independent variables: high school rank, SAT, Gender, and financial aid.

The limitations of this study include the results are based on a snapshot of the data from one university and so one should be cautious in generalizations. Further studies should include data from more periods of data, on more variables and from different universities. Further studies should also include data on intervention measures and their impact on retention and graduation rates.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Analysis of Student Retention Studies – A Text Analytical Approach

Xiaoni Zhang
Northern Kentucky University
Email: zhang@nku.edu

Toru Sakaguchi
Northern Kentucky University
Email: sakaguch@nku.edu

ABSTRACT

Many institutions are challenged by the issue of student retention and persistence. In this study we used a text analytics method to summarize recent publications on student retention research. As demographics shift, new knowledge on student retention and practice should be consolidated and reviewed. Institutions can borrow from research to develop effective retention strategies.

KEYWORDS: Retention, Persistence, student success, online education, institution programming

INTRODUCTION

The National Student Clearinghouse Research Center (2016) reports that 30 percent of students who entered college in the fall of 2014 did not return in the second year. According to the National Center for Education Statistics, students enrolled in higher education in 2011-2012, only 11% of this cohort earned an associate degree and 7% earned a bachelor's degree within six years, 9% earned a certificate, and 12% were still working to complete a credential. Half of these students left their first two- or four-year college without earning a degree or certificate (Sparks, 2019).

Colleges and universities are challenged by multiple stakeholders such as state legislatures, the federal government, the broader public, and students to improve retention and graduation rates. To respond to the pressure of improving retention, many universities and institutions have created student success centers, student retention offices, etc. Programs and departments have been introduced across the country in an effort to increase the rates of retention in higher education. These initiatives are designed to improve student retention and support students with a variety of services (Fishman, Ludgate, & Tutak, 2017).

Universities and colleges are measured by graduation rates for their effectiveness. In order to improve graduation rates, institutions must improve retention rates as the institution's retention rate has a direct relationship with its graduation rate. The goal of retention programs is to reduce the number of students who leave the institution before completing a degree. Forty years ago, students who dropped out of college were viewed as lack of skills and motivations. Over the years, the dropout view has been changed. Institutions have to investigate their own internal processes, policies and procedures to understand a broad picture of students' dropout.

As colleges and universities are continually seeking ways to identify predictors of academic success and effective means to retain students, many institutions spend more money on recruiting students as opposed to retaining existing students (Fike & Fike, 2008). A prior study reports that retention of current students is far less expensive than the recruitment of new students (Lau, 2003). To help students complete college successfully, more resources should be invested in retention.

Though much retention research has been conducted (Robbins, 2004; Rendón, Jalomo, & Nora, 2000), retention and graduation statistics have remained relatively stable over the years (Sparks, 2019). Freshman attrition rates are commonly as high as 20-30% (Barclay, 2018) and 20% of all first-time, first-year students enrolled in a four-year institution were not returning to college in their second year (Feldman, 2017). Different generations have different characteristics and therefore their attitudes towards higher education and needs are different. Considering many years of research on the retention issue and the changing demographics of the student population, expectations of multiple stakeholders, we need to continuously understand retention dynamics. Summarizing recent findings could help us to better translate research into practice.

Given student retention is a persistent issue, understanding retention strategies and practices could shed some light on how to improve retention. This study intends to achieve the following: 1) review the literature on student retention; 2) perform a content analysis and a text mining analysis on the abstract of 30 recent publications on student retention; 3) summarize recent findings on student retention using a text analytic approach.

LITERATURE REVIEW

Predicting Retention

Tinto proposes a number of factors contribute to a student's decision to drop out: student characteristics, the extent of their academic, environmental, and social integration in an institution (Tinto 2006). Prior studies have identified many factors contributing to student retention. In general, the important predictors are socioeconomic status, high school GPA, college assessment entry scores when combined with institutional commitment, academic goals, social support, and academic self-confidence (Robbins, 2004). Fike & Fike (2008) report the developmental education programs and internet-based courses impact student persistence, financial aid, parents' education, the number of semester hours enrolled in the first fall semester, and participation in the Student Support Services program as strong predictors for first year persistence. (Droddy, Smith, & Guarino, 2009) conclude that school accountability status, ACT scores and first year GPA are strong predictors for first year college retention. Milea (2018) find students who were academically prepared, received grants or scholarships, and were in smaller classes have higher retention and graduation rates. Milea (2018) also suggests that universities could improve graduation and retention rates by investing in scholarships, smaller class sizes, and financial aid infrastructure.

Though many studies reported ACT/SAT scores and high school GPA are strong predictors for retention (Robins, 2014; Fike &Fike 2008), Saunder-scott et al. (2018) find that perceived stress and grit were good predictors for retention but ACT/SAT and high school GPA are strong predictors for college GPA but not for retention. Barclay and Barclay (Dewberry & Jackson, 2018) report psychological variables are important predictors for retention. Dewberry & Jackson (2018) find that theory of planned behavior variables predict student' voluntary dropout well.

These findings suggest effective retention strategies should be built upon understanding individual students, and their psychological factors.

Practices

Universities have designed a wide variety of programs including university wide, first year seminars, advising alteration, mentor/role model programs, and altering gateway courses to improve student retention. Traditionally, the most common method of retention focuses on study skills, time management, and habits. While study skill habits are effective in improving student's academic performance, retention rates have remained the same over the past decade (Scott et al., 2004). This begs for new programs to be developed. In general, retention programs fall into roughly four categories: academic, interpersonal skills, social situation, and financial situation. Nadeem, Abbas, & Javed (2015) find student orientation has a positively effect on satisfaction and retention. Cox (2005) states that enhancing first-year experience improves student retention. In addition, mentor/role model and advising programs contribute to retaining students. In a study of a two-year nursing program, effective practice in increasing retention includes adding a face to face meeting within the first two weeks of the semester as part of the advising duties plus maintaining and documenting contact at least once every 16 weeks, and responding to inquires within three days (Harrell & Reglin, 2018).

Students At-Risk

The literature produces fruitful results on how to improve retention for students at-risk. The typical risk factors are rural, low socio-economic background, minority, first-gen college, academic performance, gender etc. Chan (2015) finds that different types of interaction had different effects upon GPA. To retain low income and first-gen students, a structured first-year and learning community programs are effective when they are designed to meet the specific needs and characteristics of students from low income and first-generation backgrounds (Thayer, 2000). Peer tutors increase retention for low-income students and it is also important to realize that building self-confidence is essential for this group of students (Mangan, 2015). For rural students, three factors affect their college persistence: community's and family's values of education, the pull of family obligations, and the students' struggle to master college-level coursework (Hlinka, 2017). In the study of under-represented minority group, there is a positive impact on African-American male academics specifically to increased grade point averages with a mixed method (Brooks, Jones, & Burt, 2013). Black males have the highest dropout rate of any demographic (Strayhorn, 2014). In a study of black males attending a predominantly White institution, grit is positively related to college grades for Black males. Grit explains 24 % of the variance in Black male's college grades (Strayhorn, 2014). As the freshmen year is crucial to engage black males, universities should establish support structures such as a mentor program, same race peers, and faculty relationship to engage black males (Lucas, 2018). In order for retention strategies to be effective, the retention program should be specific and customized to the particular group at-risk (Hernandez, 2018; Brooks, 2012).

METHODS

We used a qualitative method to examine our research purposes; one is the content analysis of the title of the recent publications on retention, and the other is a text analytics approach to exact themes.

Data Collection

We performed a keyword search on student retention using ABI database. After reading the abstract of the publications, we've decided to use 30 papers published in the past four years on student retention as these papers cover a broad perspective of student retention and their findings are current. We plan in the future to include more papers for analysis.

Content Analysis

A content analysis of the title of the 30 papers was performed. We identified the themes as shown in Table 1 below. Recent research has investigated a wide variety of issues that are related to student retention: complex system, internal process, student satisfaction, macro influence, financial status, faculty personality, faculty employment status, strength based approach, professional staff, digital badges, peer mentoring, academic library, early alert systems, specialized courses etc.

Table 1: Topics
Complex system
Internal process
Student satisfaction
Macro influence
Online student success
Faculty and student perceptions
Student loans/financial stress
Faculty student lifestyle habits
Faculty employment status
Strength-based approach
Faculty personality
Professional staff
Online learning
Learning analytics and digital badges
Expectations and experiences
Theory of planned behavior
Peer mentor characteristics
Specialized courses
Predictors of retention
Early alert systems
Academic library
Technology acceptance factor
STEM majors
Folk high school

Table 2 shows the number of papers published in the selected year window of study.

Table 2: Counts of Publication by Year	
Year	
2015	7
2016	10
2017	4
2018	7
2019	2

Text Analytics

After performing a content analysis of the title of the publications and identified themes, we then proceeded with a text analysis. Knime 3.7 was used to perform a text analysis on the abstract of the 30 papers. To analyze the text data, we must complete the data wrangling process. The data wrangling process includes tokenization and normalization. Tokenization is the process of breaking content into smaller pieces, words, or tokens. An important part of tokenization is the removal of “stop words.” Stop words include prepositions and conjunctions (after, for, of, but, nor) which will not improve our understanding of the overall sentiment of a paper. Normalization is the process of reducing a group of words to its base or stem. For example, retain, retentions, retention, retaining are all variations of the same stem, retain. We want to group variations like these into one stem; a normalization process is also called word stemming. We computed inverse document frequency (IDF) and term frequency (TF) and used the computed metric TFIDF to create a word cloud that visually represents the importance of retention studies in the past four years as TFIDF is one of the most popular term-weighting schemes (Beel, Gipp, Langer, & Breitinger, 2016)

Figure 1. Word Cloud of the Abstract

Figure 1 shows the word cloud. As shown in this word cloud, the highly occurred words are online, quality, recommendations, research, education etc. As more and more online programs and courses are offered, institutions now focus on how to retain students in the online environment. More research is needed in the online quality of education. To help students persist through college, universities must understand students’ profile, their issues/problems and take a student-centered approach to design course offerings, learning experiences, campus life, making college experience enjoyable and keeping students engaged.

CONCLUSION

To be financially healthy, universities and colleges should treat retention and recruitment as their top priorities. All institutions have implemented a variety of programming to retain students. In this study we performed text analytics on recent publications to better understand recent knowledge on retention strategies and practices. Our study shows that recent publications have examined internal and external processes, faculty personality/employment status, and a strength-based approach in the area of student intention. We can borrow knowledge and practice that are implemented in other institutions and adapt them to the specific education programs and student populations.

Limitations and Future Directions

This study is limited to the text analysis of the 30 abstracts. Student retention will continue to be of research interest as generation Z enters college. In the future we plan to expand the sample size and study the difference on retention research between 4-year universities vs 2-year colleges, and online programs vs traditional face-to-face programs. In addition, the method of text mining is limited to word clouds, where only the word count is taken into account. We will use the sentiment analyses, including the text network and the topic modeling for generating explanatory variables and statistical models to analyze significance of factors and explore the relationships between factors.

REFERENCES

Beel, J., Gipp, B., Langer, S., & Breitinger, C. (2016). Research-paper recommender systems : a literature survey. International Journal on Digital Libraries, 17(4), 305–338.

Dewberry, C., & Jackson, D. J. R. (2018). An application of the theory of planned behavior to student retention. Journal of Vocational Behavior; Philadelphia, 107, 100.

Droddy, J., Smith, W., & Guarino, A. (2009). School Accountability Status, Student Core

Gpa and Act as Predictors for First Year College Retention. Allied Academies International Conference. Academy of Educational Leadership. Proceedings; Arden, 14(1), 11.

Fike, D. S., & Fike, R. (2008). Predictors of First-Year Student Retention in the Community College. Community College Review, 36(2), 68–88.

Fishman, T., Ludgate, A., & Tutak, J. (2017). Improving student success in higher education | Deloitte Insights, accessed May 28, 2019, available at https://www2.deloitte.com/insights/us/en/industry/public-sector/improving-student-success-in-higher-education.html.

Harrell, J. C., & Reglin, G. (2018). Evaluation of a Community College’s Nursing Faculty Advising Program Relative to Students’ Satisfaction and Retention. College Student Journal, 52(1), 33–48.

Hlinka, K. R. (2017). Tailoring Retention Theories to Meet the Needs of Rural Appalachian Community College Students. Community College Review, 45(2), 144–164.

Lucas, Y. (2018). Support Structures: Challenges and Potential Solutions Toward Persistence in Earning a Baccalaureate Degree for Black Male Students Attending Predominantly White Institutions. The Journal of Business Diversity; West Palm Beach, 18(1), 68–78.

Mangan, K. (2015). To Improve Retention, Community Colleges Teach Self-Esteem. The Chronicle of Higher Education.

Nadeem, M. S., Abbas, Q., & Javed, U. (2015). Assessing Relationship between Improvement in Internal Process, Student Satisfaction and Student Retention; Does Student's Orientation Mediates the Relationship? Global Management Journal for Academic & Corporate Studies; Karachi, 5(1), 124–136.

Rendón, L. I., Jalomo, R. E., & Nora, A. (2000). Theoretical considerations in the study of minority student retention in higher education. Reworking the student departure puzzle, 1, 127–156.

Robbins, S. (2004). Lotkowski, V., Robbins, S., & Noeth, R. (2004). The role of academic and non-academic factors in improving college retention. ACT Policy Report. Iowa City, IA: ACT, Inc.

Sparks, S. D. (2019). Persistence, Retention, and Attainment of 2011-12 First-Time Beginning Postsecondary Students as of Spring 2017. Education Week, 38(26), 5–5.

Strayhorn, T. L. (2014). What Role Does Grit Play in the Academic Success of Black Male Collegians at Predominantly White Institutions? Journal of African American Studies, 18(1), 1–10.

DECISION SCIENCES INSTITUTE

Use of Auxiliary Variable for Efficient Estimation of Average Production of Peppermint Yield

S. K. Yadav
Babasaheb Bhimrao Ambedkar University
Email: drskystats@gmail.com

Dinesh K. Sharma
University of Maryland Eastern Shore
Email: dksharma@umes.edu

Kate Brown
University of Maryland Eastern Shore
Email: kbrown@umes.edu

ABSTRACT

In this study, we consider an auxiliary variable to improve the estimation of the population mean of the primary variable, and propose another group of estimators for improved estimation of the average yield of peppermint oil. Also, we compare the proposed estimators with competing estimators theoretically and derive the conditions for the suggested class to be an improvement. Finally, original data collected from the Siddaur Block of Barabanki District at Uttar Pradesh State (India) are used to compare empirically the proposed class of estimators with the competing classes of estimators.

KEYWORDS: Study Variable, Auxiliary variable, Regression-cum-Ratio estimator, Bias, MSE, PRE.

INTRODUCTION

Auxiliary information enhances the efficiency of the sampling design. This information is achieved from the auxiliary variable which has a high degree of positive or negative correlation with the study variable. Many authors have utilized the auxiliary information for elevated estimation of the population mean. Some of the contributions in this direction is due to Sharma and Singh (2013), Subramani and Kumarapandiyan (2012, 2013), Yadav and Mishra (2015), Yadav et al. (2016), Subramani (2016), Abid et al. (2016), Gupta and Yadav (2017, 2018), Subramani and Ajith (2017), Subzar et al. (2017), Srija and Subramani (2018), Ijaz and Ali (2018) and Yadav et al.(2018**).**

In this study, we consider enhanced estimation of the population mean of the main variable utilizing known information acquired from an auxiliary variable. We introduce a new family of estimators for the elevated estimation of average production of peppermint oil at Siddaur Block of Barabanki District at Uttar Pradesh State in India. We study large sampling properties using Taylor series expansion up to the first order of approximation both theoretically and empirically.

LITERATURE REVIEW

A number of improved estimators of the population mean have been suggested by various authors. Some of the estimators under comparison are presented in Table 1.

Table 1: Different estimators of population mean by various authors

Year	Authors	Journal
2004	Kadilar and Cingi	Applied Mathematics and Computation
2006	Kadilar and Cingi	Hacettepe Journal of Mathematics and Statistics
2012a	Subramani and Kumarapandiyan	International Journal of Probability and Statistics
2012b	Subramani and Kumarapandiyan	American Journal of Mathematics and Statistics
2016	Abid et al.	Pakistan Journal of Statistics and Operations Research
2017a	Subzar et al.	World Applied Sciences Journal
2017b	Subzar et al.	International Journal of Modern Mathematical Sciences
2018a	Subzar et al.	Applied Mathematics and Information Sciences Letters
2018b	Subzar et al.	Statistics in Transition New Series

PROPOSED ESTIMATOR

Motivated by Subzar et al. (2017b) and Jerajuddin and Kishun (2016), we suggest the following family of estimators of the population mean using non-conventional parameters and the ninth decile of the secondary variable as,

$$t_{pi} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x} + \xi_j)}(\bar{X} + \xi_j),\ i = 1,2,...,8 \text{ and } j = 1,2,...,8$$

$$t_{pi} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x}\rho + \xi_j)}(\bar{X}\rho + \xi_j),\ i = 9,10,...,16 \text{ and } j = 1,2,...,8$$

$$t_{pi} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x}C_x + \xi_j)}(\bar{X}C_x + \xi_j),\ i = 17,18,...,24 \text{ and } j = 1,2,...,8$$

The biases and the mean squared errors of the above estimators up to the first order of approximation are, respectively,

$$B(t_{pj}) \cong \frac{1-f}{n}\frac{S_x^2}{\bar{Y}}R_{pj}^2,\ j = 1,2,\ldots,24$$

$$MSE(t_{pj}) \cong \frac{1-f}{n}[(R_{pj}^2 S_x^2 - (1-\rho^2)S_y^2],\ j = 1,2,\ldots,24$$

Where,

$$R_{pi} = \frac{\bar{Y}}{\bar{X} + \xi_j},\ i = 1,2,...,8 \text{ and } j = 1,2,...,8$$

$$R_{pi} = \frac{\bar{Y}\rho}{\bar{X}\rho + \xi_j},\ i = 9,10,...,16 \text{ and } j = 1,2,...,8$$

$$R_{pi} = \frac{\bar{Y}C_x}{\bar{X}C_x + \xi_j}, \ i = 17, 18, ..., 24 \text{ and } j = 1, 2, ..., 8$$

and,

$\xi_1 = (QD \times D_9)$, $\xi_2 = (DM \times D_9)$, $\xi_3 = (TM \times D_9)$, $\xi_4 = (MR \times D_9)$, $\xi_5 = (HL \times D_9)$, $\xi_6 = (G \times D_9)$, $\xi_7 = (D \times D_9)$, $\xi_8 = (S_{pw} \times D_9)$

EFFICIENCY COMPARISON

In this section, the suggested class of estimators is compared with the competing families and the theoretical conditions for the suggested family to be more efficient are given in Table 2.

Table 2: Efficiency conditions of proposed estimator over the competing estimators

S. No.	Estimator by	Efficiency condition of suggested estimator to be more efficient than other if,
1.	Kadilar and Cingi (2004)	$MSE(t_{ai}) - MSE(t_{pj}) > 0$
2.	Subramani and Kumarapandiyan (2012a, 2012b)	$MSE(t_{ci}) - MSE(t_{pj}) > 0$
3.	Abid et al. (2016)	$MSE(t_{di}) - MSE(t_{pj}) > 0$
4.	Subzar et al. (2017)	$MSE(t_{ei}) - MSE(t_{pj}) > 0$
5.	Subzar et al. (2018)	$MSE(t_{hi}) - MSE(t_{pj}) > 0$

NUMERICAL STUDY

In this section a real population has been considered and the data has been collected from a block of the Barabanki district of Uttar Pradesh State of India.
The study and the auxiliary variables under considerations are as follows:
Y : The production (Yield) of peppermint oil in kilograms
X : The area of the field in Bigha (2529.3 Square Meter)

Table 3 represents the parameters of the population under consideration.

Table 3: Parameters of the population under consideration

$N = 150$, $n = 40$, $\gamma = 0.018333$, $\bar{Y} = 79.58$, $\bar{X} = 6.5833$, $\rho = 0.9363$, $C_y = 0.781333$, $C_x = 0.661726$, $S_y^2 = 3866.165$, $S_x^2 = 18.97791$, $\beta_1 = 1.4984$, $\beta_2 = 5.408$, $Q_1 = 4$, $M_d = 5$, $Q_3 = 10$, $D_1 = 2$, $D_2 = 3$, $D_3 = 4$, $D_4 = 5$, $D_5 = 5$, $D_6 = 6$, $D_7 = 8$, $D_8 = 10$, $D_9 = 13$, $QD = 3$, $DM = 6.22$, $TM = 6$, $MR = 11$, $HL = 7$, $G = 8.2298$, $D = 9.2542$, $S_{pw} = 9.3707$

The MSEs of some competing and suggested estimators (see Table 2) of the population mean for the mentioned natural population are presented in the following Table 4.

Table 4: MSE of competing and suggested estimators

Estimators	MSE	Estimators	MSE
t_{a1}	59.5772	t_{p1}	9.7970
t_{c1}	31.1269	t_{p6}	8.8734
t_{d1}	22.1784	t_{p7}	8.8702
t_{e1}	50.6571	t_{p12}	8.8234
t_{e18}	56.1919	t_{p16}	8.8545
t_{h1}	14.3495	t_{p21}	8.8427
t_{h18}	12.1847	t_{p24}	8.7972

CONCLUSION

In the present study, we investigated the enhanced estimation of the population mean using an auxiliary variable. We demonstrated that the suggested class of estimators may be used for improved estimation of the population mean. We derived the expressions for the bias and mean squared error of the suggested estimators up to an approximation of degree one. The suggested estimators are compared theoretically with the competing estimators using auxiliary variable and the efficiency conditions for the suggested estimators to be more efficient are also obtained. The theoretical conditions are verified through natural primary data on peppermint oil production collected from a block of the Barabanki District of Uttar Pradesh in India. The numerical study shows that the suggested estimators have the least MSEs among the class of competing estimators of the population mean. Thus the suggested family may be used for enhaced estimation of the population mean of similar products.

REFERENCES

Abid, M., Abbas, N. Sherwani, R.A.K., & Nazir, H. Z. (2016). Improved Ratio Estimators for the population mean using non-conventional measure of dispersion. *Pakistan Journal of Statistics and Operations Research*, XII (2), 353-367.

Kadilar, C., & Cingi, H., (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151, 893–902.

Kadilar C., & Cingi, H. (2006). An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe Journal of Mathematics and Statistics*, 35 (1), 103–109.

Subramani, J. & Kumarapandiyan, G. (2012a). Estimation of population mean using co-efficient of variation and median of an auxiliary variable. *International Journal of Probability and Statistics* 1, 4, 111–118.

Subramani, J., & Kumarapandiyan, G. (2012b). Estimation of population mean using known median and co-efficient of skewness. *American Journal of Mathematics and Statistics*, 2, 5, 101-107.

Subramani, J., & Kumarapandiyan, G. (2013). A new modified ratio estimator of population mean when median of the auxiliary variable is known. *Pakistan Journal of Statistics and Operation Research*, Vol. 9 (2), 137–145.

Subramani, J., & Master Ajith, S. (2017). Improved Ratio cum Product Estimators for Finite Population Mean with Known Quartiles and Their Functions. *SM J. Biometrics Biostat*., 2, 1, 1008.

Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean. *Statistics in Transition New Series*, 17, 4, 1-14.

Subzar, M., Maqbool, S., Raja, T.A., & Shabeer, M. (2017a). A New Ratio Estimators for estimation of Population mean using Conventional Location parameters. *World Applied Sciences Journal*, 35 (3), 377-384.

Subzar, M., Maqbool, S., Raja, T. A., Pal, S.K., & Sharma, P. (2018b). Efficient Estimators of Population Mean Using Auxiliary Information under Simple Random Sampling. *Statistics in Transition new series*, 19, 2, 219-238.

Yadav, S.K., Sharma, Dinesh K., Mishra, S.S., & Shukla, A.K. (2018). Use of Auxiliary Variables in Searching Efficient Estimator of Population Mean. *International Journal of Multivariate Data Analysis*, 1, 3, 230-244.

(A complete list of references is available upon request)

DECISION SCIENCES INSTITUTE

Investment fund managers – rational or not – Polish capital market example

Rafal Wolski
University of Lodz
Email: rafal.wolski@uni.lodz.pl

Monika Bolek
University of Lodz
Email: monika.bolek@uni.lodz.pl

ABSTRACT

The aim of this paper is to assess the rationality of certified investment advisors employed as mutual fund managers operating on the Polish capital market by examining the decision biases to which they are subject. Investment fund managers are expected to make rational decisions while managing the capital of individuals and this hypothesis is verified in the paper. The method of tasks identifying the behavioural traps is used in the survey. As the result of the survey it is concluded that investment fund managers are rational basic on the answers they gave to the questions analyzing their decisions.

KEYWORDS: Rationality, Investment fund managers, Behavioural biases

INTRODUCTION

Investment fund managers make their decisions in the environment of continuous process related to the information coming from different local and international institutions. The information process is not coordinated and there are periods with more and less intensive news appearing on the market. That news affect rates of return and its volatility in a positive and negative way. There is a wide scope of evidence related to the overreaction and under reaction of the market on the available information indicating, that the model of making decision is not a simple project but more complicated process that can be described only in a small part. Decisions that investors make on the market are not only the result of the external information but also the internal motivations and preferences. In case of investment fund managers it can be related to the corporate governance, knowledge and experience all together with heuristics they are affected by. In this paper the behavioural biases are examined but it is only a small part of the rationality assessment.
The effects of certainty and reverse in a prospect theory, isolation (framing), disposition, overconfidence, illusion of control, ambiguity aversion, attribute substitution, sunk costs and fast thinking are analysed in the presented paper in the group of investment fund managers on the Polish capital market to verify the hypothesis that their decisions are rational. The survey is made on an emerging capital market but it can be expected, that there is a little difference between investment fund managers on other markets since the capital flow, information and knowledge is universal in a global financial world.

LITERATURE REVIEW

Rationality is a multidimensional issue that can be considered from different points of view. In case of papers in the area of finance authors stress on the economic approach to this issue. But the concept of rationality differs between psychology, philosophy, economics, and biology. For psychologists and philosophers, the emphasis is on the process by which decisions are made: rational beliefs are arrived at by reasoning and contrasted with beliefs arrived at by emotion, faith, authority or arbitrary choice. Economists emphasize consistency of choice, regardless of the process and the goal. Biologists use a concept that links both previous ideas. (Kacelnik, 2006). Moreover no such thing exists as a rationality that is not the rationality of some tradition. Each tradition is developed within a particular historical context and sought to resolve particular conflicts. Allegiance to one tradition can allow for meaningful contact with other traditions in a way that can lead to understanding, vindication, or revision of that tradition in its continuing form. Thus, only by being grounded in the history and opposing traditions we are able to restore rationality and intelligibility to our moral attitudes and commitments today (Macintyre, Macintyre,1988). Analysing the rationality is a complicated project, but every survey may add something to the general theory.

Rational behaviour refers to a decision-making process that is based on making choices that result in the optimal level of benefit or utility for an individual. Most conventional economic theories are based on the assumption that all individuals taking part in an action or activity are behaving rationally. Rationality is based on wisdom, logical reasoning and choice of means and methods of action, naturally within the scope of knowledge (Penc, 2007). Rational behaviour of managers on the market may be accompanied by negative externalities, described by microeconomics as market weaknesses (Fornalczyk, 2007).

There are as many types of rationality as codes of conduct. Analytical rationality includes every action based on the analysis of elements of a given situation. Holistic rationality is based on a multidimensional view of the problem. The limited rationality corresponds to the information available and the actual human calculating abilities: the decision-maker first determines the intervals of satisfactory grades, and then rests on the first option considered from such a range. When analysing decisions, one needs to be aware of what type of rationality we are dealing with (Szapiro, 1993).

Dean and Perlman (1998) point out that the one of the opponents of the traditional concept of homo oeconomicus was Leibenstein, the author of selective rationality, i.e. selective tendency and ability to calculate. Leibenstein pointed to the gradability of rationality (strict and loose) and assumed that people act as they please, or as they think they must act or choose a compromise solution between these two possibilities. Behaviours that do not maximize usability are by no means irrational. Sensible behaviour requires only marginal rationality, not total rationality at all times (Adamska, 2008).

By observing and describing the real behaviour of people in organizations, the researchers tried to find out what makes real decision-makers behave differently than the models of perfect rationality recommend them. There are many reasons for this state of affairs. They can be divided into two groups (Bolesta-Kukułka, 2003):

- psychological reasons, i.e. disturbances of rationality resulting from cognitive and emotional distortions that man's common thinking is subjected to;
- organizational reasons resulting from the very nature of the organization.

Bounded ethicality is based on the term bounded rationality, which explains how rationality is limited. Bounded rationality was first used by Simon (1986) who explained how people act rationally in their decision-making, their rationality is limited by the information they have, their own mental limitations and the limits on the time frame available for the decision-making. As such, Simon noted, people are more focused on reaching a satisfactory decision, rather than the optimal one.

Simon stressed that rationality have different meaning in economics and psychology and economy from the beginning was treating people as rational while psychology always concerned people as irrational. People make decisions on the purpose driven by motivations

and goals. In the investment process there are some goals that can be mentioned like to maximize the return, reduce the risk or look for a balance. The problem of achieving the goal is another fact since the behavioural biases may influence the decisions and affect the results of investment, therefore economic approach may emphasize irrationality while the psychological approach explain it in another way looking for the answer what constitute the rationality, not what it is in itself. The assessment of rationality should be in a context of situation, goals and the way they will be attended. In economics rationality is viewed in terms of choices it produces, in the other social sciences it is viewed in the process it employs.
Bounded rationality (Kahneman, 2003) was presented following Simon's ideas and states that decision makers are bounded rational and the utility maximization was replaced by satisfaction in this proposal. Kahneman and Tversky explored three programs of research: first heuristics that people use and the biases to which they are prone in tasks of judgement under uncertainty (Kahneman Tversky 1973, Tversky and Kahneman 1974, Kahneman et al. 1982). The second was concerned with prospect theory, a model of choice under risk (Kahneman and Tversky 1979, Tversky and Kahneman 1992) and loss aversion with riskless choice (Kahneman et al 1990, 1991, Tversky and Kahneman 1991). The third line was dealing with framing effect (Tversky and Kahneman 1981, 1986). Rational models are psychologically unrealistic and the previous areas are extended by unified treatment of intuitive judgements and choice based on the previous study of the relationship between preferences and attitudes (Kahneman et al 1999) and extends the model of judgement heuristics proposed by Kahneman and Frederic (2002). Kahneman (2003) claims that the most choices and judgments are made intuitively, rules that govern intuition are generally similar to the rules of perception. Cognitive approach and distinguishing the reasoning and intuition is one step toward understanding the decision process.In case of investment fund managers, who are professional and certified advisors, hard thinking may be the common behaviour that can affect the cognitive process of decision making. Intuition and reasoning can be joined in the systems of decision making in the cognitive process and one can be fast and intuitive, while the second is slower and effortful. Effort is the main difference and in case of investment fund managers it is common to use an effort. People who are occupied by demanding mental activity are much more likely to respond to another task by blurting out whatever comes to mind (Gilbert 1989). Accessibility is another issue that can be defined as ease with which mental contents come to mind (Higgins 1996). The acquisition of skills gradually increase the accessibility of useful responses to the effortless level. Intuitive decisions described by Klein (1998) show that experienced decisions makers working under pressure rarely need to choose between options because only one answer comes to their minds. Rational agent makes his choices in a comprehensively inclusive context with all relevant details of the present situation and expectations about all future opportunities and risks. People are characterized by narrow framing, mental accounting, decision bracketing but it can help in case of investment fund managers to make fast decision basing on hard thinking they are subject to.
Investors' decisions about a particular investment appear to be considered in isolation from the remainder of investors portfolio (Barberis, Thaler 2003). The time horizon investors evaluate their investment is very short and affects the equity premium puzzle (Benartzi, Thaler 1995). Narrow framing is far from risk neutrality than decision made in a more inclusive context. Tversky and Kahneman in 1974 stated that people rely on limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predict the values to simpler judgemental operations. In general these heuristics are useful but sometimes they lead to severe and systematic errors. Heuristics such as representativeness, availability and anchoring were described to explain such bias in judgement under uncertainty. In 2002 Kahneman and Frederic proposed attribute substitution, changing the judgement under uncertainty.

There are many heuristics affecting the decisions of investors on the market. The group investment managers are in a specific situation since they do not make decisions over their money but behave as agents toward the owners of capital, providing rate of return in relation to their expectations. Even though they are professionals they decisions may be irrational but it can be expected, from the other hand, that the knowledge and experience may reduce the irrationality and influence the process of decision making in a positive way. Investment fund managers can be replaced by automatic systems and machines making decisions (Teixeira, De Oliveira, 2010) but we cannot be certain, that artificial intelligence built on the scheme of human brain will not suffer the behavioural traps.

In the literature there are many proposals describing and identifying the behavioural biases. It seems like this area of behavioural economics is trying to implement as much as possible from the science of psychology to explain the market inefficiency and investors irrationality. Loss aversion or prospects theory related to the individual's desire to avoid losses than comparable profits (Kahneman, Tversky 1979) is a trap that investors are a subject to when making decisions. The disposition effect is a behavioural trap that means a tendency to quickly getting rid of shares in upward trends and to excessively hold shares of companies whose prices are falling (Shefrin and Statman, 1985). Overconfidence is a bias that affects the decision of individual as well as corporate world. People have a propensity to overestimating their abilities and avoid taking the help of others in decision making process. These people are totally rely on their abilities and therefore they search less help and direction during the decision making process. According to Shefrin "pertains how well the people understand their own abilities and the limits of their knowledge". Excessive optimism' refers to the overconfident and overestimation of the favourable outputs rather than unfavourable outputs (Shefrin 2007). Confirmation bias is a cognitive trap or the ability of people to understand information in such a way that it confirms the previous ideas while avoiding explanation which disagrees with previously held beliefs (Shefrin, 2007). Illusion of control is defined as the propensity of people to believe that they can control or affects the outputs that in reality they have no affects over (Shefrin 2007). Illusion of control appears when people give wrong impression that their personal involvement influenced the outputs but the reality is quite different. Mental accounting is defined from the other hand as the process in which the people code, categorize and evaluate the economic outputs (Thaler, 1980). There are only examples of biases and the methods of recognizing them seems to be adequate for implementation in the presented survey.

The dominant paradigm regarding the response of asset prices to new information (first articulated by Fama, 1971) is that, since markets are efficient, asset prices should react immediately and in an unbiased manner to new information. Orphanides, Williams (2005) state that the learning process lead to different behaviour than in rational expectations equilibrium. A number of studies have considered the impact of the unexpected component of scheduled news released on individual markets (Balduzziet al. 1999). More information may induce not more but less clarity and common understanding among market participants, as there are limits to how much information can be digested effectively (Kahneman, 2003).

Every person has cognitive limits, i.e. a constraint to absorb and process information only up to a certain point. Such limits bind utility-maximising rationality, as conventionally conceived in economics. What is more, the investment-making process includes quasi-rational motives, which satisfy psychological needs . In the latter case, the paramount need is likely to be the desire to continue being able to decide at all, followed by the desire to control actions and to minimise cognitive dissonance, for example, with regard to previous decisions (Shefrin and Statman, 1985).

Efficiency of the market is based on the assumption that investors are rational, irrational investors behave chaotic and their decisions match and eliminate themselves, irrational investors decisions are eliminated by arbitrageurs (Shleifer, 2000). Brenner et al (2009) suggest that some of this evidence is consistent with existing theoretical arguments in the literature, e.g., those emphasizing the role of investors' trading activity.

Market sentiment is the overall attitude of investors toward a particular security or financial market. Market sentiment is the feeling or tone of a market, or its crowd psychology, as revealed through the activity and price movement of the securities traded in that market. In broadest terms, rising prices would indicate bullish market sentiment, while falling prices would indicate bearish market sentiment.

Abreu and Brunnermeier (2003) argue that noise traders are subject to animal spirits, fads and fashions, overconfidence and other psychological biases. Chung, Hung, and Yeh (2012) emphasize that sentiment-driven overpricing is more prevalent than underpricing due to the limits of arbitrage and short-sales constraints. All their evidence suggests that sentiment affects the stock price reaction to monetary policy news. Crucially measures that are orthogonalized to a set of macroeconomic variables, and thus, the state dependence is related to optimism or pessimism that is unwarranted by developments in economic fundamentals. It highlights the existence of a separate behavioural channel.

There is a wide scope of evidence related to the overreaction and under reaction of the market (Barberies, Shleifer, Vishny 1998) on the available information indicating, that the model of making decisions is not a simple project but more complicated process. Khan (2018) stated that behavioural approach is related to many aspects. Norms influence decisions of decisionmakers can be divided for social, market and legal. From their perspective there is a chance of simplifying the decision making process. Behaviour of other influence the decisions as well from the internalization, identification and conformity perspectives.

Investment fund managers may be subject to rules related to corporate governance. The theory of the agency was formulated by Alchian, Demsetz (1972), and it was developed by Jensen and Meckling (1976) who refer to it in the relationship between principals (shareholders) and agents (managers). According to the theory of the agency, the principals hire agents to accomplish their goals. Investors expect some actions of managers, who in turn may avoid implementing them (Ross 1973). Another proposal to explain the problem is corporate governance is the theory of stewards, which was presented by Davis et al. in 1997. Its authors stated that the steward aims to maximize and protect shareholders, because by doing so, it maximizes its own utility function. Fama (1980) and Shleifer and Vishny (1997), however, argue that managers' actions are motivated by building their own reputation on the market and questioning the assumptions of the steward theory at the same time. Both theories can be applied to the investment fund managers whose motivations can be related to the role they play in the relation to investors and their expectations.

Brzeszczynski et al (2015) stated that the role of institutional investors has grown substantially in emerging markets. Developed and developing countries policy-makers alike have promoted institutional investors as a pillar of their financial systems. Among many different objectives, they are expected to invest for the long term, follow market fundamentals, provide liquidity to countries and companies overlooked by other financial markets participants and reduce many of the shortcomings of the financial system. Although the significance of this group of investors is still relatively small in some emerging markets, they are nevertheless important players there and, therefore, they are subject of numerous research studies which document their positive role but sometimes also negative effects of their actions. Study presented in this paper extend the findings of Brzeszczynski et al. in the field of behavioural biases of managers on the emerging market.

HYPOTHESES/MODEL

The survey was conducted in 2017 and 2018, a total of 17 questionnaires were collected. The surveyed were managers of investment funds. The survey was divided into three parts. The first is a metric in which managers gave answers to questions about gender, the fund's time horizon and investment strategy. In the second part of the survey, the in-depth interview was conducted on the functioning of the fund. The questions concern both the fund's strategy and the basis for making decisions. In this part, the authors of the survey wanted to learn more about the nature of the fund and how it was managed. In total, 16 questions were asked. Both in the first and in the second part of the survey, there are closed questions, where the respondent has a choice of standardized answers as well as open questions. In the last part, the analysis of behavioural biases was conducted and 10 questions were asked in the form of tasks.
Task 1 measures the effect of certainty, or tendency to reevaluate certain events in relation to highly probable events. This task measures whether a person will choose a certain profit, but with a lower expected value (Option B - vulnerability) or choose an option more rational, potentially higher expected value (A) (Kahneman 2003).
Task 2 measures the reverse effect, which means that majority of people are characterized by risk aversion in the area of profits and risk in the area of losses. Selecting option A the tested person is susceptible to the reversal effect, especially when he chose answer B in Task 1 (Kahneman 2003).
Task 3 is a task for the effect of isolation (framing), that is related to different forms of presenting the same decision problem that may affect other decisions of the respondents (Kahneman 2003).
Task 4 is a task for the disposition effect measuring the tendency of investors to sell assets that have increased in value, while keeping assets that have dropped in value with the only rational option A, while others mean the effect of disposition. It may also examine the degree of "severity" of this disposition effect (Szyszka, 2009).
Task 5 is a task for overconfidence, the so-called effect of being better than average. Here, it is analyzed whether the respondents in various spheres will consistently indicate that they are above-average better in a given field (Czerwonka, 2017)
Task 6 is a task measuring the manifestation of overconfidence called "the illusion of control" which is based on the illusory conviction of many people that they can affect the course of future events of a random nature (Heath, Tversky 1991).
Task 7 is measuring the paradox of Ellsberg, analyzing the phenomenon of aversion to ambiguity (Segal, 1987).
Task 8 is a modified version of the Linda and attribute substitution problem assessing the representativeness heuristics (Tversky, Kahneman 1983).
Task 9 is a task for the sunk costs, managers often face (Kahneman, Lovallo, 1993, Szyszka 2009).
Task 10 is a task for fast and short thinking affecting the decisions (Kahneman, Egal 2011).
The essence of the survey was to determine the occurrence of behavioural heuristics in the group of people professionally associated with the management of investment fund portfolios. A research hypothesis has been put forward that these people with appropriate education and competences will be aware of the occurrence of specific behaviours and therefore will not be subject to behavioural heuristics. The study was aimed at examining the occurrence of the following effects: the effect of certainty, reversal, isolation (framing), disposition, overconfidence, ambiguity aversion - avoiding ambiguities, heuristics of representativeness, sunk costs and the effect of fast thinking.

METHODS

In the survey that is composed of two parts, in the first one the consistency of the questionnaire is analysed and in the second the statistical analysis of the answers to the questions that investment fund managers gave is made.
Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of items are as a group. It is considered to be a measure of scale reliability. A "high" value for alpha does not imply that the measure is unidimensional. Cronbach's alpha is not a statistical test – it is a coefficient of reliability (or consistency). Cronbach's alpha can be written as a function of the number of test items and the average inter-correlation among the items. Below, for conceptual purposes, we show the formula for the Cronbach's alpha:

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N-1) \cdot \bar{c}}$$

Here N is equal to the number of items, c-bar is the average inter-item covariance among the items and v-bar equals the average variance. [https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/]

RESULTS

The first stage of work on the results of the research was to conduct a reliability analysis allowing to assess how many questions are actually able to identify the occurrence of certain behavioural effects. The analysis was carried out in three variants. For the sensitivity analysis, the output data should be presented in the form of a matrix presented in Table 1.

Table 1. Number of responses in subsequent variants of the analysis.

variant	task and number of answers																
I	**1**	**2**	**3a**	**3b**	**4**	**5a**	**5b**	**5c**	**5d**	**5e**	**5f**	**6**	**7a**	**7b**	**8**	**9**	**10**
A	11	8	9	5	5	15	12	8	9	10	1	6	13	12	10	10	16
B	5	9	2	5	1	2	5	9	8	7	16	9	1	3	7	7	1
C	0	0	5	5	2	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0
II	**1**	**2**	**5a**	**5b**	**5c**	**5d**	**5e**	**5f**	**6**	**7a**	**7b**	**8**	**9**	**10**	**x**	**x**	**x**
A	11	8	15	12	8	9	10	1	6	13	12	10	10	16	x	x	x
B	5	9	2	5	9	8	7	16	9	1	3	7	7	1	x	x	x
III	**1**	**2**	**3a**	**3b**	**4**	**5**	**6**	**7a**	**7b**	**8**	**9**	**10**	**x**	**x**	**x**	**x**	**x**
A	11	8	9	10	5	55	6	13	12	10	10	16	x	x	x	x	x
B	5	9	7	5	10	47	9	1	3	7	7	1	x	x	x	x	x

Source: own study

The questionnaire contained 10 questions, but question 3 had two sub-tasks in which one could choose answers A, B and C. In addition, the fourth question had four possible answers. This led to the necessity of equating in the test the number of possible answers for all questions. In place of additional answers, it was always entered 0. The matrix prepared in this way was tested using the first variant of the study.

Table 2. Cornbach's alpha test results in option 1.

Avarage	69	Sum	275

Standard deviation	73,39	Variance	5386
Bias	0,59	Curtosis	-2,64
Minimum	7	Maximum	160
Cronbach's Alpha	0,96	Standardized alpha	0,96
Average correlation between positions	0,85		

source: own study in the Statistica package

In the first variant, the Cornbach alpha assumes a high value of 0.96. Confirmation of the correctness of Cornbach's alpha calculation is lower than the first alpha value calculated on standardized positions. Because the alpha is standardized at the same level, the test may not be fully reliable. The test was repeated in variant 2. In variant 2, the analysed questions with more than two possible answers were removed. This meant omitting questions 3 and 4.

Table 3. Cronbach's alpha test results in option 2			
Avarage	115	Sum	230
standard deviation	36,77	Variance	1352
Bias	0	Curtosis	0
Minimum	89	Maximum	141
Cronbach's Alpha	0,69	Standardized alpha	0,66
Average correlation between positions	1		

source: own study in the Statistica package

The estimated Cornbach alpha dropped significantly to 0.69. It is a value on the verge of recognition. The standardized alpha is lower, which makes the test more credible. The reasons for the drop in alpha value can be found in the construction of tasks. The fifth question consisted of six similar tasks, which always, according to the accepted research hypothesis, were expected to answer "no" marked as "B" in the matrix. A similar situation took place in question 7, where the same answer was also expected from respondents, this time in line with the adopted hypothesis was the answer "A". The similarity and the number of these questions may have resulted in lower reliability measured with alpha.

The third option assumed the answers in question 3 and 4 to the values "A" and "B". These questions checked the occurrence of the zero-one feature - that is, the bias occurs or does not. In question 3a, the trait occurred for the answer "A", while for the answer "B" and "C" the trait did not occur. The authors counted "B" and "C" answers to one category, adding them together. In question 3b, the trait existed for the answer "B", while for the answer "A" and "C" the trait did not exist. The answer number "C" was therefore added to the answer "A". Similarly in question 4. The trait did not occur for the "A" answer, the trait appeared for the "B", "C" and "D" responses. In this case, the "B", "C" and "D" links were added together and saved in the matrix as the "B" response. Thus, the number of rows of matrices was reduced to two, equalizing them for all questions. In addition, for question 5, all six variants were added to one case, together counting all the answers to "yes" - "A" in the matrix and "no" - "B" in the matrix. The results of the analysis for option 3 are presented in the table 4.

Table 4. Cronbach's alpha test results in option 3.			
Avarage	138	Sum	276
Standard deviation	38,18	Variance	1458

Bias	0	Curtosis	0
Minimum	111	Maximum	165
Cronbach's Alpha	0,85	Standardized alpha	0,73
Average correlation between positions	1		

source: own study in the Statistica package.

In the third variant, Cornbach alpha was obtained at 0.85 and standardized alpha at 0.73. On the basis of the tests carried out, it can be said that the questions allow to make correct inferences, hence a further interpretation of the collected data and generalization of results is possible. In the further part of the analysis, the data were aggregated in such a way as to check how many respondents answered the questions so that they confirmed the existence or not of the given effect. The results of such aggregation of data are presented in Table 5.

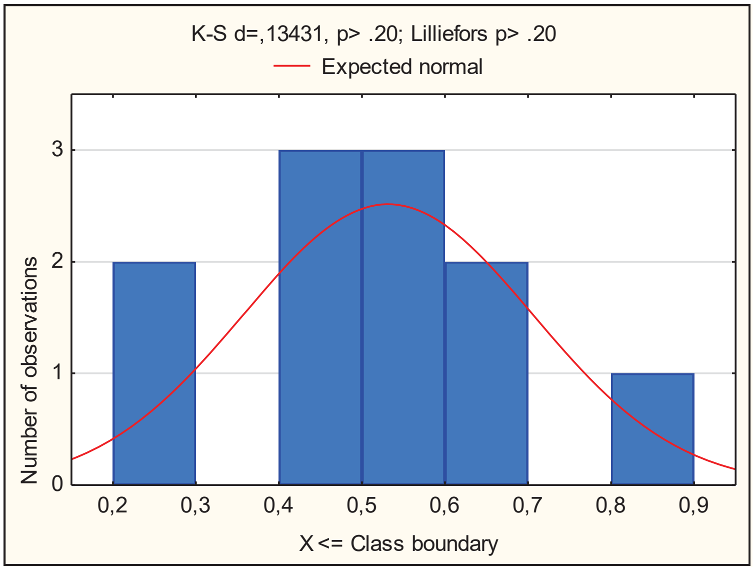
Table 5. Behavioural effects in subsequent questionnaires.

		Effect exists	Effect does not exists	Effect exists	Effect does not exists	Effect exists	Effect does not exists
Task1	Certainty effect	6	12	33,33%	66,67%	0	1
Task 2	Reverse effect	8	9	47,06%	52,94%	0	1
Task 3a	Isolation (framing) effect together with certainty and reverse	9	7	56,25%	43,75%	1	0
Task 3b	Isolation (framing) effect together with certainty and reverse	5	10	33,33%	66,67%	0	1
Task 4	Disposition effect	12	5	70,59%	29,41%	1	0
Task 5	Overconfidence effect	55	45	55,00%	45,00%	1	0
Task 6	Overconfidence effect	6	9	40,00%	60,00%	0	1
Task 7	Ambiguity aversion	12	5	70,59%	29,41%	1	0
Task 8	Represantivness heuristic	9	7	56,25%	43,75%	1	0
Task 9	Sunk cost	7	10	41,18%	58,82%	0	1
Task 10	Fast thinking	2	15	11,76%	88,24%	0	1
					Total	5	6

Źródło: opracowanie własne.

Although in 6 cases out of 11 it was confirmed that there is no selected behavioural effect, this analysis seems too vague to draw conclusions. Hence, a study was carried out on the distribution of the percentage of respondents responding in a way that negates the occurrence of subsequent effects. This distribution is shown in Figure 1.

Figure 1: The distribution of answers confirming the lack of successive effects.



source: own study in the Statistica package.

The study shows that the distribution of the percentage of responses negating the occurrence of a given effect has no normal distribution. Descriptive statistics are presented in table 6.

Table 6. Descriptive statistics of the percentage of respondents confirming the lack of successive effects.

N	11	Variance	3,05%
Avarage	53,15%	Standard deviation	17,45%
Median	52,94%	Bias	0,45
Minimum	29,41%	Curtosis	0,26
Maximum	88,24%		

source: own study in the Statistica package.

Skewness of 0.45 indicates a right-sided distribution, while positive kurtosis indicates a leptokurtic distribution. This information combined with information on the average of 53.15% and the median of 52.94% suggests that the majority of answers given by respondents show that they lack behavioural effects.

DISCUSSION AND CONCLUSIONS

Investment fund managers are expected to be rational and use their knowledge and experience to manage effectively the capital of individual investors, who trust their competences. The assumption of rationality is not under doubt and the knowledge and experience that the certified investment advisors must have seems to be sufficient to prove their rationality. In this survey the research toward verifying this hypothesis is presented.

The majority of answers given by respondents show that they lack behavioural effects. Thus, the hypothesis can be considered as positively verified. However, the results must be approached with a great deal of caution for at least several reasons. The first is a relatively small sample of respondents. In addition, the rightness of the distribution of response rates is not dominant, which is confirmed by the mean and median values. The Cronbah alfa was analysed to prove that the questionnaire that was used for the research is built properly. Investment advisors have all the information and are able to use them effectively because they are rational. Therefore, if there were only investment funds on the market, the market could be more effective. Direct comparison with investors operating on other markets could prove that there is no difference between rationality of investment fund managers on a global market.

REFERENCES

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and crashes. Econometrica, 71(1), 173-204.
Adamska M. (ed.), (2008), Bankructwa gospodarstw domowych. Perspektywa ekonomiczna i społeczna, Difin, Warszawa
Alchian, A. A., Demsetz, H. (1972). Production, information costs, and economic organization. The American economic review, ss.777-795.
Balduzzi, P., Elton, E. J., & Green, T. C. (2001). Economic news and bond prices: Evidence from the US Treasury market. Journal of financial and Quantitative analysis, 36(4), 523-543.
Bolesta-Kukułka K., (2003), Decyzje menedżerskie. , Polskie Wydawnictwo Ekonomiczne, Warszawa
Brzeszczyński, J., Gajdka, J., & Kutan, A. M. (2015). Investor response to public news, sentiment and institutional trading in emerging markets: A review. International Review of Economics & Finance, 40, 338-352.
Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance, 1, 1053-1128.
Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment1. Journal of financial economics, 49(3), 307-343.
Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. The quarterly journal of Economics, 110(1), 73-92.
Brenner, M., Pasquariello, P., & Subrahmanyam, M. (2009). On the volatility and comovement of US financial markets around macroeconomic news announcements. Journal of Financial and Quantitative Analysis, 44(6), 1265-1289.
Chung, S. L., Hung, C. H., & Yeh, C. Y. (2012). When does investor sentiment predict stock returns?. Journal of Empirical Finance, 19(2), 217-240.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. psychometrika, 16(3), 297-334.
Czerwonka, M. (2017). Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities. International Journal of Management and Economics, 53(3), 48-66.
Davis, J. H., Schoorman, F. D., Donaldson, L. (1997). Davis, Schoorman, and Donaldson reply: The distinctiveness of agency theory and stewardship theory, ss.611-613.
Dean, J. W., & Perlman, M. (1998). Harvey Leibenstein as a pioneer of our time.
Fama, E. F. (1971). "Risk, Return, and Equilibrium". Journal of Political Economy, 79(1), 30-55.
Fama, E. F. (1980). Agency Problems and the Theory of the Firm. The journal of political economy, ss.288-307.
Fornalczyk A., (2007) Biznes a ochrona konkurencji. , Oficyna a Wolters Kluwer business, Kraków
Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. Unintended thought, 26, 481.
Heath, C., & Tversky, A. (1991). "Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty". Journal of Risk and Uncertainty, 4(1), 5-28.

Higgins, E. Tory. “Knowledge Activation: Accessibility, Applicability, and Salience,” in E. Tory Higgins and Arie W. Kruglanski, eds., Social psychology: Handbook of basic principles. New York: Guilford Press, 1996, pp. 133–68.
Jensen, M. C., Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of financial economics, 3(4), ss.305-360.
Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), Rational animals? (pp. 87-106). New York, NY, US: Oxford University Press.
Khan, A. (2018). A Behavioral Approach to Financial Supervision, Regulation, and Central Banking. International Monetary Fund.
Kahneman, D. (2003) "Maps of Bounded Rationality: Psychology for Behavioral Eco-nomics." American Economic Review, 93:5, 1449-75.
Kahneman, D., & Egan, P. (2011). “Thinking, Fast and Slow” (Vol. 1). New York: Farrar, Straus and Giroux.
Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological review, 80(4), 237.
Kahneman, D. Tversky A. (1979). “Prospect Theory: An Analysis of Decisions Under Risk”. Econometrica, 47, 278.
Kahneman, D., & Tversky, A. (1986). Rational choice and the framing of decisions. Journal of business, 59(4), 251-278.
Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. Heuristics and biases: The psychology of intuitive judgment, 49, 81.
Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. Management science, 39(1), 17-31.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. Journal of political Economy, 98(6), 1325-1348.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. Journal of Economic perspectives, 5(1), 193-206.
Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). Well-being: Foundations of hedonic psychology. Russell Sage Foundation.
Klein, G. Sources of power: How people make decisions. Cambridge, MA: MIT Press, 1998.
MacIntyre, A. C., & Macintyre, A. C. (1988). Whose justice? Which rationality? (Vol. 239). London: Duckworth.
Orphanides, A., & Williams, J. C. (2005). The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations. Journal of Economic dynamics and control, 29(11), 1927-1950.
Penc J., Nowoczesne kierowanie ludźmi. Wywieranie wpływu i współdziałanie w organizacji, Difin, Warszawa 2007, s. 89
Ross, S. A. (1973). The economic theory of agency: The principal's problem. The American Economic Review, ss.134-139.
Segal, U. (1987). “The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach”. International Economic Review, 175-202.
Shefrin, H., & Statman, M. (1985). “The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence”. The Journal of Finance, 40(3), 777-790.
Shefrin, H., & Statman, M. (1994). “Behavioral Capital Asset Pricing Theory”. Journal of Financial and Quantitative Analysis, 29(03), 323-349.
Shefrin, H., & Statman, M. (2000). “Behavioral Portfolio Theory”. Journal of Financial and Quantitative Analysis, 35(02), 127-151.
Shefrin, H. (2007). “How The Disposition Effect and Momentum Impact Investment Professionals”. Journal of Investment Consulting, 8(2), 68-79.
Shleifer, A. (2000). Inefficient markets: An introduction to behavioural finance. OUP Oxford.

Shleifer, A., Vishny, R. W. (1997). A survey of corporate governance. The journal of finance, 52(2), ss.737-783,
Simon, H. A. (1986). Rationality in psychology and economics. Journal of Business, S209-S224.
Szapiro T. (1993), , Co decyduje o decyzji. , Wydawnictwo Naukowe PWN, Warszawa
Szyszka, A. (2009). „Finanse behawioralne: nowe podejście do inwestowania na rynku kapitałowym" (Behavioral Finance: A New Approach to Investing in The Capital Market). Wydawnictwo Uniwersytetu Ekonomicznego.
Teixeira, L. A., & De Oliveira, A. L. I. (2010). A method for automatic stock trading combining technical analysis and nearest neighbor classification. Expert systems with applications, 37(10), 6885-6890.
Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), 39-60.
Tversky, A., & Kahneman, D. (1983). "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment". Psychological Review, 90(4), 293.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. science, 185(4157), 1124-1131.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty, 5(4), 297-323.
Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. The quarterly journal of economics, 106(4), 1039-1061.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453-458.

DECISION SCIENCES INSTITUTE

Personality Types as Performance Predictors in Learning a Complex Enterprise Resource Planning (ERP) System

Bih-Ru Lea
Missouri University of Science & Technology
Email: leabi@mst.edu

Katherine Yu
University of Missouri - Columbia
Email: kayn9h@mail.missouri.edu

Dinesh Mirchandani
University of Missouri – St. Louis
Email: mirchandanid@msx.umsl.edu

Mary Sumner
University of Missouri – St. Louis
Email: sumnermb@umsl.edu

Nhu Thai
Missouri University of Science & Technology
Email: ndtkxf@mst.edu

ABSTRACT

This research explores the influence of personality types on ERP system learning performance. The four personality types are determined using cluster analysis based on Big Five personality traits: conscientiousness, neuroticism, extraversion, openness, and agreeableness. The results suggest that prior computer experience, learning motivation, demographic factors (gender, age, year in school), and social culture influence moderate relationships between personality types and ERP learning performance.

KEYWORDS: Personality traits, Information Systems, ERP, Learning Performance

INTRODUCTION

Enterprise Resource Planning (ERP) systems are large-scale, complex, heavily procedural-, process-, and detail-oriented computer systems (Kumar &Van Hillegersberg, 2000; Shtub, 2001; Tenhiala & Helkio, 2015). Due to their intricate nature, it can be challenging for beginners to familiarize themselves with and become adept at utilizing ERP systems. The lack of ERP-relevant technical, analytical, and business process management skills in IT professionals and end-users can hinder the adoption and successful use of such systems (Sumner, et al., 2006). As a result, the need for a skilled ERP workforce put pressure on universities to educate students with ERP knowledge (Charland, et. al., 2015).

Finding an efficient mechanism for universities to educate students with ERP knowledge and for various companies to recruit talented and qualified workers has presented itself as a difficult task. Although a multitude of factors contributes to the success of students learning ERP systems, the ability to identify factors with the highest correlation to success, or particular clusters of these factors could be a significant advancement. As personality traits are perceived to be correlated with job performance, companies often utilize personality assessment in their quest for skilled IT workers (Lui, et. al., 2008). This study aims to determine to determine if there is an ideal personality type (a combination of the Big Five personality traits) that correlate to the success of learning ERP systems.

LITERATURE REVIEW

Challenges in Complex Information System Education

ERP system is “an integrated system where a unique database provides flow for information continuously and consistently for the entire company” (Wadate, 2014; Klaus & Gable, 2000) and becomes a strategic tool that helps an organization to gain a competitive edge through integrating all business processes and optimizing the available resources available. As an ERP system integrates all information systems across an organization, it creates an interconnection between functions and various processes (Barki & Pinsonneault 2002) and that contributes to the complexity in learning and operating the system.

The widespread implementation of ERP systems over the last two decades has led to the challenge for companies to recruit a skilled ERP workforce and for universities to educate graduates with ERP knowledge (Charland, et al., 2015). As most IT failures stem from a lack of user acceptance rather than poor technical quality (Nelson and Cheney, 1987), the successful implementation of ERP systems requires effective, reinforced, and updated training of personnel (Markus et al., 2000; Gargeya and Brady, 2005). Because of the complexity of an ERP system, the learning curve is steep and has become a challenge for both educators and learners. Furthermore, the steep learning curve and heavy infrastructure investment needed by the educational institutes result in few academics who have intimate knowledge of the strengths and weaknesses of ERP (Becarra-Sanchez, et. al., 2002) and that had limited the ability of the U.S. universities in addressing this industry critical need. Such a shortage had led to IT outsourcing that hinders our nation’s competitiveness and economy.

Léger (2006) and Seethamraju (2011) believe that the challenge for students to learn the business process and ERP software in the classroom is the lack of knowledge and understanding about real-world business processes as well as the limited IT skills to operate an ERP software application. To be successful in learning an EPR system, Boyle and Strong (2006) report that a student needs to possess not only technical skills such as operating system, system analysis or data management, but also technology management knowledge, business functional knowledge, interpersonal skills and teamwork skills. Researchers have observed correlations between personality traits and one’s interpersonal skills and teamwork skills (Riggio, 1986; Wodarski, et al., 1988; Neuman & Wright, 1999). As a result, personality traits and assessment are often utilized by companies in their quest for skilled IT workers (Lui, et. al., 2008).

Personality Traits and Academic Learning Performance

Since the late twentieth century, psychologists have used the Big Five Trait model to assess and describe people's unique personality styles. The traits included are conscientiousness, neuroticism, extraversion, openness, and agreeableness. Each trait is presented as a spectrum, on which subjects can score moderately or to either extreme. Conscientiousness is a measure of how organized, disciplined, and achievement-oriented a person is (Komarraju, et. al., 2011). People who are highly conscientious have good self-control, determination, purposefulness, dependability, and the will to achieve (Grehan, et al., 2011). Neuroticism is known as the degree of emotional stability, anxiety, and impulse control a person has. Extraversion refers to a person's degree of sociability, talkativeness, and assertiveness (Komarraju, et. al., 2011). People who are highly extraverted will have a tendency to be cheerful, assertive, and social (Grehan, et al.,, 2011). Openness is the measure of a person's intellectual curiosity and preference for novelty and variety. When a person is highly open they are receptive to new ideas, attentive to inner feelings, and prefer varied sensations (Grehan, et al., 2011). The last trait, agreeableness, is characterized by how helpful, cooperative, and sympathetic a person is.

Meta-analyses have shown a connection between personality traits and academic success. For example, certain personality traits have been linked to learning styles such as deep and shallow processing. Individuals who are highly conscientiousness, open, extraverted, and emotionally stable tend to process information with the goal of mastery through deep understanding and long-term retention as opposed to the shallow goal of performing well on assessments. (Furnham, 1996; Komarraju et. al., 2011). Conscientiousness, openness, and agreeableness each show a positive relationship to academic success (Paunonen & Ashton, 2001; Laidra et. al., 2007) with conscientiousness having the highest relation and being the strongest predictor (Poropat, 2009). Meanwhile, neuroticism and extraversion were negatively related to success in academics (Chamorro-Premuzic & Furnham, 2003; Nguyen et. al., 2005; O'Connor & Paunonen, 2007; De Raad & Schouwenburg, 1996). Agreeableness and openness were also shown to have an impact on fostering collaborative learning (Komarraju et. al., 2011).

Although personality traits & academic performance are clearly linked, learning styles and academic environment may complicate how these connections manifest. When in specific academic settings, the impact of certain traits on academic success varies. In learning environments and contexts that require knowledge acquisition, conscientiousness was shown to be the best predictor of academic success. However, in graduate settings that tend to emphasize skill acquisition and application, neuroticism was shown to be linked to increased performance (Ferguson, 2014).

Researchers further classify Big Five personality traits into higher-order factors such as stability (Neuroticism, Agreeableness, Conscientiousness) and plasticity (Extraversion, Openness) (Buruk, et al., 2017; Davies, 2014; DeYoung, 2006; DeYoung, et al., 2002). As personality traits often exist as dynamic systems of several personality traits for an individual (Specht and Luhmann, 2014), personality types (or combinations of personality traits) can provide a more comprehensive and detailed picture of an individual's character than do the individual traits (Specht and Luhmann, 2014; Tyseen, et al., 2007). Cluster analysis is often used to derive personality types (Specht and Luhmann, 2014) and common personality types are resilients, overcontrollers, and undercontrollers (Gerlach, et al., 2014; Donnellan and Robins, 2010). Resilients are individuals with high scores on all of the Big Five personality traits and are comparatively emotionally stable, extraverted, open to experience, agreeable, and conscientious. Undercontrollers are individuals with particularly low on conscientiousness and agreeableness. Overcontrollers are those who with particularly introverted, neurotic, and less open (Donnellan and Robins, 2010).

Prior Computer Experience, Demographic factors, Motivation, and Culture/Social Influence

Academic motivation comes in different forms and scales and is one of important predictors that has significant impact on academic achievement (Fairchild, et al., 2005). Research has also provided empirical evidence that experience is a predictor of performance (Avolio, et. al., 1990; Schmidt & Hunter, 1998), but that the relationship between these two measures is moderated by length of experience and job complexity (Mcdaniel, et. al. 1998). Gender (Wawrzynski, 2003; Li, et. al., 2015), prior knowledge (Avolio, et. al., 1990; Mcdaniel, et. al, 1998), academic performance (e.g., GPA), academic experience (Cheung & Kan, 2002; Wawrzynski, 2003), and cultural and social influences (Wawrzynski, 2003; Fatemifar, et. al.,2015) have also shown influence in IT education.

RESEARCH QUESTIONS, HYPOTHESES, AND METHODOLOGY

The objective this research is explore if there is an ideal personality type (i.e., a combination of the Big Five personality traits) that predict success in learning ERP systems and concepts. Prior computer Experience, culture or social influence, learning motivation, age, gender, and year in school are included as moderating variables in this study. Prior computer experience is measured by three participant self-reported measures including (1) years of information system (IS) experience outside of classroom, (2) number of ERP classes completed previously, and (3) the computer savvy level. The degree of excitement on the selected course, on IT in general, and on using an ERP system are used to represent learning motivation with a reliability score of Cronbach's Alpha 0.838. Culture and social influence are represented by participants' school location, race, and English proficiency. Participant demographic characteristics are represented by age, gender, and year in school.

H1. Personality types have influence on ERP learning performance.

H2. Prior Computer Experience moderates the relationships between personality types and ERP learning performance

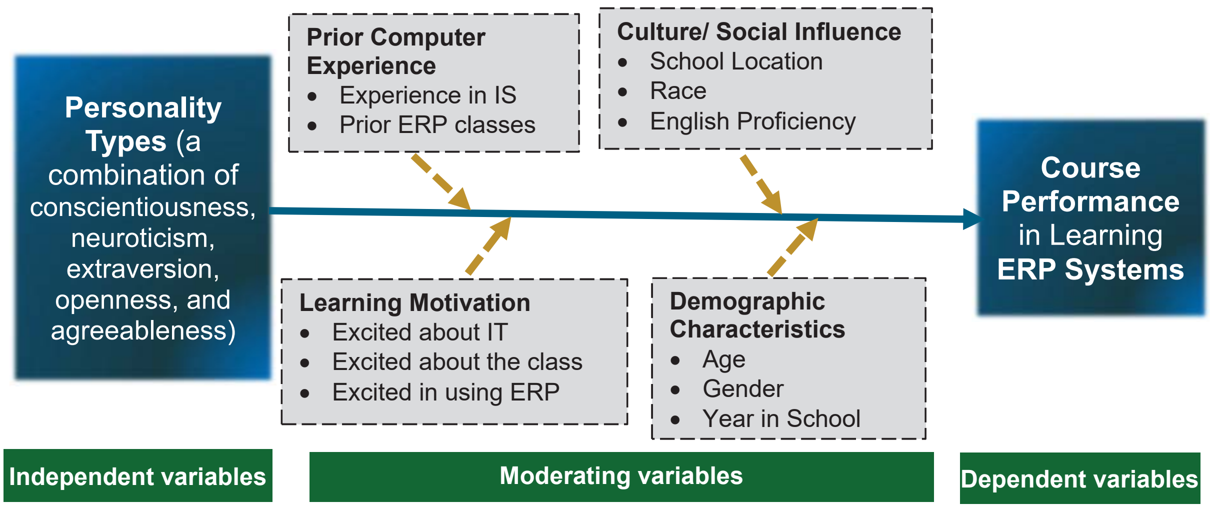
H3. Learning motivation moderates the relationships between personality types and ERP learning performance

H4. Culture and Social Influence moderates the relationships between personality types and ERP learning performance

H5. Demographic characteristics moderate the relationship between personality types and ERP learning performance

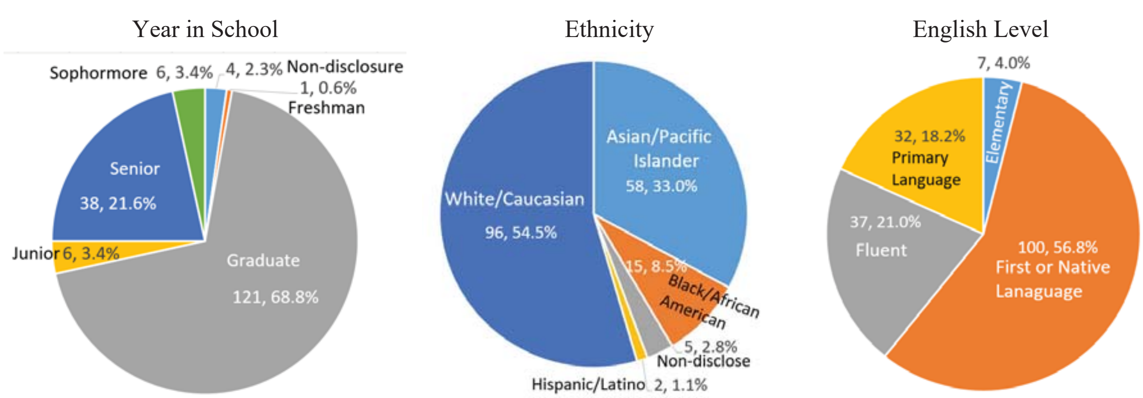
The success in learning ERP concepts and concepts is measured by course grade. The personality types are derived from cluster analysis based on the Big Five personality traits of conscientiousness, neuroticism, extraversion, openness, and agreeableness. The Mini-IPIP. This twenty-item assessment is an abbreviated version of the original International Personality Item Pool (IPIP) which contains only four questions to assess each trait. Studies conducted to test the reliability and validity of the Mini-IPIP have confirmed both; the Mini-IPIP is recommended as an efficient yet effective personality assessment tool (Donnellan, et al., 2006). Each item in the 20-item questionnaire were rated on a 5-point likert scale from 1 is Very Inaccurate to 5 is Very Accurate. The assessment also includes demographic questions about age, gender, year in school, self-reported technology level, ethnicity, and other relevant data. The research model is summarized in Figure 1.

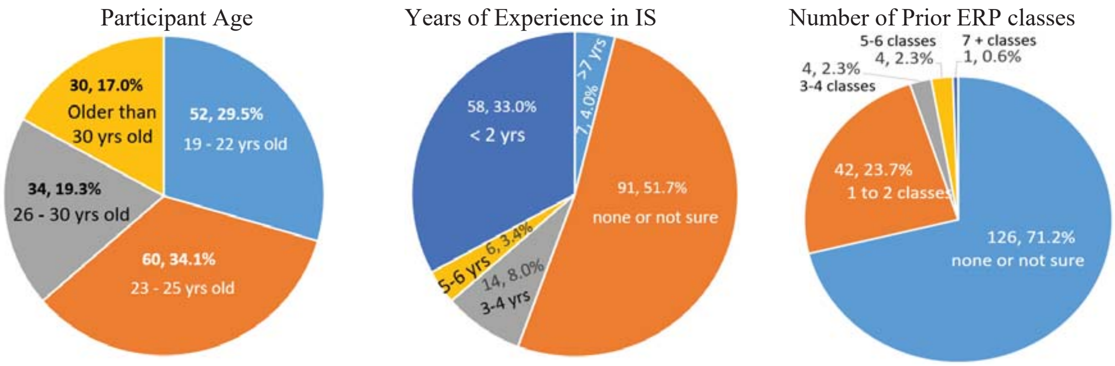
Figure 1 Research Model



Data were collected from two universities located in a rural university town and in a metropolitan area in the Midwest of the United States over three semesters: Fall 2016, Spring 2017 and Fall 2018. Among a total of 176 students with valid responses, female, male, and no disclosure students are 38.1%, 59.7%, and 2.3% respectively. Participants enrolled in the rural area university and the metropolitan area university accounts for 48.9% and 51.1% respectively. Graduate students, undergraduate, non-disclosure students are 68.8%, 28.9%, and 2.3% respectively. About 54.5% of participants are white or Caucasian, 33% are Asian or Pacific Islander, 8.5% are Black or African American, 1.1% is Hispanic or Latino, and about 2.9% are other or non-disclosure. Participants indicate that English is their first or native language and primary language is 56.8% and 18.2% respectively. About 21% of participants speak English fluently and about 4% speak at the elementary level. Figure 2 provides a summary of participant information.

Figure 2. Year in School, Ethnicity, and English Ability Information among Participants





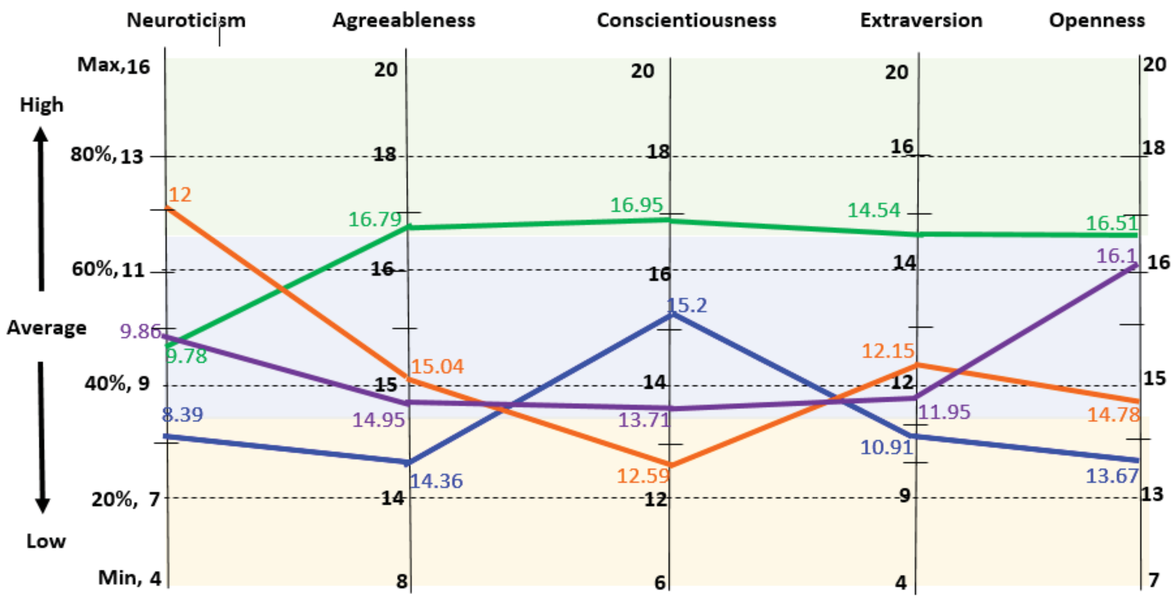
RESULTS AND OBSERVATIONS

Cluster analysis was conducted with the number of clusters ranged from three to ten. A four-cluster is selected based on the Elbow methods and the cluster centers, the distance between clusters, and the number of cases in each cluster. As shown in Figure 3, the following general description is derived for each personality type:

- Personality type one scores low on agreeableness, extraversion, openness, and neuroticism which means they are not cooperative, withdrawn, play it safe, and are emotionally stable. Their conscientiousness score is relatively high, meaning they are relatively disciplined and studious.
- Personality type two scores high on agreeableness, conscientiousness, extraversion, and openness which means that they are highly cooperative, organized, outgoing, and willing to take risks in trying new things. They score average on neuroticism which means their emotional stability is typical.
- Personality type three scores relatively low on agreeableness, conscientiousness, extraversion, and openness meaning they are not cooperative, studious, outgoing, or curious. Their neuroticism scores are comparatively high meaning they are very anxious and emotionally unstable.
- Personality type four scores relatively low on agreeableness, conscientiousness, and extraversion meaning they do not care much about group cohesion, are not very organized and are reserved. The relatively high score on openness indicates that they are imaginative and creative and the average score on neuroticism indicates that they have a typical level of emotional stability.

Generalized Linear Model analysis was conducted and results are summarized in Table 1. The influence of personality types (clusters) on ERP learning performance is significant at the α = 0.05 level. The moderating effects of English proficiency, race, years of IS Experience, degree of computer savvy, gender, age, school level, excited about the class, excited about using ERP system, and excited about IT on ERP learning performance are statistically significant at α = 0.05 level. The moderating effects of school location and number of prior ERP classes on relationships between personality types and ERP learning performance are not statistically significant at α = 0.10 level.

Figure 3 Parallel Diagram of Cluster Average



Personality Types (number of cases)	Neuroticism	Agreeableness	Conscientiousness	Extraversion	Openness
1 (46)	8.39 (Low)	14.30 (Low)	15.20 (Average)	10.91 (Low)	13.67 (Low)
2 (63)	9.78 (Average)	16.79 (High)	16.95 (High)	14.54 (High)	16.51 (High)
3 (46)	12.00 (High)	15.04 (Average-**Low**)	12.59 (Low)	12.15(Average)	14.78 (Average - **Low**)
4 (21)	9.86 (Average)	14.95 (Average - **Low**)	13.71 (Average - **Low**)	11.95 (Average - **Low**)	16.10 (Average - **High**)

Table 1 Tests of Model Effects

	Source	Type III Wald Chi-Square	df	Sig.
	(Intercept)	78.642	1	0.000
	Personality Types/Clusters	14.758	3	**0.002***
Prior Experience	Personality Types/Clusters * Years of IS Experience	95.709	12	**0.000***
	Personality Types/Clusters * Prior ERP class	3.558	7	0.829
	Personality Types/Clusters * Computer Savvy	56.558	11	**0.000***
Demographic Characteristics	Personality Types/Clusters * Gender	19.790	4	**0.001***
	Personality Types/Clusters * age	28.811	9	**0.001***
	Personality Types/Clusters * School Level	56.215	8	**0.000***
Culture / Social Influence	Personality Types/Clusters * English	21.597	9	**0.010***
	Personality Types/Clusters * Race	94.130	10	**0.000***
	Personality Types/Clusters * English * Race	35.073	7	**0.000***
	Personality Types/Clusters * Location type	3.728	2	0.155
Learning Motivation	Personality Types/Clusters * Excited about class	27.818	8	**0.001***
	Personality Types/Clusters * Excited about ERP	44.302	8	**0.000***
	Personality Types/Clusters * Excited about IT	67.280	8	**0.000***

* significant at α = 0.05 level

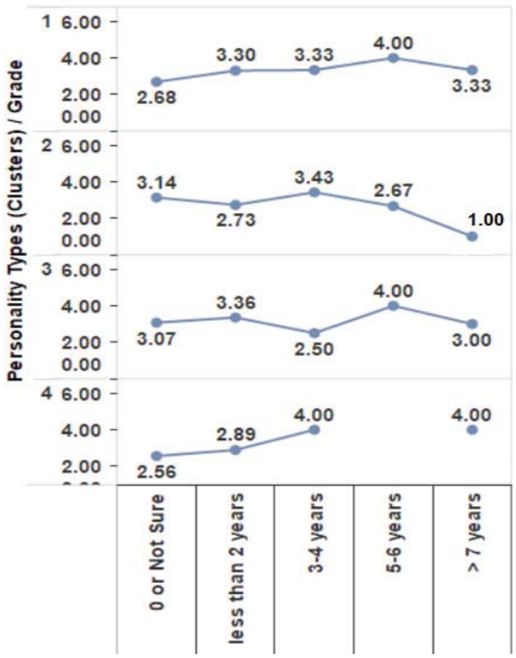
Prior Computer Experience and Personality Types on ERP learning performance

As shown in Table 1, prior computer experience represented by years of IS experience, number of prior ERP classes, and the self-reported level of computer savvy significantly moderate relationships between personality types and ERP learning performance. As shown in figure 4a, the learning performance decrease as participants' prior IS experience increase for type two personality participants. For participants with the other three personality types, participants with some level of prior experience in information system (IS) had better learning performance than those who do not have any experience. However, the impact of prior IS experience becomes negative for participants with more than seven years' experience. Although more analysis and research are needed to identify the root causes, it is possible that those participants with longer IS experience may have been to a management position and their IS experiences are more relevant to managerial and analytical and less relevant to the procedural and detailed oriented ERP tasks. This finding can also result from the complexity of ERP system as reported by McDaniel, et al. (1998) that that the length of experience and job complexity moderate experience and performance (McDaniel, et al., 1998)

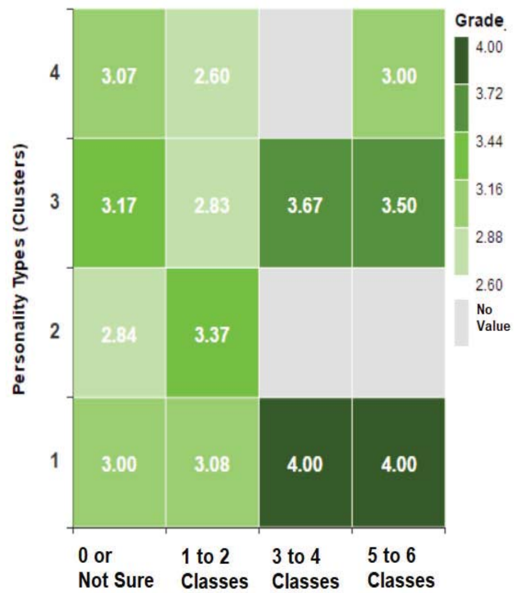
As shown in Figure 4b, learning performance improve as the number of prior ERP classes increase regardless of personality types. However, the performance seems to not differ significantly after having three or more prior ERP classes. One interesting observations from Figure 4b is that participants who ranked themselves as least computer savvy (strongly disagree) had the strongest performance. Participants with personality type two and are not computer savvy (strongly disagree or disagree) performed better than those who ranked them as computer savvy (agree and strongly agree). On the other hand, participants ranked themselves as not computer savvy (disagree) had very different performance level. Participants are not computer savvy (disagree) with personality types one and four had the lowest performance level while those with personality types two and three had strong performance.

Figure 4 Years of IS Experience, Number of Prior ERP Classes, and Computer Savvy Level

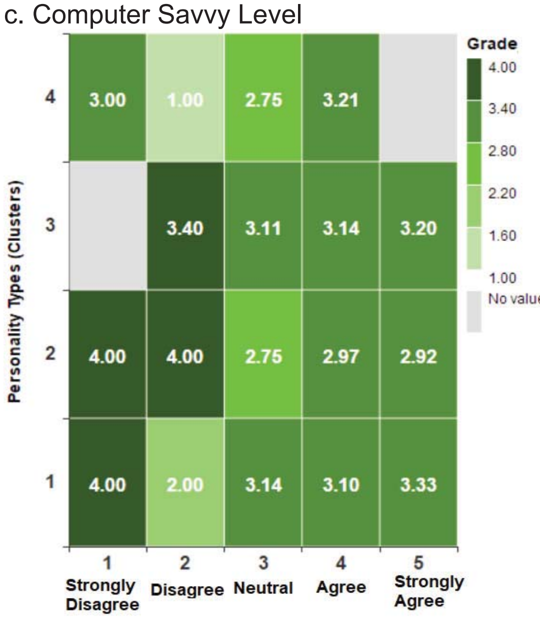
a. Years of IS Experience



b. Number of Prior ERP Classes



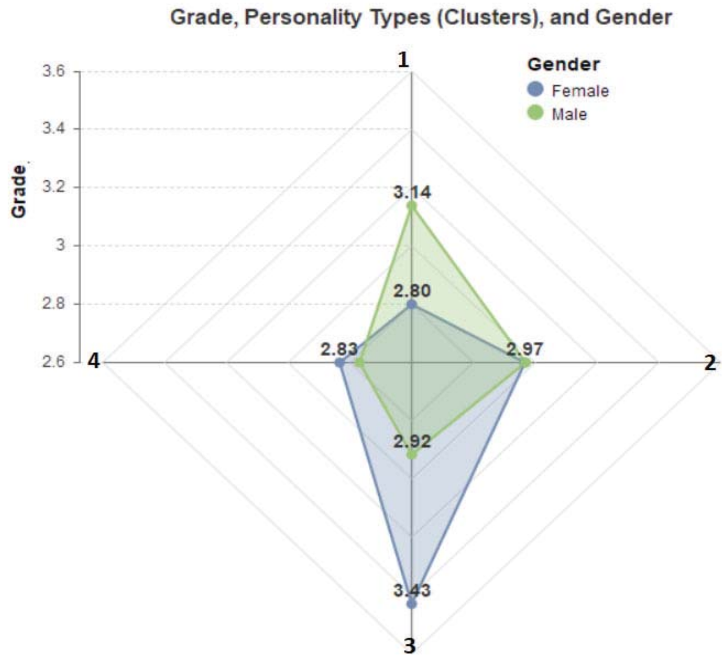
c. Computer Savvy Level



Demographic Characteristics and Personality Types on ERP Learning Performance

The results indicate that demographic characteristics of gender, year in school, and age moderate relationships between personality types and ERP learning performance. As shown in Figure 5, except personality type one participants, female participants typically performed better than male participants. An interesting observation is the performance between male and female participants are completely opposite. For type one participants, female participants performed significantly worse than male participants. On the other hand, female participants performed significantly better than male participants for those who with a personality type three. Although additional studies are needed to identify causes, it is observed that closely aligned personality type and gender characteristic result in better performance. Comparing with type three participants, personality type one participants have relatively low neuroticism, agreeableness, extraversion, and openness, but exhibit a higher level of conscientiousness, so they are more emotionally stable, more studious, more cooperative, less outgoing, and less open to new ideas and are more closely aligned with a typical male profile. As a result, male participants outperformed female participants. Type three participants are more emotionally unstable, more cooperative, more outgoing, more willing to try new things, and less studious, and those are more aligned with the characteristics of a female. As a result, female participants thrive more than male participants.

Figure 5 Gender, Personality Types, and ERP Learning Performance

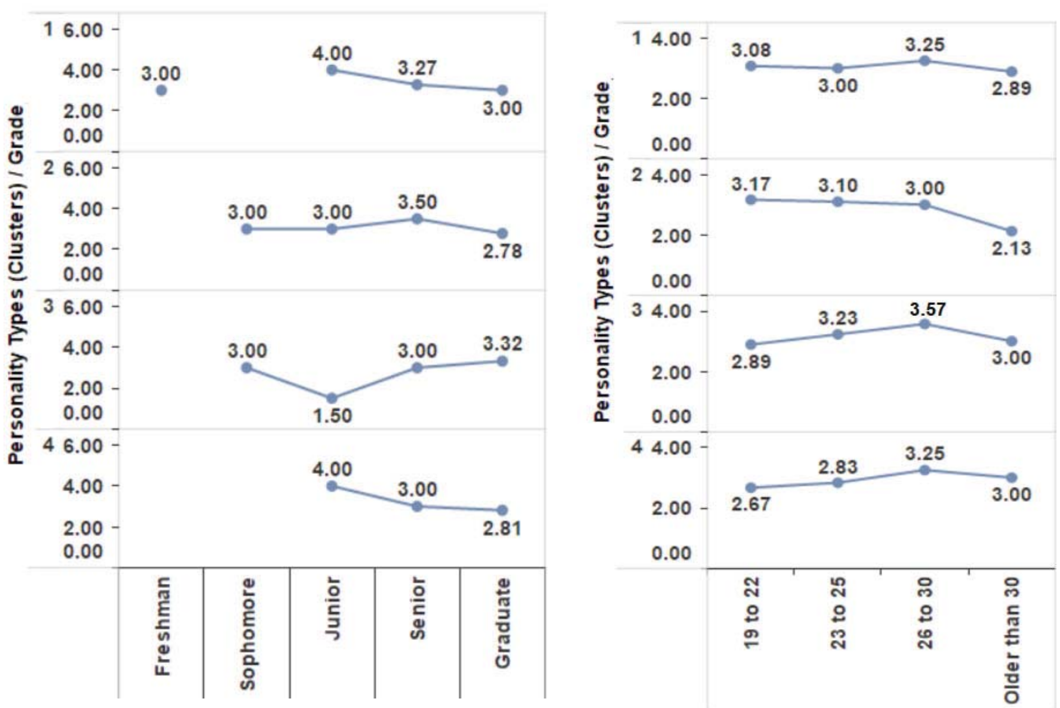


For personality type one and type four participants, results indicate that ERP learning performance decrease as participants advance through their academic majors and that senior students outperform graduate students, as shown in Figure 6a, For type two participants, it is also observed that senior students tend to outperform other groups of students. Other than personality type three, graduate students had lower performance than undergraduate students for the other three personality types. Boyle and Strong (2006) report that a student needs to possess not only technical skills such as operating system, system analysis or data management but also technology management knowledge, business knowledge, problem-solving, analytical and team skills to be successful in learning ERP systems. As the curriculum for the participating undergraduate students require both business core classes (accounting, finance, marketing, management, business capstone) and information technology core classes (two programming courses, Management Information System, digital media, ERP Introduction), undergraduate students are more likely to be successful in learning ERP than graduate students who specialized in either information technologies or business administration (MBA students).

For personality type two participants who exhibit high level of agreeableness, conscientiousness, extraversion, and openness, it is observed that younger participants had better ERP learning performance than older participants and the performance decrease as the age of participants increase, as shown in Figure 6b. For participants younger than 30 years old with a personality type of one, three, and four, learning performance seems to increase as participants' age increase. Participants who are older than 30 years old perform worse than other age groups for all four personality types.

Figure 6 Year in School, Age, Personality Types, and ERP Learning Performance

a. Year in School b. Age



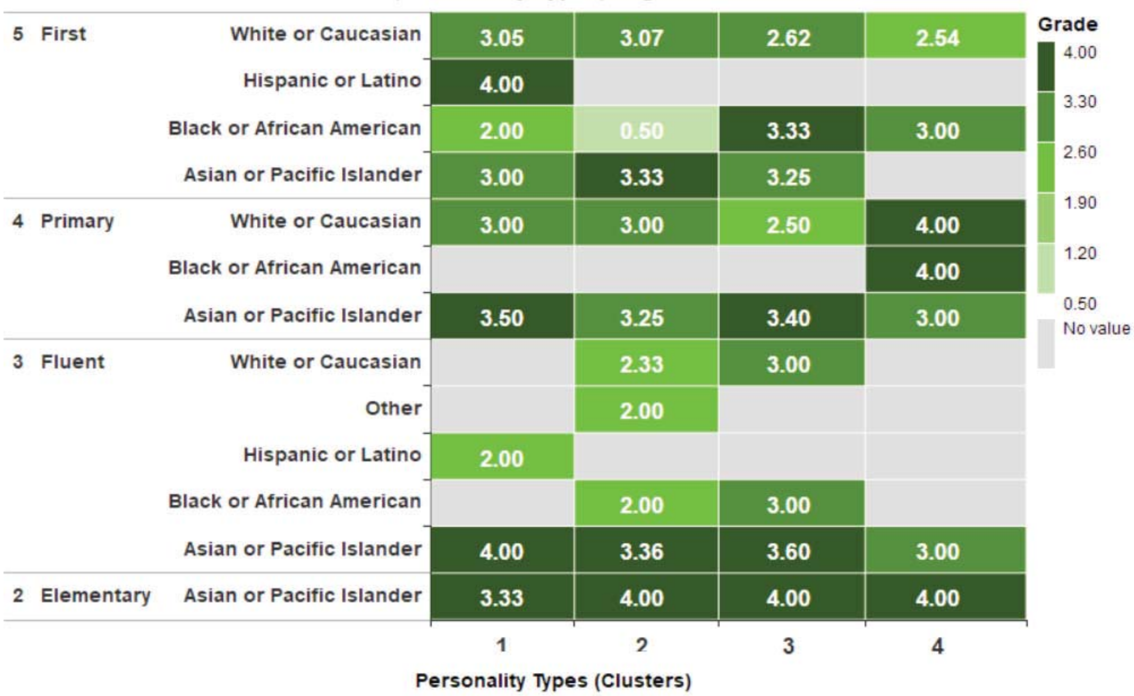
Culture and Social Influence and Personality Types on ERP Learning Performance

As shown in Figure 7, it is observed that participants who indicate English is their first or native language have lower ERP learning performance than those with lesser English proficiency. Participants with the least English proficiency (i.e., elementary level) outperform participants who have better English proficiency. In general, ERP learning performance decrease as English proficiency increase regardless of personality types and ethnicity.

Further analysis reveals that participants with Asian or Pacific Islander origin tend to self-report lower English proficiency, but their performance is significantly better than other ethnicities regardless of personality types, as shown in Figure 8. Although additional studies are needed to identify causes, Asian or Pacific Islander origin students tend to exhibit stronger mathematics, reasoning, and memorization skills. As an ERP system is complex and highly procedural, stronger mathematics, reasoning, and memorization skills often proof beneficial in operating or troubleshooting the system.

Figure 7 Influence of English Proficiency and Personality Types

Figure 8 Influence of Personality Types, English Proficiency, and Race on Course Grade



Learning Motivation and Personality Types on ERP Learning Performance

Participants who reported a positive excitement level (agree or strongly agree) perform better than participants who had a negative excitement level (disagree or strongly disagree), as shown in Figure 9. Furthermore, results indicate that learning performance improves as participants' level of excitement increase regardless of personality types. However, ERP learning performance of personality type 2 participants is not consistent. For the "excited about the class" measure, those who responded "strongly disagree" perform better than those who with other responses. For the "excited about IT in general" measure, those who responded "disagree" had the lowest performance than participants who had other responses.

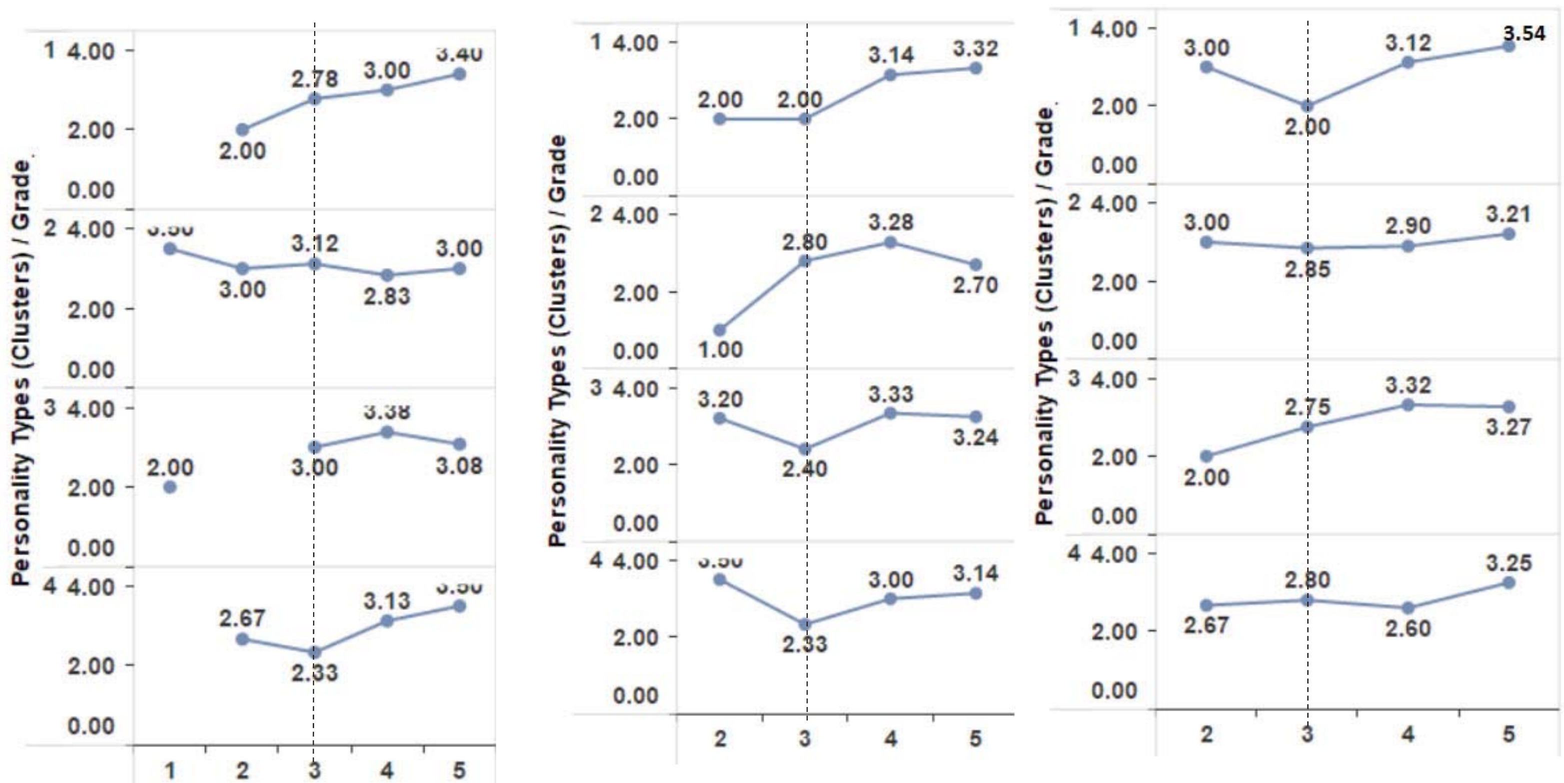
SUMMARY, LIMITATIONS, AND FUTURE RESEARCH DIRECTIONS

As personality types can provide a more comprehensive and detailed picture of an individual's character than individual personality traits (Specht and Luhmann, 2014), this study identifies four different personality types derived from Big Five personality traits (i.e., conscientiousness, neuroticism, extraversion, openness, and agreeableness) through cluster analysis. The results suggest that participants' prior computer experience, demographic factors of age, gender, and year in school, culture and social influence, and learning motivation moderate the relationships between personality types and ERP learning performance.

Only college students were included in this study is one of limitations. Studies that include additional participant groups such industry employees or high school students are invaluable in generalize results. Although this study presented many interesting observations, additional studies need to be conducted to identify causes, to derive theories, and to create decision models to those observations are invaluable.

Figure 9 Learning Motivation, Personality Type, and ERP Learning Performance

a. Excited About the Class b. Excited About IT in General c. Excited About Using ERP



1: Strongly Disagree 2: Disagree 3: Neither Disagree nor Agree 4: Agree 5: Strongly Agree

REFERENCES

Boyle, T. A., & Strong, S. E. (2006). Skill requirements of ERP graduates. *Journal of Information Systems Education*, 17(4), 403.

Barki, H., & Pinsonneault, A. (2002). Explaining ERP implementation effort and benefits with organizational integration. *Cahier du GReSI no*, *2*(01).

Buruk, P., Şimşek, Ö. F., & Kocayörük, E. (2017). Higher-order Traits and Happiness in the Workplace: The Importance of Occupational Project Scale for the Evaluation of Characteristic Adaptations. *The Journal of general psychology*, 144(4), 245-263.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of research in personality*, *37*(4), 319-338.

Charland, P., Leger, P., Cronan, T. P., & Robert, J. (2015). Developing and Assessing ERP Competencies: Basic and Complex Knowledge. Journal of Computer Information Systems, 56(1), 31-39.

Davies, S. E. (2014). Lower and higher order facets and factors of the interpersonal traits among the Big Five: Specifying, measuring, and understanding extraversion and agreeableness.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of personality*, *10*(5), 303-336.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of personality and social psychology*, 91(6), 1138.

DeYoung, C. G. (2015). Cybernetic big five theory. Journal of Research in Personality, 56, 33-58.

DeYoung, C. G., & Krueger, R. F. (2018). A cybernetic theory of psychopathology. Psychological Inquiry, 29(3), 117-138.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health?. *Personality and Individual differences*, 33(4), 533-552.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*, *18*(2), 192.

Donnellan, M. B., & Robins, R. W. (2010). Resilient, overcontrolled, and undercontrolled personality types: Issues and controversies*. Social and Personality Psychology Compass*, 4(11), 1070-1083.

Fairchild, A. J., Horst, S. J., Finney, S. J., & Barron, K. E. (2005). Evaluating existing and new validity evidence for the Academic Motivation Scale. Contemporary Educational Psychology, 30(3), 331-358.

Ferguson, E., Semper, H., Yates, J., Fitzgerald, J. E., Skatova, A., & James, D. (2014, 02). The 'Dark Side' and 'Bright Side' of Personality: When Too Much Conscientiousness and Too Little Anxiety Are Detrimental with Respect to the Acquisition of Medical Knowledge and Skill. *PLoS ONE, 9*(2).

Furnham, A. (1996). The FIRO-B, the learning style questionnaire, and the five-factor model. *Journal of Social Behavior and Personality*, 11(2), 285.

Gargeya, V. B., & Brady, C. (2005). Success and failure factors of adopting SAP in ERP system implementation. *Business process management journal*, *11*(5), 501-516.

Gerlach, M., Farb, B., Revelle, W., & Amaral, L. A. N. (2018). A robust data-driven approach identifies four personality types across four large data sets*. Nature Human Behaviour*, 2(10), 735.

Grehan, P. M., Flanagan, R., & Malgady, R. G. (2011). Successful graduate students: The roles of personality traits and emotional intelligence. *Psychology in the Schools*, *48*(4), 317-331.

Klaus, H., Rosemann, M., & Gable, G. G. (2000). What is ERP?. *Information systems frontiers*, *2*(2), 141-162.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and individual differences*, *51*(4), 472-477.

Kumar, K., & Van Hillegersberg, J. (2000). ERP experiences and evolution. *Communications of the ACM*, 43(4), 22-22

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and individual differences*, *42*(3), 441-451.

Léger, P.-M. (2006). Using a Simulation Game to Teach Enterprise Resource Planning Concepts. *Journal of Information Systems Education*, 17(4), 441-447.

Lui, K. M., & Chan, K. C. (2008). Rescuing troubled software projects by team transformation: A case study with an ERP project. *IEEE Transactions on Engineering Management*, *55*(1), 171-184.

Markus, M. L., Axline, S., Petrie, D., & Tanis, S. C. (2000). Learning from adopters' experiences with ERP: problems encountered and success achieved. *Journal of information technology*, *15*(4), 245-265.

Morgeson, F. P., Reider, M. H., & Campion, M. A. (2005). Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge. *Personnel psychology*, 58(3), 583-611.

Nguyen, N. T., Allen, L. C., & Fraccastoro, K. (2005). Personality predicts academic performance: Exploring the moderating role of gender. *Journal of Higher Education Policy and Management*, *27*(1), 105-117.

Nelson, R. R., & Cheney, P. H. (1987). Training end users: An exploratory study. *MIS quarterly*, *11*(4), 547-559.

Neuman, G. A., & Wright, J. (1999). Team effectiveness: beyond skills and cognitive ability. *Journal of Applied psychology*, 84(3), 376.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual differences*, *43*(5), 971-990.

Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior*. Journal of personality and social psychology*, 81(3), 524.

Pickering, A. D. (2004). The neuropsychology of impulsive antisocial sensation seeking personality traits: From dopamine to hippocampal function. On the psychobiology of personality: Essays in honor of Marvin Zuckerman, 453-476.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2), 322.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and social Psychology*, 51(3), 649.

Robinson, M. D., Moeller, S. K., & Ode, S. (2010). Extraversion and reward-related processing: Probing incentive motivation in affective priming tasks. Emotion, 10(5), 615.

Seethamraju, R. (2011). Enhancing Student Learning of Enterprise Integration and Business Process Orientation through an ERP Business Simulation Game. *Journal of Information Systems Education*, 22(1), 19-29.

Specht, J., Luhmann, M., & Geiser, C. (2014). On the consistency of personality types across adulthood: Latent profile analyses in two large-scale panel studies. Journal of personality and social psychology, 107(3), 540.

Shtub, A. (2001). A framework for teaching and training in the Enterprise Resource Planning (ERP) era. *International Journal of Production Research*, *39*(3), 567-576.

Sumner, M., Watson, E., and Corbitt G., (2006) "ERP Competencies: What Does Industry Need?" Proceedings of the Americas Conference on Information Systems (AMCIS).

Tenhiälä, A., & Helkiö, P. (2015). Performance effects of using an ERP system for manufacturing planning and control under dynamic market requirements. *Journal of Operations Management*, *36*, 147-164.

Tyssen, R., Dolatowski, F. C., Rovik, J. O., Thorkildsen, R. F., Ekeberg, O., Hem, E., ... & Vaglum, P. (2007). Personality traits and types predict medical school stress: a six-year longitudinal and nationwide study. *Medical education*, 41(8), 781-787.

Wadate, J. D. (2014). Enterprise Resource Planning (ERP) in Universities. *International Journal of Informative & Futuristic Research (IJIFR)–Volume*, *2*.

Wodarski, J. S., Pippin, J. A., & Daniels, M. (1988). The effects of graduate social work education on personality, values and interpersonal skills. *Journal of Social Work Education*, *24*(3), 266-277.

DECISION SCIENCES INSTITUTE

Pricing the Cloud: A QoS-based Optimal Design

Yang Lu
Old Dominion University
College of Charleston
Email: y1lu@odu.edu

ABSTRACT

Cloud computing has a major impact on the IT industry. How to price and allocate cloud resources to meet users' requirements is an important problem. This paper proposes a dynamic mechanism to price cloud services, which can work in complex environments such as distributed system and uncertain budget constraints. A directly relationship between QoS and price is established. The approach uses an optimization technique to estimate the potential transaction price in the distributed network. It can allocate cloud resources under uncertainties, where providers can optimize their revenues, and consumers can obtain the resources at a relatively low price.

KEYWORDS: QoS (quality of service), Reliability, Availability, Robust optimization auction, Optimal revenue, Budget constraint

INTRODUCTION

"Cloud computing: a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." (Mell & Grance, 2011) Cloud services include Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS) (Josep, et al., 2010). Cloud service can be treated as an ordinary commodity that has spread through the Internet. Currently, leading companies are gearing up to use cloud services for their businesses: Amazon's AWS (Amazon Web Services), Google's GAE (Google App Engine), Microsoft's Azure, and IBM's Cloud. For instance, Amazon Elastic Compute Cloud (Amazon EC2) offers seven instance purchasing options: On-demand Instances, Reserved Instances, Scheduled Instances, Spot Instances, Dedicated Hosts, Dedicated Instances, and Capacity Reservations. But they have not yet provided roust optimal design directly related to QoS metrics (reliability and availability). Based on the SLA (service level agreement), the overall performance of cloud service is guaranteed to some extent, but the relationship between price and QoS metrics still isn't clear. Usually, customers can purchase cloud service via various strategies, at certain prices which are determined by costs and by an economy proxy (supply and demand). As a special commodity, the cloud not only has costs that are similar to those of other goods, but it also has QoS-related costs that aren't easy to estimate. It is possible that customers bid on cloud services based on the QoS metrics that both provider and customer really care about.

Cloud computing provides us with on-demand and remote QoS-embedded services that are scalable, elastic, complex, and potentially decentralized (Zhang, Cheng, & Boutaba, 2010). Quality of service is crucial to the cloud industry. On the one hand, providers seek to improve the overall performance of cloud service to compete in the industry; on the other hand, customers expect cloud service to have high QoS standards, especially in its reliability and availability. The reliability is actually the probability that a system will be operational in a given time interval without any failures (Armbrust, et al., 2010). The availability is actually the probability that the system will be up and functional correctly in a certain time period (Armbrust, et al., 2010). High availability, with high costs and price, is essential to guarantee QoS, to maintain customer's confidence, and to attract more customers. More customers are involved in the use of cloud services, a lower price each customer needs to pay. In other words, it is the economies of scale. The percentile of multiple Nines and Fives is defined to express the different level of availability in general, e.g., 3 Nines is 99.90% and 4N5 is 99.995%. In our study, we will briefly illustrate the relationship between reliability and availability by MTBF (Mean Time between Failure) and MTTR (Mean Time to Repair). A function between availability and price then will be established as the reference for participants to estimate the price of cloud services.

One customer's requirement for a specific service is not independent of those of other customers. Indeed, a customer using a specific service has some connection with the needs of other customers. In an auction, cloud services are auctioned concurrently, and bidders can compete for services. Since QoS is important, both for the provider and the customer, there is the potential to add QoS metrics (availability) into the pricing algorithms. A mechanism allocates services, associated with the expected QoS, to customers.

Depending on the customer's different requirements for cloud services, suppliers need a viable and efficient pricing mechanism that is critical to the allocation and optimization of the available cloud resources. For suppliers, customers will bid for their own budget for better service, and the suppliers will select customers based on their bids, provide resources for higher-priced customers and guarantee the cloud's QoS. For customers, they can freely choose resources according to their own needs. However, due to the liquidity of marginal customers, whose bidding is uncertain, and it is difficult to estimate the demand based on existing bids. The cloud resource scheme based on uncertain bid auction has the potential to improve both economic efficiency and optimal profits.

Our design is a QoS-based robust optimization mechanism that includes multiple customers bidding for a variety of cloud services with QoS guarantees offered by different providers in a distributed network. The reasons why our pricing model combines QoS metrics with auction design are (1) cloud computing is a new and advanced computing technology. Cloud services are different from ordinary products or services, which can be evaluated relatively easily based on supply and demand. However, the cloud has unique QoS features, such as reliability and availability. An auction is a proper approach to evaluate not-easy-estimated parameters of cloud. (2) Cloud pricing is still in its early development stage (Zheng, et al., 2014). There is no direct reference to pricing cloud services based on QoS metrics. Hence, an auction algorithm is an appropriate way to estimate the valuation of certain cloud services and to allocate resources efficiently. (3) An auction design can sell services in a way that customers come to expect, specifically based upon Service Level Agreement (SLA) and Quality of Service (QoS). There are no obvious QoS-based pricing algorithms in the extant literature. The most likely exponential relationship expresses the intrinsic connection between price and QoS metrics and offers a straightforward pricing reference to both provider and customer.

In this paper, we address the issue of designing a robust optimal mechanism, through which cloud services are distributed between provider and customer, along with QoS. The goals are (1) to propose QoS-embedded availability to price and to allocate cloud services from providers to customers, and (2) to flexibly set up a dynamic pricing mechanism to enable budget constraint. Our contribution is to build a QoS-based dynamic auction model to effectively allocate cloud resources. Specifically, as a critical indicator of QoS, availability is used to estimate the value of cloud service in the pricing algorithms. Because of the unique features of the cloud, the relationship between price and QoS availability is not simply linear. We apply a robust optimal scheme to allocate resources and to optimize revenues among participants.

The paper has the following structure. A complete overview of the auction design for cloud computing (fixed pricing strategy, dynamic pricing strategy, and auction pricing strategy) and discussion of QoS-embedded optimal pricing strategy are addressed in Section 2. Section 3 depicts the relationship between QoS metrics (availability) and price. We describe and explain the detailed stages of functions and scheme in Section 4. Section 5 explores future directions. The final section summarizes the paper.

CLOUD PRICING MECHANISM

Cloud-related usage is become more and more popular in the daily life. Pricing is one of the important issues in the cloud industry. In this paper, we divided cloud pricing mechanisms into three categories: fixed pricing strategies, dynamic pricing strategies, and auction pricing strategies. Specifically, the three popular fixed pricing strategies are pay-as-you-go, subscription, and pay-for-use. Dynamic pricing strategies have many different algorithms, such as Genetic Model, Financial Optional Theory, Markov Decision Process, etc. Auction pricing strategies consist of various auction design, such as English auction, sealed-bid auction, double auction, and combinatorial auction. A summary of cloud pricing scheme is addressed in Table 1.

Fixed Pricing Strategy

Provider presents the price of cloud service, and customer pays the amount if the customer expects to use the service. The price is stable. Although the fixed pricing strategy is straightforward to both provider and customer, the mechanism is unfair to all potential customers who don't have the same needs (C. S. Yeoa, S. Venugopalb, X. Chua and R. Buyyaa, "Autonomic Metered Pricing for a Utility Computing Service", Future Generation Computer Syst., vol. 26, no. 8, (2010).). Even to provider, the mechanism is not an optimal strategy, especially when demand is higher than supply.
Pay-as-you-go, which charges customers based on their overall resource usage. The customers pay for the amount they consume of a product or the amount of time they use a certain service. Pay for resources, which charges customers based on the technical features of cloud services, such as the storage and bandwidth. Subscription, customers subscribes with certain service for a time period. Subscription is another type of fixed pricing, in which the customer pays a fixed amount of money to use the service for longer periods at any convenient time or amount. Provider presents the price of cloud service, and if a customer wishes to use the service, the customer pays the amount according to the agreement. In the cloud industry, three pricing mechanisms are widely employed, pay-as-you-go, subscription, and pay-for-use. Pay-as-you-go charges customers based on the overall resource usage for a certain time period. Subscription is another type of fixed pricing, in which a customer pays a fixed amount in advance to use the service for a longer period of time. Depending on the technical features of cloud services, such as storage and bandwidth, a customer is required to make pay-for-use fees. The common feature of these three pricing mechanisms is that price remains stable over the contact time. While fixed pricing strategies are straightforward to both provider and customer, the mechanism is inappropriate to all customers because not all customers have the same needs (Al-Roomi, et al., 2013; Luong, et al., 2017). Even for providers, the mechanism is not optimal, especially when demand is higher than supply in the real world. Our pricing design is embedded with an auction algorithm that can offer all customers with equal opportunities and competition associated with dynamically changing prices.
The most prevalent method of pricing in cloud is pay per use, which is based on Units with constant price. Another common pricing model is subscription, whereby users sign a contract (subscribe) based on constant price of service unit and a longer period of time, for example six months or a year. Obviously, customers and providers would like to use static and simple pricing model in order to ease payment prediction. Nevertheless, for high-value services dynamic pricing can be more efficient.

Dynamic Pricing Strategy

The final price is calculated by a dynamic pricing mechanism. In this pricing strategy, prices are changing with respect to market condition or status. An appropriate pricing algorithm is an effective way to estimate the value of cloud service and to allocate the available resources. Aiming at providing high QoS, Sharma et al. employed the financial option theory and the related economic model to capture the real value of cloud service within the lower and upper boundaries.

Macias and Guitart [16] proposed a genetic model for pricing in cloud computing markets. Choosing a good pricing model via their genetic algorithms involved three main steps: define a chromosome, evaluate it, and finally select the best pairs of chromosomes for reproduction and discarding those with the worst results. The results of the simulation illustrated that genetic pricing acquired the highest revenues in most of the scenarios. Service providers employing genetic pricing achieved revenues up to 100% greater than the other dynamic pricing strategies and up to 1000% greater than the fixed pricing strategy. The proposed genetic model with a flexible genome was proven to be more stable against noise and earned more money than the one with the rigid genome. The proposed genetic model is easy to implement, flexible, and easily adapted to a set of various parameters that influence pricing. The genetic pricing approach can be further explored by defining relations between the parameters that influence pricing.

Li et al., [27] proposed a pricing algorithm for cloud computing resources. This proposal used the cloud bank agent model as a resource agency because it could provide the proper analysis and assistance for all members. The authors used a price update iterative algorithm to determine the price. It analyzed the historical utilization ratio of the resources, iterated current prices constantly, assessed the availability of resources for the next round, and determined the final price. The model included a user request broker (GCA), cloud banking, a cloud service agent (CSA), and a cloud resource agent (GRA). The proposed pricing model was comparatively fixed because it could not adapt to the rapid changes that typically occur in the market. However, it could reduce the costs to providers and maximize their revenues, allowing resources to be used more effectively.

In particular, Xu and Li (2013) study practices of AWS spot instances—spot prices are changed dynamically by AWS based on real-time demand and idle capacity. Truong-Huu and Tham (2014) consider competition between service providers and formulate the competition through a Markov Decision Process. However, those studies have not examined and compared the pricing schemes that are considered in our study, and they have not studied the service providers' optimal choices of pricing schemes and the impact of the chosen pricing schemes on the customers' choices of service providers.

Another major recent research on cloud pricing is dynamic pricing algorithms. The market-clearing price is computed by a dynamic pricing mechanism. In this pricing strategy, prices are changing with respect to market condition or status. An appropriate pricing algorithm is an effective way to estimate the value of cloud service and to distribute the available resources. The dynamic pricing mechanism can accurately estimate the price based on market status and present a reasonable price for both provider and customer. For instance, Macias and Guitart (2011) proposed a Genetic Model for pricing in cloud computing markets. Sharma et al. (2012) employed the Financial Option Theory and the related economic model to capture the real value of cloud service within the lower and upper boundaries. Truong-Huu and Tham (2014) used a Markov Decision Process to illustrate the competition between providers and to compute the price. Dynamic pricing strategy mainly focuses on changing prices based on supply and demand, but the strategy ignores the important cloud feature, especially the QoS metrics. The value of cloud is directly affected by QoS metrics and operational process. Although dynamic pricing strategy can offer a reasonable price, it is difficult for practitioners to estimate the actual price of cloud, because there are no QoS-related parameters in the extant dynamic pricing designs. Our pricing design added the QoS availability into pricing models that clearly indicates how a customer can bid directly on the expected QoS level with the related price.

Auction Pricing Strategy

Theoretically, prices should be determined by the interaction between supply and demand; however, in practice, it seems like that the companies employ specific auction mechanisms in setting up the prices, in order to make optional profits and to attract more potential customers. From Amazon’s initial effort of using auction-based allocation, it is reasonable to expect that cloud providers will be interested in more efficient allocation and pricing schemes in the near future.

This is a market mechanism for service allocation that enables users and providers to deal through double-sided combinational auction. This mechanism is suitable for cases requiring various services and where many participants exist. Both users and service providers should be satisfied by the resource allocation mechanism. A double-sided auction model and K-pricing scheme were used in this mechanism.

Basically, all participants accurately know the nature of the available resources and the distribution of these possible resources. Each participant, whether the participant is a provider or a customer, needs a bidding strategy (Milgrom and Weber, 1982). Reasonable uncertainties from participants and markets make auction design more attractive and more desirable through some mathematical algorithms and bidding strategies.

Many classic auction designs were employed to price cloud services. When they were applied to pricing cloud services, the mechanism will be NP-hard. Wang, Wang, & Huang (2012) proposed an English auction. The mechanism focuses on how to maximize the seller’s revenue and shorten the execution time. Another study about dynamic auctions was conducted by Wang, Liang, & Li (2013) to cope with changes in cloud market. Asymptotic optimization, incentive compatibility, and computational complexity were pursued, through their capacity allocation scheme. A double auction model has been used to distribute cloud resources to multiple buyers and sellers (Shang, et al., 2010). A framework uses marginal bid to evaluate cloud capacity distribution and to generate reasonable revenue for the cloud provider (Lin, et al., 2010).

In practice, leading cloud service companies, such as Amazon AWS or Microsoft Azure, lack the adoption of a desirable auction mechanism to efficiently allocate resources. For instance, Amazon EC2 offers three categories of pricing schemes to sell its cloud services: Reserved Instances, On-Demand Instances (Fixed Price), and Spot Instances. Spot Instance is one type of auction-style pricing policy, through which bidders can periodically bid on offered resources dynamically. Successful bidders can use these instances until the auction price exceeds their bids at a later date. This design has attracted significant attention from practitioners and researchers, and it has prompted a number of studies on auction-based pricing design. Zaman and Grosu (2013) believed that an auction design is better than the current fixed-price mechanism. Depending on the experimental results, the Auction-Greedy is a better choice than the Auction-Linear Programming. The Auction-Greedy can generate higher revenue and resource utilization in a limited time. But these mechanisms are only applied to certain types of items. Extended from Zaman and Grosu (2013), Samimi, Teimouri & Mukhtar (2016) proposed the Combinatorial Double Auction Resource Allocation, focusing on the economic efficiency of the cloud computing environments. Providers are more concerned about what economic profits they can obtain. Thus, based on individual rationality and incentive compatibility, our study will address two other important issues, QoS-based pricing mechanisms and revenue approximation.

Table 1. Summary of Pricing Scheme for Cloud Allocation
Fixed Pricing Strategy Classification: Pay-as-you-go, Subscription, Pay for Resource Characteristics: Stable price for the time of usage. Easily implemented, without auctioneer. Example: Amazon EC2, Microsoft Azure
Dynamic Pricing Strategy (Macias and Guitart, 2011; Sharma et al., 2012; Du, Das and Ramesh, 2012; Truong-Huu and Tham, 2014) Classification: Genetic Model, Financial Option Model, Markov Decision Process Characteristics: Price is changed based on supply and demand. Multiple providers to multiple customers.
Auction Pricing Strategy
English Auction (Wang, Wang, & Huang, 2012) Characteristics: One-sided ascending auction. Winning buyer pays the bidding price. One single provider to multiple customers, with an auctioneer.
Double Auction (Zaman & Grosu, 2013; Samimi, Teimouri & Mukhtar, 2016) Characteristics: Both provider and customer propose bids. Multiple providers to multiple customers, with an auctioneer
Combinatorial Auction (Zaman & Grosu, 2013; Samimi, Teimouri & Mukhtar, 2016) Characteristics: Customer can bid for customized bundle cloud services. The auction design construct computation complexity. Multiple providers to multiple customers, with an auctioneer

In a cloud environment, practitioners are concerned about Quality of Service (QoS). The sharing of resources, software, and information makes up the basic functions of cloud computing. In an auction design, economic models can be used to reduce unnecessary costs, to regulate available resources, to provide incentives for providers, and to stimulate buyers to choose the preferred services that are associated with rational evaluation and QoS criteria. Due to the heterogeneity of cloud computing, a well-designed auction is a basic step toward allocating available resources to meet supply and demand, via the Internet. A QoS-based auction design is a good attempt at finding an efficient and effective way to benefit from cloud computing and to guarantee the performance of the cloud.

Quality of Service describes the requirements that a service provider should provide to its customers, such as service availability, reliability, security, privacy, scalability, and integrity (Zheng, Xu, & Chai, 2017). Our model allows bidders to submit their own valuations and QoS-based bidding strategies. The transaction price will be calculated based on the basis of each participant’s bid directly related to QoS availability, although the economic perspectives of supply and demand in the cloud market still impact participants' auctioning strategies. What we did was using an exponential function to price cloud services based on QoS metrics (availability) and bids from customers. In this way, available cloud services can be distributed to clients with heterogeneous QoS expectations and portfolio requirements. Only buyers submit bids and QoS preferences for a certain item. A QoS-based auction mechanism is a good pricing strategy that could enrich the classic auction design, especially for the bidding targets with unique or not-easy-to-estimate features, e.g., cloud service has QoS criteria.

THE MODEL OF QOS AND PRICE

Reliability and Availability

The reliability and availability of cloud computing are the crucial QoS parameters, but they are difficult to quantifiably analyze because cloud services are implemented through the entire serving process in a complicated network that integrates software, hardware, and the related techniques. Reliability focuses on the downtime of cloud service, while, availability represents the uptime of cloud services (Armbrust, et al., 2010; Mell & Grance, 2011). Reliability and availability are interrelated with each other. A better performance of reliability will lead a better performance of availability. A more stable system will decrease the number of failures and lessen the time of repairing. In the real world, reliability and availability are integrated with each other greatly impacting the overall performance of cloud services.
The relationship between reliability and availability is:

$$A = \frac{MTBF}{MTBF+MTTR} = \frac{Uptime}{Uptime+Downtime} \tag{1}$$

A is referred to availability. MTBF (Uptime) and MTTR (Downtime) are the two parameters of reliability. MTBF is the Mean Time between Failure, and MTTR is the Mean Time to Repair. Thus, based on the above function, the availability can be increased either by increasing the average time interval between repairs (MTBF) or by decreasing the repair time (MTTR). The intuitive way to represent availability is using the downtime associated with Nines and Fives. The following table is an example.

Table 2. Representation of Availability

Availability	Downtime
99% (2-nines)	3.65 days/year
99.5% (2N5)	1.825 days/year
99.9% (3-nines)	8.76 hours/year
99.95% (3N5)	4.38 hours/year
99.99% (4-nines)	52 minutes/year
99.995% (4N5)	26 minutes/year
99.999% (5-nines)	5 minutes/year

The Relationship between Availability and Price

The QoS metrics is the availability of cloud service ($A_{(x)}$), and the bidding price is the average price (y) of the same item from different customers' bids.
The function that expresses the availability is:

$$A_{(x)} = 1 - 10^{-x+g(x)} \tag{2}$$

Where *x* represents the level of availability as Nines and Fives, e.g., x=2, it means "2Nines" as 99%; x=3.5, it means "3N5" as 99.95%. x = {x/x=2, 2.5, 3, 3.5, 4, 4.5, …}. $g(x)$ is a function that is potential to calculate for the result of *x*. Since we know all possible (x, y): (2, 99%), (2.5, 99.5%), (3, 99.9%), …, the function can be addressed as:

$$A_{(x)} = 1 - 10^{-x+(0.5+\ln 0.5)sin^2(\pi x)} \quad (3)$$

Therefore, *x* is expressed by $A_{(x)}$ as:

$$x = \begin{cases} -\ln\left[1 - A_{(x)}\right], if\ x\ is\ an\ integer \\ (0.5 + \ln 0.5) - \ln\left[1 - A_{(x)}\right], if\ x\ is\ not\ an\ integer \end{cases}$$

One Example of The Relationship between Availability and Price

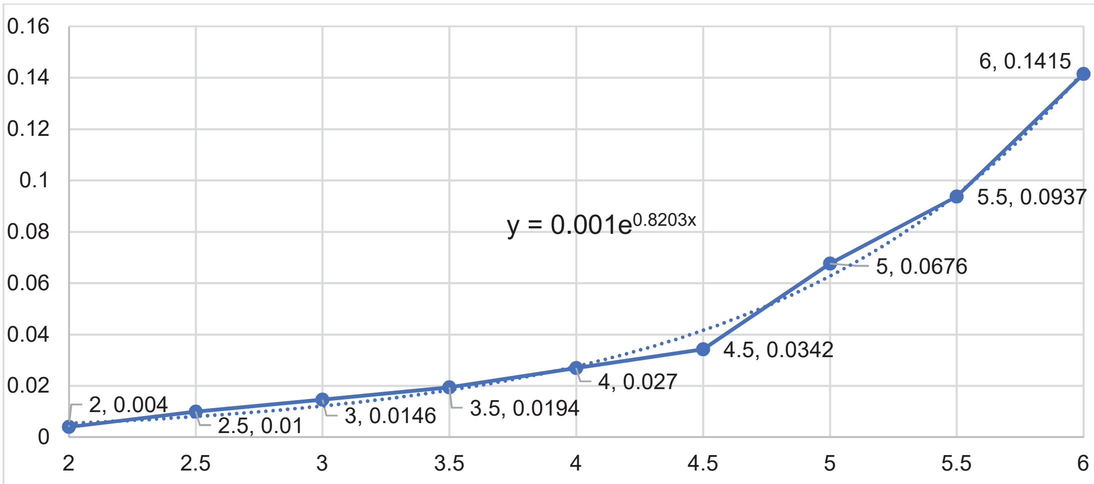
The exponential function that expresses price by the QoS availability is:

$$y = a + cb^{h_{(x)}} \quad (4)$$

Where, *a* is the constant and *c* is the slope. *b* is the base of the exponentiation, and $h_{(x)}$ is the exponent. Now, we can compute the exponential function based on *x* and *y*. In fact, there exist many different ways to establish the relationship between QoS metrics and cloud price. This relationship can be represented by a linear function that is easy to implement but not accurate. Because even with a slight increase in reliability and availability, the costs associated with cloud services increase significantly. Hence, an exponential algorithm is a better approach to describe the relationship between QoS metrics and cloud price. Figure 1 illustrates one example of the exponential relationship between price and availability. Based on the different levels of availability (*x*) that customers bid, such as 2 (99%), 2.5 (99.5%), 3 (99.9%), 3.5 (99.95%), 4 (99.99%), 4.5 (99.995%), 5 (99.999%), 5.5 (99.9995%), 6 (99.9999%), firstly the average price (*y*) of each level of cloud service is calculated accordingly: \$0.004, \$0.01, \$0.0146, \$0.0194, \$0.027, \$0.0342, \$0.0676, \$0.0937, \$0.1415. Then, based on function (4), the relationship between QoS availability and price is calculated as:

$$y = 0.001e^{0.8203x}$$

Figure 1: One Example of the Exponential Relation



ROBUST OPTIMIZATION SCHEME

Robust Auction Design

Symbols and Assumptions

This section addresses the related parameters and mathematical symbols used. Suppose there are *n* customers, *n* = {*1, 2, …, n*}, and *ith* represents the customer *i*. There are *m* providers, *m* = {*1, 2, …, m*}, and *jth* represents the resource *j*. $T_{m,j}$ represents the *mth* supplier offers the resource *j* or not. $T_{m,j}$ is a binary factor, $T_{m,j}$ (=1) means that the mth supplier is capable of offering the resource *j*, $T_{m,j}$ (=0) means that the *mth* supplier doesn't provide the resource *j*. $w_{i,j,m}$ illustrates that the customer *i* obtains the resource *j* provided by supplier *m*, and all the bidding resources are set in the matrix *W*.

In the system, the customer is the buyer, the provider is the seller. The buyer's bid is $b_{i,j,m}$, represented by vector B, and the final price is $p^b_{i,j,m}$. The customer *i*'s value is V_i, $v_{i,j,m}$ represents the valuation of the resource *j* provided by the seller *m*. The cost of the resource *j* is $c_{i,j}$. $D_{i,j}$ represents the demand of the buyer *i* on the resource *j*. The capability of the resource *j* from the seller *m* is $C_{j.m}$. The more the buyer bids, the utility (*μ*) is higher. Specifically, $\mu_{i,j}$ = $h_{i,j}v_i\ (b_{i,j}) - c_i\ (b_{i,j})$.

The customer's bid has two general constraints. (1) a customer needs to bid incrementally, (2) all customers are truthful bidders. The set (*U*) of all possible bids is: U_j = {($b_{1,j}$, $b_{2,j}$, …, $b_{n,j}$) | -Γ ≤ $\frac{\sum_{i=1}^{n} v_{i,j} - n\mu_j}{\sqrt{n}\,\sigma_j}$ ≤ Γ} . Where μ_j and σ_j are the bidding prices of two different time, respectively.

Optimal Auction Model

The overall auction model (*M1*) is:

$$max \sum_i \sum_j \sum_m x_{i,j,m} b_{i,j,m} \tag{5}$$

Subject to

$$\sum_m x_{i,j,m} = 1, \forall\, i, j \tag{6}$$

$$\sum_i x_{i,j,m} D_i T_{m,j} \leq C_{j,m}\,, \forall\, m, j \tag{7}$$

$$\sum_m \sum_j \overline{b_{i,j}} - \overline{p_i} \leq \sum_m \sum_j b_{i,j} x_{i,j,m} - p_i \quad (8)$$
$$c_{i,j,m} \leq b_{i,j,m} \leq v_{i,j,m}, \forall\, i,j,m \quad (9)$$
$$x_{i,j,m} \in \{0,1\}, \forall\, i,j,m \quad (10)$$

Function *M1* is the maximum revenue for all the bidding resources. Constraint (6) indicates that there must exists at least one seller (*m*) can provide the resource *j* as the buyer (*i*) expects. Constraint (7) indicates that the possible allocation of a seller's resources is less than the seller's capacity. Constraint (8) indicates that all buyers are trustworthiness. Constraint (9) indicates that the bid is between the cost and the value of the resource.
Function *M1* is an integer (0-1) programming problem. This problem belongs to NP-hardness (Karp, 1972). It is transformed to a solvable linear programming problem to find the optimal solution.

Solution to the Model

Algorithm 1

The *M1* model is based on the fact that $b_{i,j,m}$ can be estimated and controlled, but in reality, each user's pricing strategy and actual bid are dynamic and difficult to estimate. This study focuses on how to allocate resources to users based on their bidding strategies to meet the needs of users and resource providers. For the problem of uncertain optimization, a robust optimization method can be used to solve the problem. There are four criteria: optimal coordination based on probability, optimistic criteria, conservative criteria, and optimistic and conservative weighting (Beyer & Sendhoff, 2007; Bertsimas, Brown & Caramanis, 2011; Chaithanya & Dimitris, 2014). This study uses conservative criteria. It is seeking the worst result of the original problem and then finding the optimal case under the worst scenario.
Since $b_{i,j}$ is uncertain in model M1, $\widehat{b_{i,j,m}}$ is used to represent the conservative bid of user *i*, and the user *i* has to pay this amount to get the required resource *j*. Constraint (8) can be converted to its equivalent. In order to achieve a reasonable resource allocation, each customer first proposes a conservative bid, and then provider can obtain a higher return based on the bid. Using a conservative bid ($\widehat{b_{i,j,m}}$) converts the original 0/1 problem into a linear programming problem. *M1* is converted to *M2*:

$$max \sum_i \sum_j \sum_m x_{i,j,m} \widehat{b_{i,j,m}} \quad (11)$$

Subject to

$$\sum_m x_{i,j,m} = 1, \forall\, i,j$$
$$\sum_i x_{i,j,m} v_{m,j} \leq B_{i,m}\,, \forall\, m,j \quad (12)$$
$$\sum_i x_{i,j,m} D_i T_{m,j} \leq C_{j,m}\,, \forall\, m,j$$
$$\sum_m x_{i,j,m} \widehat{b_{i,j,m}} \leq \sum_m \sum_j x_{i,j,m} \mu_{i,j,m}\,, \forall\, i,j \quad (13)$$
$$c_{i,j,m} \leq \widehat{b_{i,j,m}} \leq v_{i,j,m}, \forall\, i,j,m \quad (14)$$
$$x_{i,j,m} \geq 0, \forall\, i,j,m \quad (15)$$

Now, *M2* is converted into *M3*, which is the dual problem of M2.

$$min \sum_i \sum_j \varepsilon_{i,j} + \sum_i \sum_m (\delta_{m,j} B_{i,m} + \theta_{m,j} \sum_j x^*_{i,j,m} U_{i,j}) \quad (16)$$

Subject to

$$\varepsilon_{i,j} + z_{i,j,m}\delta_{m,j} + z_{i,j,m}\theta_{m,j} \geq z_{i,j,m} \quad (17)$$
$$\varepsilon, \delta, \theta \geq 0 \quad (18)$$

$z_{i,j,m}$ is the bid of the worst scenario, $x^*_{i,j,m}$ is the solution. The user i's uncertain bid is $U_{i,j}$. Specifically,

$$U_i = argmin \sum_m \sum_j x^*_{i,j,m} \mu_{i,j} \quad (19)$$

The user's conservative bid ($\widehat{b_{i,j,m}}$) can be represented by $z_{i,j,m}$. Since $z_{i,j,m}$ is the bid under the worst scenario,

$$\widehat{b_{i,j,m}} = \varepsilon^*_{i,j} + z_{i,j,m}\delta^*_{m,j} + z_{i,j,m}\theta^*_{m,j} \quad (20)$$

ε^*, δ^*, and θ^* are the parameters of the dual problem, respectively. These parameters can be solved by the model M3. After the results of M3 are solved, the algorithm 1 can be employed to solve $\widehat{b_{i,j,m}}$.

Algorithm 1. Estimated Bidding Algorithm

Algorithm 1 Calculating the Estimated Bidding Price

Input: The Uncertain Price Set U

Output: $\widehat{b_{i,j,m}}$ and $x^*_{i,j,m}$

Based on model M2, calculate the bid of the worst scenario ($z_{i,j,m}$) and the associated allocation result ($x_{i,j,m}$);
Based on model M3, calculate the parameters ε^*, δ^*, and θ^*;

To each of the involved suppliers

Based on Function (20), calculate the estimated bidding price ($\widehat{b_{i,j,m}}$).

Algorithm 2

After we get $\widehat{b_{i,j,m}}$ and $x^*_{i,j,m}$, the Algorithm 2 will solve the allocation and the final price. The intermediary variable $y^b_{i,j,m}$ represents adopted allocation, and the final allocation result is $a^b_{i,j,m} = x^*_{i,j,m} + y^b_{i,j,m}$, $a^b_{i,j,m}$ is the final allocation results. Specifically, bidding vector is *B*, the final price is $p^b_{i,j,m}$. $y^{b-k}_{i,j,m}$ is the temporary intermediate allocation.

$$y^b_{i,j,m} = argmax \sum_i \sum_j \sum_m y_{i,j,m}(b - \widehat{b_{i,j,m}}) \quad (21)$$

$$y_{i,j,m}^{b-h} = argmax \sum_{i \neq k} \sum_{j} \sum_{m} y_{i,j,m} (b - \widehat{b_{\iota,J,m}}) \tag{22}$$

The user's flexible bidding policy:

$$P = arg \begin{Bmatrix} \sum_{i} \sum_{m} y_{i,j,m} \leq 1 - \sum_{i} \sum_{m} x_{i,j,m}^{*} \\ \sum_{j} y_{i,j,m} \mu_{i,j,m} \leq B_{i,m} - \sum_{j} x_{i,j,m}^{*} \widehat{b_{\iota,J,m}} + \sum_{j} x_{k,j,m}^{*} \theta^{*} U_{i,j,m} \end{Bmatrix} \tag{23}$$

Since the final price is determined by the user *i*'s bid and other users' bids, we define Q_k is the allocation of the user *k*.

$$Q_k = arg \begin{Bmatrix} \sum_{i \neq k} \sum_{m} y_{i,j,m} \leq 1 - \sum_{i} \sum_{m} x_{i,j,m}^{*} \\ \sum_{j} y_{i,j,m} \mu_{i,j,m} \leq B_{i,m} - \sum_{j} x_{i,j,m}^{*} \widehat{b_{\iota,J,m}} \end{Bmatrix} \tag{24}$$

Function (21)'s constraint is function (23), and function (22)'s constraint is function (24). The final allocation ($a_{i,j,m}^{b}$) is

$$a_{i,j,m}^{b} = x_{i,j,m}^{*} + y_{i,j,m}^{b} \tag{25}$$

The final price ($p_{k,m}^{b}$) is

$$\begin{aligned} p_{k,m}^{b} = & \sum_{j} y_{k,j,m}^{b} \widehat{b_{\iota,J,m}} + \sum_{j} x_{k,j,m}^{*} \widehat{b_{k,J,m}} - \sum_{j} x_{k,j,m}^{*} \theta^{*} U_{k,j,m} \\ & + \sum_{i \neq k} \sum_{j} \sum_{m} y_{i,j,m}^{b-k} (b - \widehat{b_{\iota,J,m}}) - \sum_{i \neq k} \sum_{j} \sum_{m} y_{i,j,m}^{b} (b - \widehat{b_{\iota,J,m}}) \end{aligned} \tag{26}$$

Algorithm 2. Resource Allocation Algorithm

Algorithm 2 Resource Allocation Algorithm

Input: The bidding vector b and the estimated prices $\widehat{b_{\iota,J,m}}$ and $x_{i,j,m}^{*}$

Output: The final allocation vector $a_{i,j,m}^{b}$ and the final price $p_{i,j,m}^{b}$

To each supplier m

 If $b_{i,j,m} \notin U_{i,j}$

  Won't provide any resource to the user I;

Based on function (21) and (22), calculate the allocation $y_{i,j,m}^{b}$ and the temporary variable $y_{i,j,m}^{b-k}$ that satisfies function (23) and (24);
Based on Function (25) and (26), calculate the final allocation vector $a_{i,j,m}^{b}$ and the final price $p_{i,j,m}^{b}$.

The solution to M2 is solved by Algorithm 1 and Algorithm 2. We get the final allocation $a^b_{i,j,m}$ and the final price $p^b_{i,j,m}$.

Theoretical Proof

In our study, the final price is similar to the transaction price in VCG (Vickrey, 1961; Varian & Harris, 2014). The transaction price is associated with each users' bids, which consist of the winning bids and unsuccessful bids. In this section, we will illustrate that the solutions to Algorithm 1 and Algorithm 2 are also the solution for the original model *M1*.
Lemma 1. The unknown variables (z, x^*, ε^*, δ^*, θ^*, $\hat{b}$) of Algorithm 1 and Algorithm 2 satisfy the following relations.

$$\sum_m \sum_i x^*_{i,j,m} \leq 1, \forall\, j \tag{27}$$
$$\sum_j x^*_{i,j,m} z_{i,j,m} \leq B_{i,m}\,, \forall\, i, m \tag{28}$$
$$\sum_m \sum_j x^*_{i,j,m} z_{i,j,m} \leq \sum_m \sum_j x^*_{i,j,m} z_{i,j,m}\,, \forall\, i \tag{29}$$
$$x^*_{i,j,m} \widehat{b_{i,j,m}} = x^*_{i,j,m} z_{i,j,m}, \forall\, i, j, m \tag{30}$$
$$\sum_j x^*_{i,j,m} \widehat{b_{i,j,m}} \leq B_{i,m}\,, \forall\, i, m \tag{31}$$
$$\sum_i \sum_j \sum_m x^*_{i,j,m} \widehat{b_{i,j,m}} = \sum_i \sum_j \varepsilon^*_{i,j} + \sum_i \sum_m (\delta^*_{m,j} B_{i,m} + \theta^*_{m,j} \sum_j x^*_{i,j,m} U_{i,j}) \tag{32}$$

According to Functions (27)-(29), $z_{i,j,m}$ is a fixed term. The optimal model *M2* can be converted to the followings:

$$max \sum_i \sum_j \sum_m x_{i,j,m} z_{i,j,m}$$

s.t.

$$\begin{gathered}\sum_m \sum_i x_{i,j,m} \leq 1, \forall\, i, j \\ \sum_j x_{i,j,m} z_{i,j,m} \leq B_{i,m}, \forall\, i, m \\ \sum_m \sum_j x_{i,j,m} z_{i,j,m} \leq \sum_m \sum_j x_{i,j,m} \mu_{i,j,m}\,, \forall\, i\end{gathered} \tag{33}$$

Since $x^*_{i,j,m}$ is one solution to the optimal problem, Function (33) can be transformed to another question of *M3*.

$$max \sum_i \sum_j \sum_m x_{i,j,m} z_{i,j,m}$$

s.t.

$$\sum_m \sum_i x_{i,j,m} \leq 1, \forall\, i, j$$

$$\sum_j x_{i,j,m} z_{i,j,m} \le B_{i,m}, \forall\, i, m$$
$$\sum_m \sum_j x_{i,j,m} z_{i,j,m} \le \sum_m \sum_j x^*_{i,j,m} \mu_{i,j,m}\,, \forall\, i \tag{34}$$

If $\bar{x}$ is the solution of function (34), and x^* is the solution of function (33), x^* is also one solution of function (34).

$$\sum_m \sum_i \sum_j \bar{x}_{i,j,m} z_{i,j,m} \ge \sum_m \sum_j \sum_i x^*{}_{i,j,m} z_{i,j,m} \tag{35}$$

And

$$\sum_m \sum_j \bar{x}_{i,j,m} z_{i,j,m} \le \sum_m \sum_j x^*{}_{i,j,m} \mu_{i,j,m}\,, \forall\, i \tag{36}$$

Based on function (35) and (36),

$$\sum_m \sum_j \sum_i x^*{}_{i,j,m} z_{i,j,m} \le \sum_m \sum_j x^*{}_{i,j,m} \mu_{i,j,m} \tag{37}$$

Thus, $x^*{}_{i,j,m}$ is the optimal solution to Function (21). Function (34) can be converted to the followings:

$$\sum_m \sum_j x^*{}_{i,j,m} z_{i,j,m} \le \sum_m \sum_j x^*{}_{i,j,m} \mu_{i,j,m} \tag{38}$$

Hence, $x^*{}_{i,j,m}$ is an optimal solution to Function (33). U_i is computed from Function (19). We have the following:

$$max \sum_i \sum_j \sum_m x_{i,j,m} z_{i,j,m}$$

s.t.

$$\sum_m \sum_i x_{i,j,m} \le 1, \forall\, i, j$$
$$\sum_j x_{i,j,m} z_{i,j,m} \le B_{i,m}, \forall\, i, m$$
$$\sum_m \sum_j x_{i,j,m} z_{i,j,m} \le \sum_m \sum_j x^*_{i,j,m} U_i\,, \forall\, i \tag{39}$$

Obviously, $x^*{}_{i,j,m}$ is the optimal solution to Function (39). The Dual Problem of Function (39) can be the followings:

$$min \sum_i \sum_j \varepsilon_{i,j} + \sum_i \sum_m (\delta_{m,j} B_{i,m} + \theta_{m,j} \sum_j x^*_{i,j,m} U_{i,j}) \tag{40}$$

s.t.

$$\varepsilon_{i,j} + z_{i,j,m}\delta_{m,j} + z_{i,j,m}\theta_{m,j} \geq z_{i,j,m}, \varepsilon, \delta, \theta \geq 0 \tag{41}$$

ε, η, and θ represent the parameters of three constraints. Define ε^*, δ^*, θ^* are the values corresponding to the three parameters, when $x^*_{i,j,m}$ is the optimal solution. According to Complementary Relaxation Theorem, we have the followings:

$$\begin{gathered}\varepsilon^* \sum_m \sum_i x_{i,j,m} = \varepsilon^*, \forall\, j \\ \delta^* \sum_j x_{i,j,m} z_{i,j,m} = \delta^* B_{i,m},\ \forall\, i, m \\ \theta^* \sum_m \sum_j x_{i,j,m} z_{i,j,m} \leq \theta^* \sum_m \sum_j x^*_{i,j,m} U_{i,j},\ \forall\, i \\ x^*_{i,j,m}\left(\varepsilon_{i,j} + z_{i,j,m}\delta_{m,j} + z_{i,j,m}\theta_{m,j}\right) = x^*_{i,j,m} z_{i,j,m}\end{gathered} \tag{42}$$

$x^*_{i,j,m}$ is the optimal solution to Function (34), according to the Strong Dual Theorem, we get,

$$\sum_i \sum_j \sum_m x^*_{i,j,m} z_{i,j,m} = \sum_i \sum_j \delta^* + \sum_i \sum_m \left(\delta^* B_{i,m} + \theta^* \sum_j x^*_{i,j,m} U_{i,j}\right) \tag{43}$$

Based on function (20), (42), and (43), Function (18) and (19) can be proofed. And **Lemma 1** can be proofed as well.

FUTURE DIRECTION

Based on conventional auction design, we added QoS metrics (availability) and established a robust optimal auction to price and allocate cloud resources. For the future research, there are two potential directions. An auction model could be converted to a combinatorial double auction design. In this way, customers will have the opportunity to bid for bundled (package) cloud services (Lu & Zheng, 2018). The major cloud trading platform is centralized and controlled by companies. A blockchain-based decentralized P2P cloud trading platform is potential to play important role to allocate cloud resources between customers without unnecessary third-party intervention (Lu, 2018 a & b; Lu, 2019 b). The decentralized P2P cloud trading system is a good complementary of the conventional trading system.

Providers will compete with each other and submit bids related to guaranteed QoS. The mathematical issue (NP-Hardness) and the computational complexity should be carefully considered as well (Nisan, et al., 2007). Another direction is to adjust the QoS metrics. We only employed the availability as the indicator of QoS in this study. Multiple indicators can be added to represent the exponential relationship between price and QoS. Such as security. More QoS metrics are added into the auction algorithm, more accurate and practical the estimates will be. The relationship between price and multiple QoS indicators will be worth investigating in real-world industries. Whereas, more QoS metrics involved may lead to the complexity of algorithms, deep learning method is a possible way to assess cloud price, according to the training layer (historical pricing records) and output layers (estimated results) (Lu, 2019 a).

CONCLUSION

In the cloud environment, QoS criteria is critical to the overall performance of cloud resources. The extant research seldom illustrates the relationship between QoS metrics and price; our design has the potential to achieve this. By a dynamic pricing mechanism, cloud resources can be traded between provider and customer with uncertainties, e.g., budget constraints. In our work, we address the robust optimal allocation of cloud services to users who are flexible in their bids for their preferred QoS metrics (availability). Our study presents us with one probable way of investigating the questions of how companies formulate prices for cloud services and how customers can utilize the pricing mechanism to bid for cloud services, based upon the QoS metrics (availability). We highlighted the QoS indicator that was employed in the auction mechanism to price cloud services among different participants. The robust optimal auction design is an appropriate pricing mechanism for cloud resources in a distributed trading system.

REFERENCES

Al-Roomi, M., Al-Ebrahim, S., Buqrais, S., & Ahmad, I. (2013). Cloud Computing Pricing Models: A Survey. *International Journal of Grid and Distributed Computing*, 6(5), 93-106.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., knowinski, A., Lee, G., Patterson, D., Rabkin, A., Stoia, I., & Zaharia, M. (2010). A View of Cloud Computing. *Communications of the ACM*, 53(4), 50-58.

Bertsimas D., Brown D. B., & Caramanis C. (2011). Theory and applications of robust optimization. *SIAM review*, 53(3), 464-501.

Beyer H. G., & Sendhoff B. (2007). Robust optimization–a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196(33-34), 3190-3218.

Chaithanya B., & Dimitris B. (2014). Optimal Design for Multi-Item Auctions: A Robust Optimization Approach. *Mathematics of Operations Research*, 39(4), 1012–1038.

Josep, A. D., Katz, R., Konwinski, A., Gunho, L. E. E., Patterson, D., & Rabkin, A. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50-59.

Karp R. M. (1972). Reducibility among combinatorial problems. *Complexity of Computer Computations*, Springer, Boston, MA, 85-103.

Mell, P., & Grance, T. (2011). NIST SP 800-145, The NIST definition of cloud computing. *National Institute of Standards & Technology*, Gaithersburg.

Lin, W. Y., Lin, G. Y., & Wei, H. Y. (2010). Dynamic Auction Mechanism for Cloud Resource Allocation. In Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing. Melbourne, Australia. *IEEE Computer Society*, 591-592

Lu, Y. (2018 a). Blockchain and the Related Issues: A Review of Current Research Topics. *Journal of Management Analytics*, 5(4), 231-255.

Lu, Y. (2018 b). Blockchain: A survey on Functions, Applications and Open Issues. *Journal of Industrial Integration and Management*, 3(04),1850015.

Lu, Y. (2019 a). Artificial Intelligence: A Survey on Evolution, Models, Applications and Future Trends. *Journal of Management Analytics*, 6(1), 1-29.

Lu, Y. (2019 b). The Blockchain: State-of-the-Art and Research Challenges. *Journal of Industrial Information Integration*, Early Access. doi.org/10.1016/j.jii.2019.04.00.

Lu, Y. & Zheng, X. R. (2018). Block Chain Based Double Auction Design. In Proceedings of the 2018 AMCIS on Strategic and Competitive Use of Information Technology (SCUIT). New Orleans. *AIS*. Available: https://aisel. aisnet.org/amcis2018/StrategicIT/Presentations/10/.

Luong, N. C., Wang, P. D. Wen, Niyato Y., & Han, Z. (2017). Resource Management in Cloud Networking Using Economic Analysis and Pricing Models: A Survey. *IEEE Communications Surveys & Tutorials*, 19(2), 954-1001.

Macías, M. & Guitart. J. (2011). A Genetic Model for Pricing in Cloud Computing Markets. In Proceedings of the 2011 ACM Symposium on Applied Computing. TaiChung, Taiwan. ACM. 113-118.

Mell, P. & Grance, T. (2011). The NIST definition of cloud computing. *NIST Special Publication*, 53(6), 50.

Milgrom, P. R. & Weber, R. J. (1982). A Theory of Auctions and Competitive Bidding. *Econometrica: Journal of the Econometric Society*, 50(5), 1089-1122. doi: 10.2307/1911865.

Nisan, N., Roughgarden, T. Tardos, E., & Vazirani, V. V. (2007). *Algorithmic Game Theory*. Cambridge: Cambridge University Press.
Samimi, P., Teimouri, Y., & Mukhtar, M. (2016). A Combinatorial Double Auction Resource Allocation Model in Cloud Computing. *Information Sciences*, 357, 201-216. https://doi.org/10.1016/j.ins.2014.02.008.
Varian H R, & Harris C. (2014). The VCG auction in theory and practice. *American Economic Review*, 104(5), 442-45.
Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1), 8-37.
Wang, W., Liang, B. & Li, B. (2013). Revenue Maximization with Dynamic Auctions in IaaS Cloud Markets. In Quality of Service (IWQoS), 2013 IEEE/ACM 21st International Symposium. Montreal, Canada. IEEE. 1-6.
Wang X. W., Wang, X. Y. & Huang, M. (2012). A Resource Allocation Method Based on The Limited English Combinatorial Auction Under Cloud Computing Environment. In Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference. Sichuan, China. IEEE. 905-909.
Zaman, S. & Grosu, D. (2013). Combinatorial Auction-Based Allocation of Virtual Machine Instances in Clouds. *Journal of Parallel and Distributed Computing*, 73(4), 495-508. https://doi.org/10.1016/j.jpdc.2012.12.006.
Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud Computing: State-Of-The-Art and Research Challenges. *Journal of internet services and applications*, 1(1), 7-18.
Zheng, X., Martin, P., Brohman, K., & L. D. Xu. (2014). CLOUDQUAL: A Quality Model for Cloud Services. *IEEE transactions on industrial informatics*, 10(2), 1527-1536.

APPENDIX

Table 3. Main Variables and Explanations (Alphabetic Order)

Variable	Description	Variable	Description
a	the constant of the exponential function	$p_{i,j,m}^{b}$	the final price
$a_{i,j,m}^{b}$	the final allocation results	Q_k	the allocation of the user k
$A_{(x)}$	the QoS availability of cloud service	$v_{i,j,m}$	the valuation of the resource j provided by the seller m
b	the base of the exponential function	V_i	the valuation of customer i
$b_{i,j,m}$	the buyer's bid	$T_{m,j}$	the mth supplier offers the resource j or not
$\widehat{b_{i,j,m}}$	the conservative bid of user i	U	the set of all possible bids
c	the coefficient of the exponential function	$U_{i,j}$	user i's uncertain bid
$c_{i,j}$	The cost of resource j	x	the level of QoS availability
$C_{j.m}$	The capability of the resource j from the seller m	$x_{i,j,m}^{*}$	the decision variable, $x_{i,j,m} \in \{0,1\}$
$D_{i,j}$	the demand of the buyer i on the resource j	y	the average bidding price of cloud service
h(x)	the exponent of the exponential function	$y_{i,j,m}^{b}$	the intermediary variable
i	the ith bidder	$y_{i,j,m}^{b-k}$	the temporary intermediate allocation
j	the jth provider	$z_{i,j,m}$	the bid of the worst scenario

m	there are m sellers	μ	the utility
n	there are n buyers	ε^*, δ^*, θ^*	the parameters of the dual problem

DECISION SCIENCES INSTITUTE
The Development of an Exercise to Incorporate a Management Dashboard into the Beer Game

Brad C Meyer
Drake University
Email: Bradley.meyer@drake.edu

Deb Bishop
Drake University
Email: deb.bishop@drake.edu

ABSTRACT

An exercise to demonstrate the utility of management dashboards was created. Students were asked to play an online version of the beer game, where supply chain information was available and could be organized and displayed in Tableau. We explain several iterations of the exercise and describe our experience and game development process.

KEYWORDS: Active learning, innovative education, supply chain management, analytics

INTRODUCTION

It is common in these annual DSI meetings to describe techniques and exercises that have been used in class with successful results. By the time the innovation is presented, the bugs have been worked out and it appears as a finished product. What may not be apparent to the hearers are all the trials and errors that have taken place to get the exercise to be effective. For those with a creative bent who find themselves driven to develop learning innovations, the story of the development of an exercise might be as useful as the description of the final exercise. With that in mind, this paper sets out to describe the development of an exercise that has students using a visualization tool, in this case Tableau, to support decision making in a digital version of the beer game. We begin by describing our history of using the beer game in an introductory operations management class, then describe the original objectives for our exercise. We then describe the development of the technology for the game, and then our first runs of the exercise and we detail what did not work as expected. We continue through revisions and retests and finally tie up with the exercise as we are conducting it at this time and our future plans.

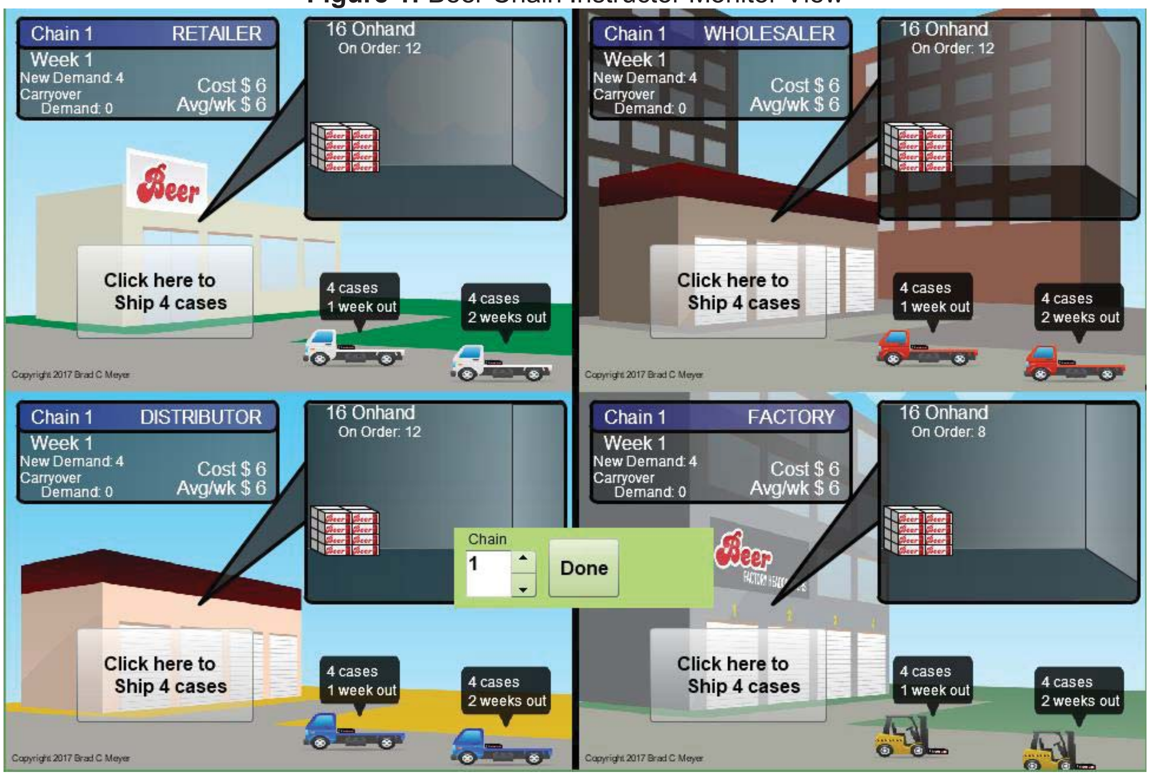
CONCEPTION OF THE EXERCISE

The beer game hardly needs an introduction, but for those unfamiliar with the exercise, we refer you to (Forrester, 1961; Sterman, 2000; and Senge, 1990). The game is a simulation of a supply chain with four echelons, not counting the end customer. The object for the players of the game is to minimize the sum of inventory and backorder costs by making order replenishment decisions. A simple demand stream with the doubling of demand consistently yields a pronounced bullwhip effect in the chain.

After playing the game with a physical board and small poker style chips for many years, one of the authors developed an online version of the game to reduce setup time and to facilitate collecting data from plays of the game. Figure 1 shows the screens for one chain from an instructor observation view.

The digital version of the game has proven to work well to illustrate the bullwhip effect and the results are comparable to the board version with typically higher levels of amplification, probably due to an interface that does not require physically sliding chips. Figure 2 shows an example of the graphs that are displayed at the end of the game. In debriefing the game, the class always concludes that performance would be better if 1) communication was allowed between the echelons and 2) they had a forecast of demand to know what to expect.

Figure 1: Beer Chain Instructor Monitor View



Student's commonly ask for another opportunity to play the game, mainly to redeem themselves from their poor performance in the first round. Using the same demand stream would be too easy, so we developed a version with a different demand stream. Demand begins around 4 per week and then ramps up by 1 every two periods till a peak and then ramps back down. To this underlying pattern is added a minor random component of plus or minus 0 to 3 units. This is a more difficult demand stream, but most teams perform better than the first round when they are given the ability to talk together as a group as the game is played.

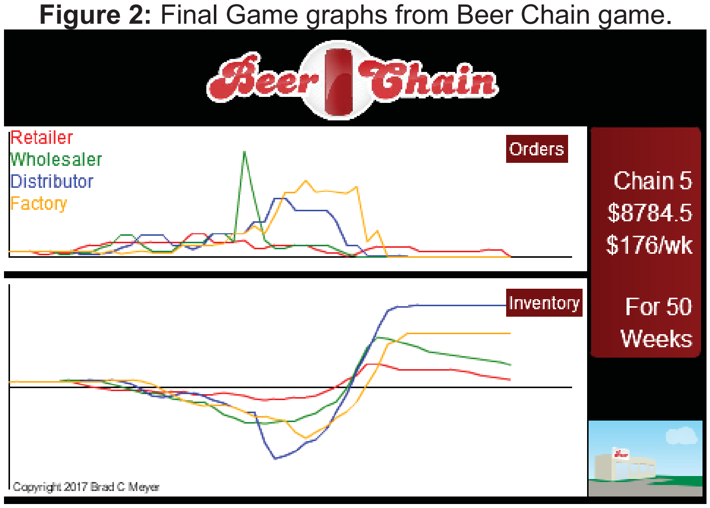
We were content with this approach: run first with no talking, then run a second time with communication but a harder demand stream, until the rise of data analytics. In fall of 2017 we began integrating Tableau into our core operations management class. We developed exercises to showcase the powers of Tableau and we introduced the idea of managing via a dashboard. It wasn't long until we hit on the idea of asking our students to create a dashboard in Tableau that could be used as the game is played. Thus began our development of this exercise.

INITIAL EXERCISE DESIGN

We began development with the following design objectives.

1. Reinforce Tableau skills
2. Give the students experience in creating a useful, albeit simple, dashboard
3. Demonstrate that managing with data yields better performance
4. Reinforce concepts of supply chain management
5. Engage students in activity that supports learning

Figure 2: Final Game graphs from Beer Chain game.



Our first technical step was to develop a way for Tableau to pull data from the database used in the game. Ideally a dashboard would show inventory at every echelon, orders at every echelon, end customer demand, and cost data. The database used by the game stored only order decision values, and they were stored in a compact way such that a direct connection to the database would require a lot of manipulation in Tableau, way above the level of our students. There would also be some permissions issues to overcome.

We settled on an approach that utilizes a Tableau Web Data Connector. This allows Tableau to access data through the internet by connecting to a web page. The web page provides data through .asp calls and a couple layers of additional coding. For those interested, code developed for the web connector is included in appendix A. This code was developed by us and is not seen or understood by the students undertaking the exercise. In the code, we use the order data and known initial values to compute the relevant inventory levels. Thus, the fields that are sent to Tableau are as shown in table 1.

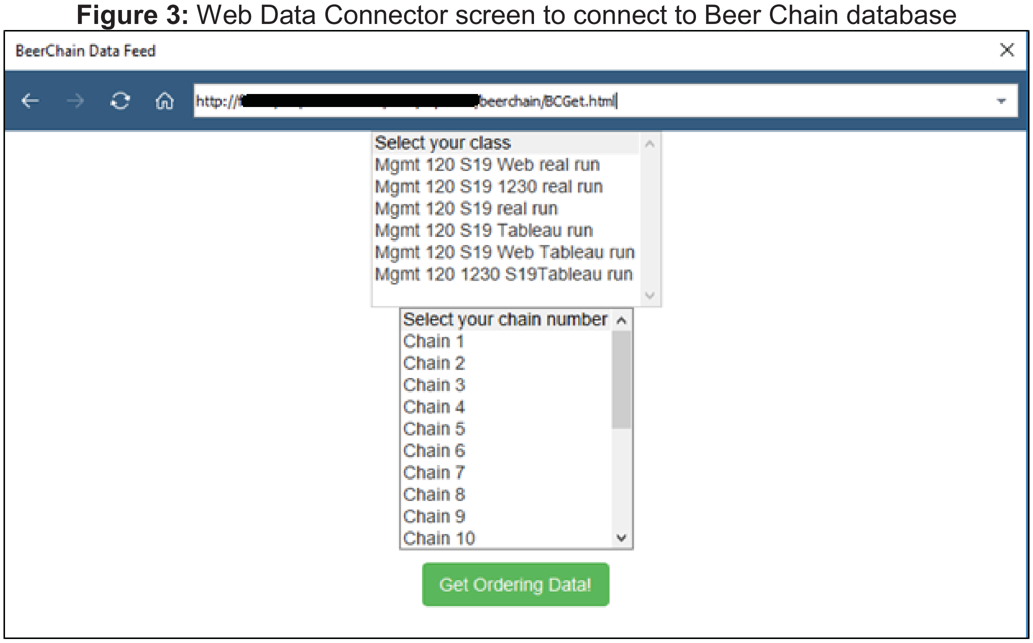
Table 1: Tableau Fields for the Beer Game Data

Field	**Description**
Period	week number, consecutive
Demand	orders from the end customer
Retailer	amount ordered by the retailer
RetailInv	current level of inventory held by retailer, positive or negative

RetailChange	the difference in inventory between current and previous period. (The sum of this column yields current inventory.)
Wholesaler	amount ordered by the wholesaler
WholesalerInv	current level of inventory held by wholesaler, positive or negative
WholesalerChange	the difference in inventory between current and previous period.
Distributor	amount ordered by the distributor
DistributorInv	current level of inventory held by distributor, positive or negative
DistributorChange	the difference in inventory between current and previous period.
Factory	amount ordered by the factory
FactoryInv	current level of inventory held by factory, positive or negative
FactoryChange	the difference in inventory between current and previous period.

From the student side, they are able to connect to the data by selecting *Web Data Connector* from the Source menu, entering in a url supplied by the instructor and then selecting their group and the chain number. Supply chain position is not entered, as each person in each chain has access to data for all the positions in that chain. Figure 3 shows the screen inside Tableau where the class and chain are selected.

Figure 3: Web Data Connector screen to connect to Beer Chain database



We wanted the students to design their own dashboard, but we demonstrated a few visualizations as starter ideas, a graph of demand, a graph of current inventory, and a chart that showed the current inventory at each echelon. Figure 4 shows a dashboard mockup with these visualizations. This figure comes from a handout given to the class for this exercise.

We designed the coding for the Web data connector so that it provides demand data from the past three years, but only has ordering and inventory data starting in year 4, the year for which they will be playing the game. We have students design their dashboard using the data from their first run of the game. The dashboard in figure 4 thus shows demand per week at 4 for the first three years and the first 4 weeks of the 4th year followed by a jump to 8 for the rest of the 4th year.

After they have designed their dashboard, we have them reconnect to the database and populate their dashboard with data from the game they are about to play. Since the database will not have any ordering data yet, all they will see on the dashboard is the demand from the past 3 years which they use to generate a forecast. Figure 5 shows this demand stream along with a forecast generated in Tableau.

Finally, we explained to the students how to refresh the data extract so they can keep their visualizations up to date as they play the game. With these design pieces in place, we began using the exercise.

Figure 4: Dashboard mockup shown to students to spur ideas

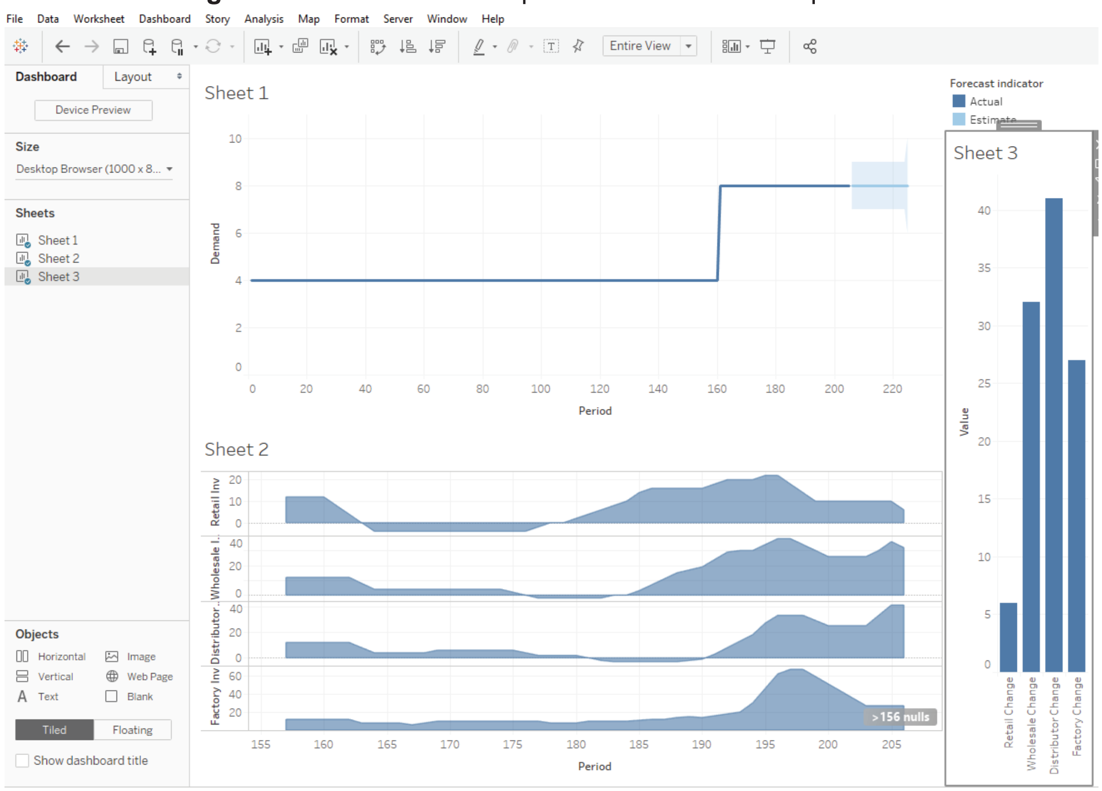


Figure 5: Historical demand for the Tableau Beer Chain exercise with forecast generated by Tableau.

GAME EXPERIENCE AND REVISIONS

The first times we used this exercise, we took 20 minutes at the beginning of a 90 minute class period to walk through the mechanics of connecting Tableau to the game database and demonstrated a few possible visualizations they could use. Then we gave the students 15 to 20 minutes to develop their dashboard and plan a strategy for playing the game. Then the game was started. We allowed the students to continue to communicate with each other while the game was being played.

The activity was engaging. The students typically reduced their costs to half that in the original 4/8 demand stream round. However, we noticed some things that were troubling with respect to our original intentions. The students relied heavily on talking to each other to make the ordering decisions. Often times a leader would arise in the group who had the major sway over the decision making across the entire chain. Most teams had one or two people who were looking at the dashboard, but there was little reliance on the dashboard. The students relied on communication, not on the analytics tool. For example, here are three responses to a debriefing question of whether their strategy to keep costs low worked.

> *"Our strategy seemed to help some. I think we kept confusing ourselves a little by talking through things too much, but for the most part, communicating helped. It was especially helpful that we could discuss our carryover demands & inventory levels (on hand) to help keep inventory levels steady and have a better balance between having too much on hand vs. too much carryover demand. Overall, the communication helped a lot. Looking at Tableau did not seem to help as much as we hoped, but it was slightly helpful to see the overall trend each of our inventory levels and demands."* (Final cost $96/wk)

> *"Our strategy did not work, as we more than doubled our cost from the last time we played. I think it was confusing for us to think ahead ten weeks, and we focused too much on the demand numbers on screen rather than talking to each other about what we needed. The graphs on our Tableau file were not very useful as they did not help us to look forward, they were simply a visualization of our errors."* (Final cost $113/wk)

Secondly, as is also apparent from the comments, many of the groups did not know how to make use of the forecast. That is, it helped them know the order quantity to expect from the end customer, but they didn't know how to plan for it down the chain, timing orders appropriately. The teams who understood the 4 week time delays built into the game and how to plan for a ramp up and ramp down of demand were able to achieve costs of $20 per week or less, total for the entire chain. The rest of the class had costs from $55 per week to $175 per week. We also realized that the students were primarily reacting in the game. They would start the game with some impressions of what should or shouldn't be done when demand started going up or when inventory started piling up, but they did not have cogent strategy going into the game.

The problem of not using the dashboards was the most troubling, given the objectives for the assignment, but we attacked the second problem first. We were using a pre-game questionnaire and we began to improve the questions to drive them to be more proactive and strategic in playing the game.

Our original questions were

- *What do you notice about historical demand?*
- *What will be your strategy to manage the chain to keep costs as low as possible?*

Answers to the strategy question were often vague, such as these examples.

> *"Be mindful of time delays. Since it costs half as much to hold inventory than it does to have backlog, err on the side of too much inventory rather than too little. In addition, obviously keep in mind the seasonality."*

> *"Try to keep inventory at zero."*

> *"We will communicate amongst ourselves to ensure that we are meeting demand and planning ahead with the information our supply chain is receiving."*

So, we tried adding a question to get them to think about timing of ordering. This questions was inserted between the two questions above.

- *How would your chain handle the case where demand was constant at 4 but you knew that in 16 periods, demand would increase to 8?*
-

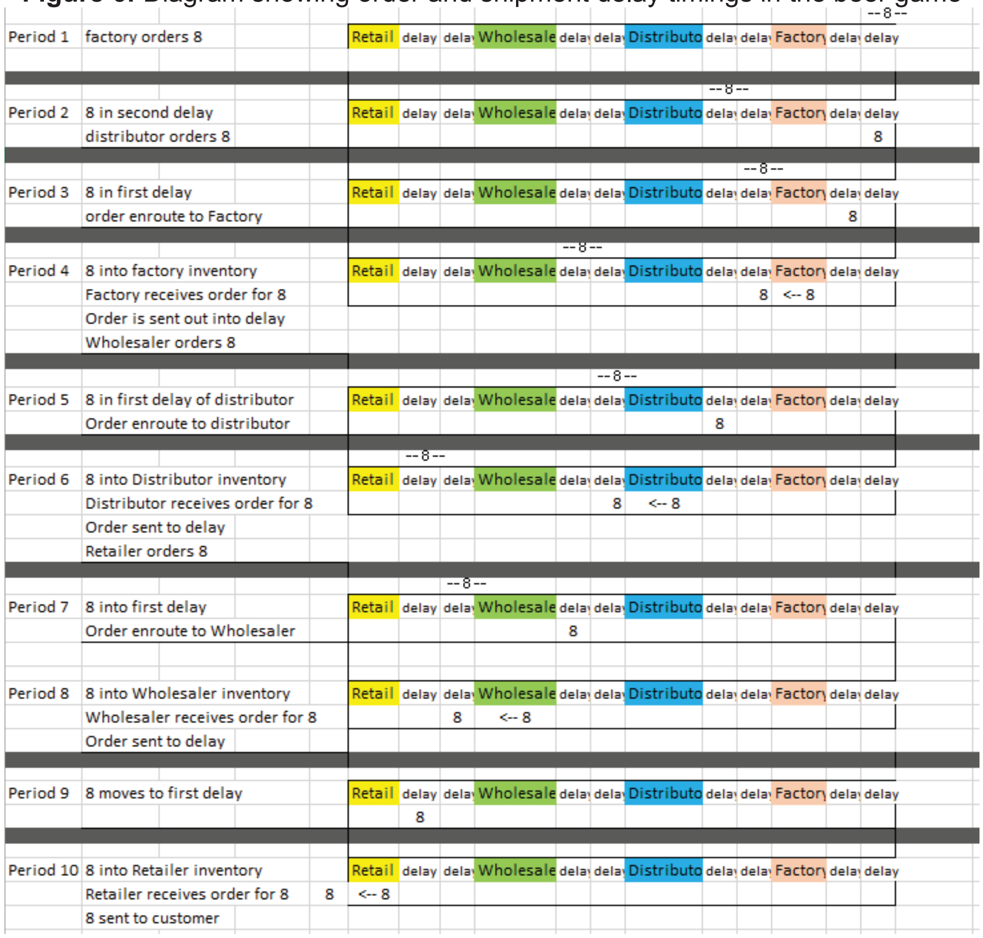
This was somewhat helpful, but we did more clarification in our next semester.

- *Look at the historical demand for the Tableau version of the game you are about to play. How would you describe the pattern of demand? How might this have an impact on how you play the game?*
- *How would your chain (the whole chain) handle the situation where end customer demand has been constant at 4 but you know that 16 periods from now, demand will jump to 8?*

- *Knowing what you know about historical demand, what is your plan for how you will determine order amounts during the Tableau round of the game? Be as specific as possible. For example, don't just say, "we will be careful not to order too much" instead describe what rules or thinking processes you will put in place so that you do not order too much. Or don't say "we will coordinate among the team members and communicate." Rather, tell what information will be communicated and how that will be used in the decision making process.*

We also hypothesized that giving the students only 20 minutes to create a dashboard and develop a strategy to play the game was not enough time to allow for deep thinking about strategy. So, we moved to briefing the exercise on one day and playing the game the following class period after giving teams time to finalize their strategy. Recently, we also added a short description of the impact of the time delays with a diagram, as shown in Figure 6.

Figure 6: Diagram showing order and shipment delay timings in the beer game



As we made these changes the answers to the strategy question improved, below are two examples.

> *"When we begin, the Factory and Distributor will order eight weeks ahead; the demand will be about 8. Wholesaler will order 6 weeks ahead, Retailer will order 4 weeks ahead. Every two weeks, we increases the order by 1 as demand increases by about half a case every week. After having ordered 16 twice, we will start decreasing our order amount by one every two weeks. So this follows the pattern of the demand so that we have the lowest inventory as possible to fulfilling all the orders."*

> *"We would order based off of the forecast. We changed the confidence interval to 90% and added a trend line so that while our data might not be quite as exact it will show some wider guidelines that we can follow and more methods to base our predictions off of. We also based our trend lines and intervals based off of what position we each had because retail would need to have a faster turn around than wholesaler. We decided as a team that we would rather over compensate for inventory so we are aiming to be towards the higher end of our confidence interval versus the lower end."*

While some of the strategies were not correct, they were at least detailed enough to guide the decision making process. In watching the game play, however, we realized that the students were still relying more on talking to each other than to looking at the dashboard, so in our next iteration, we did not allow them to talk to each other during the game. Instead, they would have to look at the dashboard to learn what was happening at the other echelons. Observing the teams playing this was revealed two main approaches. One portion of the students watch their dashboard carefully and try to make decisions based on the information displayed, another portion of students reverts to playing the game the way they played the first round where they simply watch the game screen and try to react in the best way, based on their strategy developed at the beginning of the game.

In the most recent iteration using this exercise (each iteration involves a new set of students, usually a new semester,) there was one team that achieved an average cost of $13 per week, which is about $3 per week at each of the 4 echelons. This is the group with the first of the two comments at the top of this page. This group figured out the pattern of demand and the staggered order timing necessary to maintain low costs. In fact, their strategy was good enough that they might not have even needed the dashboard. For this group that dashboard served as a status monitoring device and allowed them to make minor tweaks adapting to the randomness in demand.

At first, we were disappointed that the high achieving group was not leaning heavily on the dashboard during the game, but then we realized that this result provides us a platform to point out the difference between two primary uses of dashboards. First, a dashboard can be used for status monitoring. In fact, this is the primary use of a dashboard. In a car, the purpose for the dashboard is to notify the driver if there are any potential problems in the car system: oil level, engine temperature, battery charge, speed, etc. Adjustments can be made if some part of the system departs from normal functioning. In the beer game, the dashboard should do the same and alert the players if inventory is heading in a wrong direction, so adjustments can be made to order sizes.

The second use of a dashboard is to provide insight into a system. In our exercise, the dashboard was used *before starting the game* to display the historical demand pattern. That yielded the insight that demand was seasonal and ordering decisions should reflect that same seasonality. This insight had to be united with understanding of the time delays in the system. The delays were explained even before the first round of the game. We tried to help the students connect the dots by forcing them to think about how orders would have to be timed throughout the chain in order to get the necessary quantity to the retailer to match end customer demand. While it would be theoretically possible to determine the time delays in the system simply from historical data, that is not likely. The truth of the matter is that data driven decisions arise from a combination of system understanding and insights in the data. This exercise provides a good context to illustrate these concepts.

FUTURE DIRECTION

We are still modifying this exercise. We would like to see more advanced dashboards. For example, we have seen only a few teams display cost figures on the dashboard. The data provided allows for costs to be computed at each echelon and for the entire chain. We would also like to see triggers or warning zones on the dashboard that call attention to rising costs, to rising inventory levels, or dangerously low inventory levels. The way we plan to support these ideas is by extending the content we present when we first introduce management dashboards.

It is possible that students might use the dashboards more or in different ways if the game was played longer and customer order patterns were subject to other kinds of disturbances. We will consider options in this regards.

On the technical side, the current data feed provides a set of values for each week. Another approach that has some advantages for Tableau visualizations, is to provide a simpler data record. Each record would have 4 fields as show in Table 2. With this structure there are better options for drawing graphs that compare levels across the various roles.

Table 2: Revised Web Data Connector Structure

Field	Description
Period	week number, consecutive
Role	retailer, wholesaler, distributor, factor, or customer
Item	order, inventory, delay 1, delay 2, change in_inventory
Amount	applicable quantity

Finally, as we near the point where we are happy with the exercise itself, we will consider ways to make this assignment more easily available to instructors at other schools than our own.

APPENDIX

Below is the code for the 3 files that enabled the Tableau web data connection. These are not seen by the students. There is an html file, BCGet.html, a javascript file, BCGet.js, that defines the data fields and an .asp file, readChainTableuaMore.asp that connects to the Access database file and creates a table in JSON which is a coding format understood by Tableau.

BCGet.html - This is the file that the url in Tableau points to. Currently, this is not automatically created from the database, but has to be tweaked manually. A future version will develop the menu of classes and chain numbers from the database itself.

```
<!DOCTYPE HTML>

<html>
<head>
    <title>BeerChain Data Feed</title>
    <meta http-equiv="Cache-Control" content="no-store" />

    <link href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/css/bootstrap.min.css"
rel="stylesheet" crossorigin="anonymous">
    <script src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.1/jquery.min.js"
type="text/javascript"></script>
    <script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.6/js/bootstrap.min.js"
crossorigin="anonymous"></script>

    <script src="https://connectors.tableau.com/libs/tableauwdc-2.3.latest.js"
type="text/javascript"></script>
```

```
    <script src="BCGet.js" type="text/javascript"></script>
</head>

<body>
    <div class="container container-table">
        <div class="row vertical-center-row">
            <div class="text-center col-md-4 col-md-offset-4">
            <select id="classSelect" name="ClassNumber" size="8">
            <option value="0" selected="selected">Select your class</option>
            <option value="154">Mgmt 120 S19 Web real run</option>
            <option value="156">Mgmt 120 S19 1230 real run</option>
            <option value="158">Mgmt 120 S19 real run</option>
            <option value="159">Mgmt 120 S19 Tableau run</option>
            <option value="160">Mgmt 120 S19 Web Tableau run</option>
            <option value="161">Mgmt 120 1230 S19Tableau run</option>
            </select>
            </div>
            <div class="text-center col-md-4 col-md-offset-4">
            <select id="chainSelect" name="ChainNumber" size="11">
            <option value="0" selected="selected">Select your chain number</option>
            <option value="1">Chain 1</option>
            <option value="2">Chain 2</option>
            <option value="3">Chain 3</option>
            <option value="4">Chain 4</option>
            <option value="5">Chain 5</option>
            <option value="6">Chain 6</option>
            <option value="7">Chain 7</option>
            <option value="8">Chain 8</option>
            <option value="9">Chain 9</option>
            <option value="10">Chain 10</option>
            <option value="11">Chain 11</option>
            <option value="12">Chain 12</option>
            <option value="13">Chain 13</option>
            <option value="14">Chain 14</option>
            <option value="15">Chain 15</option>
            <option value="16">Chain 16</option>
            <option value="17">Chain 17</option>
            <option value="18">Chain 18</option>
            <option value="19">Chain 19</option>
            <option value="20">Chain 20</option>
            </select>
            </div>
            <div class="text-center col-md-4 col-md-offset-4">
                <button type="button" id="submitButton" class="btn btn-success" style="margin:
10px;">Get Ordering Data!</button>
            </div>        </div>
    </div>
```

```
</body>
</html>
```

BCGet.js – this is a javascript function containing the code necessary to define the fields for Tableau.

```
(function () {

    var myConnector = tableau.makeConnector();

myConnector.getSchema = function (schemaCallback) {
    var cols = [{
        id: "period",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "demand",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "retailer",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "wholesaler",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "distributor",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "factory",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "retailInv",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "retailChange",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "wholesaleInv",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "wholesaleChange",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "distributorInv",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "distributorChange",
```

```
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "factoryInv",
        dataType: tableau.dataTypeEnum.int
    }, {
        id: "factoryChange",
        dataType: tableau.dataTypeEnum.int
    }];

    var tableSchema = {
        id: "orderDataFeed",
        alias: "Demand so far in the Beer Chain game",
        columns: cols
    };

    schemaCallback([tableSchema]);
};

myConnector.getData = function(table, doneCallback) {
     var callObj = JSON.parse(tableau.connectionData),
     c = callObj.chain,
     g = callObj.groupID;
    //var c = "4";
    var urlString =
"http://[redacted]/BeerChain/readChainTableauMore.asp?chain="+c
+"&groupID="+g;
    $.getJSON(urlString, function(resp) {
        var feat = resp.orderDataFeed,
            tableData = [];
        // Iterate over the JSON object
        for (var i = 0, len = feat.length; i < len; i++) {
            tableData.push({
                "period": feat[i].period,
                "demand": feat[i].demand,
                "retailer": feat[i].retailer,
                "wholesaler": feat[i].wholesaler,
                "distributor": feat[i].distributor,
                "factory": feat[i].factory,
                "retailInv": feat[i].retailInv,
                "retailChange": feat[i].retailChange,
                "wholesaleInv": feat[i].wholesaleInv ,
                "wholesaleChange": feat[i].wholesaleChange,
                "distributorInv": feat[i].distributorInv,
                "distributorChange": feat[i].distributorChange,
                "factoryInv": feat[i].factoryInv,
                "factoryChange": feat[i].factoryChange
            });
```

```
        }

        table.appendRows(tableData);
        doneCallback();
    });
};

    tableau.registerConnector(myConnector);
    $(document).ready(function () {
    $("#submitButton").click(function () {
        var callObj = {
            chain: $("#chainSelect option:selected").val(),
            groupID: $("#classSelect option:selected").val()
            };
        tableau.connectionData = JSON.stringify(callObj);
        tableau.connectionName = "BeerChain Data Feed";
        tableau.submit();
    });
});

})();
```

readChainTableauMore.asp – this is the file that accesses the database and creates a JSON file according to the format required by Tableau. The demand stream in the game is generated by a pseudo-random number generator, in such a way that the values are the same for every chain. In this file, that sequence of numbers is simply listed, rather than recreating the random number generator. There are other options of demand streams built here not discussed in this context.

```
<%
Function min(a,b)
   If a < b then
        min = a
   Else
        min = b
   End If
End Function

Function max(a,b)
   If a > b then
       max = a
   Else
       max = b
   End If
```

```
End Function

Function JSONDate(myDate)
   Q = chr(34)
   d = LeadingZero(Day(myDate))
   m = LeadingZero(Month(myDate))
   y = Year(myDate)
   JSONDate = Q & y & "-" & m & "-" & d & Q
End Function

Function LeadingZero(num)
   If(Len(num)=1) Then
      LeadingZero="0"&num
   Else
      LeadingZero=num
   End If
End Function


'---- CursorTypeEnum Values ----
Const adOpenForwardOnly = 0
Const adOpenKeyset = 1
Const adOpenDynamic = 2
Const adOpenStatic = 3
'---- LockTypeEnum Values ----
Const adLockReadOnly = 1
Const adLockPessimistic = 2
Const adLockOptimistic = 3
Const adLockBatchOptimistic = 4
' Response.Buffer = False
Response.CacheControl = "no-cache"
'HTTP/1.0 Compatibility
Response.AddHeader "Pragma", "no-cache"
Response.Expires = -1

         Set MyConnection =  Server.CreateObject("ADODB.Connection")
'        FP_DumpError strErrorUrl, "Cannot create connection"

         Set MyRecordset = Server.CreateObject("ADODB.Recordset")
'        FP_DumpError strErrorUrl, "Cannot create record set"

         MyConnection.Open Application("beergame_ConnectionString")
'        FP_DumpError strErrorUrl, "Cannot open database"

Dim usualData
```

```
usualData =Array(4, 4, 4, 4, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8,
8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8,
8)
Dim sixtyData
sixtyData = Array(2,  3, 6, 5, 6, 5, 5, 8, 9, 10, 9, 9, 11, 8, 9, 11, 13, 10, 13, 11, 14, 15, 15, 17, 17,
16, 15, 18, 18, 19, 21, 19, 17, 15, 18, 15, 17, 16, 17, 13, 13, 15, 14, 13, 13, 9, 12, 12, 9, 9, 10, 8,
7, 6, 6, 5, 6, 7, 7, 2, 6 )
Dim fiftytwo
fiftytwo = Array(4, 5, 7, 6, 6, 5, 5, 8, 9, 10, 9, 9, 11, 8, 9, 11, 13, 10, 13, 11, 14, 15, 15, 17, 17,
16, 15, 17, 16, 16, 17, 15, 13, 11, 14, 11, 13, 12, 13, 9, 9, 11, 10, 9, 9, 5, 8, 8, 5, 5, 6, 4, _
4, 7, 8, 7, 8, 6, 8, 7, 8, 8, 8, 9, 11, 12, 13, 10, 14, 13, 12, 13, 13, 15, 14, 14, 17, 17, 17, 14, 15,
14, 17, 14, 14, 13, 14, 14, 12, 11, 13, 11, 8, 10, 7, 9, 8, 7, 7, 5, 4, 4, 6, 2, _
5, 5, 5, 8, 8, 6, 6, 6, 8, 8, 7, 9, 11, 9, 13, 10, 12, 13, 13, 12, 14, 13, 16, 14, 17, 15, 18, 15, 14, 15,
16, 14, 15, 13, 12, 10, 12, 11, 9, 9, 9, 11, 8, 9, 8, 8, 9, 7, 4, 4, 5, 5)
Dim fiftytoo
fiftytoo = Array (7, 8, 4, 5, 7, 6, 6, 8, 7, 6, 7, 11, 8, 9, 12, 13, 10, 13, 13, 13, 13, 13, 15, 13, 16,
16, 15, 16, 16, 14, 15, 15, 13, 14, 11, 10, 10, 11, 11, 9, 12, 10, 8, 10, 8, 6, 9, 6, 7, 4, 6, 4, 4)

chain = Request("chain")
groupID = Request("groupID")

dim record
record = chain
Q = chr(34)

sql = "SELECT * FROM groups WHERE ID = " & groupID
MyRecordSet.Open sql, Myconnection, adOpenKeyset, adLockOptimistic
weeks = MyRecordSet("weeks")
chains = MyRecordSet("chains")
MyRecordSet.Close()
' Declare a bunch of arrays to compute the inventory at each role
' Response.write weeks
dim demRet(70)
dim slRet(70)
dim elRet(70)
dim sd1Ret(70)
dim ed1Ret(70)
dim sd2Ret(70)
dim ed2Ret(70)
dim shiRet(70)
dim ordRet(70)

dim demWho(70)
dim slWho(70)
dim elWho(70)
dim sd1Who(70)
dim ed1Who(70)
```

```
dim sd2Who(70)
dim ed2Who(70)
dim shiWho(70)
dim ordWho(70)

dim demDis(70)
dim slDis(70)
dim elDis(70)
dim sd1Dis(70)
dim ed1Dis(70)
dim sd2Dis(70)
dim ed2Dis(70)
dim shiDis(70)
dim ordDis(70)

dim demFac(70)
dim slFac(70)
dim elFac(70)
dim sd1Fac(70)
dim ed1Fac(70)
dim sd2Fac(70)
dim ed2Fac(70)
dim shiFac(70)
dim ordFac(70)
  'Response.write "{ "&Q&"period"&Q&":
"&JSONDate(DateSerial(2014,2,j*7))&","&Q&"demand"&Q&": "& demnd &" }"&","
'initialize'
demRet(0)= 4  : slRet(0)= 12 : elRet(0) = 12 : sd1Ret(0)= 4 : ed1Ret(0)= 4 : sd2Ret(0) = 4  :
ed2Ret(0)= 4 : shiRet(0) = 4 : ordRet(0) = 4
demWho(0)= 4  : slWho(0)= 12 : elWho(0) = 12 : sd1Who(0)= 4 : ed1Who(0)= 4 : sd2Who(0) =
4  : ed2Who(0)= 4 : shiWho(0) = 4 : ordWho(0) = 4
demDis(0)= 4  : slDis(0)= 12 : elDis(0) = 12 : sd1Dis(0)= 4 : ed1Dis(0)= 4 : sd2Dis(0) = 4  :
ed2Dis(0)= 4 : shiDis(0) = 4 : ordDis(0) = 4
demFac(0)= 4  : slFac(0)= 12 : elFac(0) = 12 : sd1Fac(0)= 4 : ed1Fac(0)= 4 : sd2Fac(0) = 4  :
ed2Fac(0)= 4 : shiFac(0) = 4 : ordFac(0) = 4

Response.write "{ "&Q&"orderDataFeed"&Q&":[ "

 For j = 0 To 154 Step 1
  If (weeks = 52) then
     demnd = fiftytwo(j)
  Else
     demnd = 4
  End If
  Response.write "{ "&Q&"period"&Q&": "& (j+1) &","&Q&"demand"&Q&": "& demnd &" }"&","

 Next
```

```
  If (weeks = 52) then
    demnd = fiftytwo(155)
  Else
    demnd = 4
  End If
 changeSummary = ","&Q&"retailChange"&Q&": 12,"&Q&"wholesaleChange"&Q&":
12,"&Q&"distributorChange"&Q&": 12,"&Q&"factoryChange"&Q&": 12 "
 Response.write "{ "&Q&"period"&Q&": "& 156 &","&Q&"demand"&Q&": "& demnd &
changeSummary & "}"

sql = "SELECT * FROM " & groupID & " WHERE ID = " & record
MyRecordSet.Open sql, Myconnection, adOpenKeyset, adLockOptimistic

NoMoreEntries = False
first = true
k=-1
Do While Not (MyRecordSet.EOF OR NoMoreEntries)
   If
MyRecordSet("retailer")+MyRecordSet("wholesaler")+MyRecordSet("distributor")+MyRecordSet
("factory") = -200 then
      NoMoreEntries = True
   Else
      If k = -1 then
        k=k+1
      Else

          Response.write ","
          k = k+1
          If weeks = 52 then
            demnd = fiftytoo(k-1)
          Else
            demnd = usualData(k-1)
          End If
          Response.write "{ "&Q&"period"&Q&": "& (156+k) &","&Q&"demand"&Q&": "&demnd&","
          ' Response.write "{ "&Q&"period"&Q&":
"&JSONDate(DateSerial(2017,1,20+k*7))&","&Q&"demand"&Q&": "&demnd&","
          order = myRecordSet("retailer")-100
          ordRet(k) = order
          Response.write " "&Q&"retailer"&Q&":"&order&","
          order = myRecordSet("wholesaler")-100
          ordWho(k) = order
          Response.write Q&"wholesaler"&Q&":"&order&","
          order = myRecordSet("distributor")-100
          ordDis(k) = order
          Response.write Q&"distributor"&Q&":"&order&","
          order = myRecordSet("factory")-100
          ordFac(k) = order
```

```
        Response.write Q&"factory"&Q&":"&order&", "
        If k=1 then
          demWho(k)= 4 : demDis(k) = 4  : demFac(k) = 4
        Else
          demWho(k)= ordRet(k-2) : demDis(k) = ordWho(k-2)  : demFac(k) = ordDis(k-2)
        End If
        demRet(k)= demnd
        slFac(k)= elFac(k-1) : sd1Fac(k) = ed1Fac(k-1) : elFac(k)= slFac(k)+sd1Fac(k)-
demFac(k) :  ed2Fac(k)= ordFac(k-1) : sd2Fac(k) = ed2Fac(k-1) : ed1Fac(k)= sd2Fac(k)
        shiFac(k) = min(max(slFac(k),0)+sd1Fac(k),demFac(k)-min(0,slFac(k)))
        slDis(k)= elDis(k-1) : sd1Dis(k) = ed1Dis(k-1) : elDis(k)= slDis(k)+sd1Dis(k)-demDis(k) :
ed2Dis(k)= shiFac(k) : sd2Dis(k) = ed2Dis(k-1) : ed1Dis(k)= sd2Dis(k)
        shiDis(k) = min(max(slDis(k),0)+sd1Dis(k),demDis(k)-min(0,slDis(k)))
        slWho(k)= elWho(k-1) : sd1Who(k) = ed1Who(k-1) : elWho(k)= slWho(k)+sd1Who(k)-
demWho(k) :  ed2Who(k)= shiDis(k) : sd2Who(k) = ed2Who(k-1) : ed1Who(k)= sd2Who(k)
        shiWho(k) = min(max(slWho(k),0)+sd1Who(k),demWho(k)-min(0,slWho(k)))
        slRet(k)= elRet(k-1) : sd1Ret(k) = ed1Ret(k-1) : elRet(k)= slRet(k)+sd1Ret(k)-demRet(k)
:  ed2Ret(k)= shiWho(k) : sd2Ret(k) = ed2Ret(k-1) : ed1Ret(k)= sd2Ret(k)
        shiRet(k) = min(max(slRet(k),0)+sd1Ret(k),demRet(k)-min(0,slRet(k)))
        Response.write Q&"retailInv"&Q&": "&elRet(k)&","&Q&"retailChange"&Q&": "&(elRet(k)-
slRet(k))&","
        Response.write Q&"wholesaleInv"&Q&": "&elWho(k)&","&Q&"wholesaleChange"&Q&":
"&(elWho(k)-slWho(k))&","
        Response.write Q&"distributorInv"&Q&": "&elDis(k)&","&Q&"distributorChange"&Q&":
"&(elDis(k)-slDis(k))&","
        Response.write Q&"factoryInv"&Q&": "&elFac(k)&","&Q&"factoryChange"&Q&":
"&(elFac(k)-slFac(k))&"} "
      End If

   End If
   record = record*1 + chains
   MyRecordset.Close
   sql = "SELECT * FROM " & groupID & " WHERE ID = " & record
   MyRecordSet.Open sql, Myconnection, adOpenKeyset, adLockOptimistic
Loop

 Response.write "]  }"
'  Response.write "done..."
MyRecordset.Close
Set MyRecordset = Nothing
MyConnection.Close
Set MyConnection = Nothing
%>
```

REFERENCES

Forrester, J. (1961). *Industrial dynamics*. Cambridge, MA: MIT Press.

Senge, P. (2006*) The Fifth Discipline: The Art and Practice of the Learning Organization*, New York, NY., Doubleday.

Sterman, J. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*, Boston, MA., McGraw-Hill.

DECISION SCIENCES INSTITUTE

Is Milk as good as Beer? An Extension to Beer Distribution Simulation

Vivek Bhatt
Ahmedabad University
E-mail: vivek.bhatt@ahduni.edu.in

Mitesh Sharma
Ahmedabad University
E-mail: mitesh.s.bba16@ahduni.edu.in

Priyanshi Kothari
Ahmedabad University
E-mail: Priyanshi.k.bba16@ahduni.edu.in

Leena Bhagchandani
Ahmedabad University
E-mail: leena.b.bba16@ahduni.edu.in

ABSTRACT

Beer distribution simulation has served the purpose of understanding the bullwhip effect caused by communication lack among supply chain members. This paper proposes to extend the simulation with raw material and multiple products, with clarified butter (ghee), milk powder, and condensed milk as products using milk. Interestingly, the bullwhip intensifies for all three products as a result of limited raw material, which is supplied proportionately in response to the demand placed by respective product manufacturer.

KEYWORDS: Supply Chain, Beer Distribution Game, Extension of simulation, Simulation for teaching

INTRODUCTION

Beer-distribution game developed by professors at MIT has gained huge importance in supply chain and operations management classrooms. The simulation, crisply, demonstrates how the supply chain operates and thereby also helps in comprehending bullwhip effect occurring simultaneously in the supply chain. Here the over estimation of customer demand causes this bullwhip effect. Penalty for access or fewer inventories is charged. Even today, the beer game serves the purpose of demonstrating the inherent problems involved in supply chain management. However, the simulation has to adapt to changes to remain relevant and to display a better picture of the supply chain in other industries also. The proposed study asks the following changes to make the original model look more suitable to the current scenario:

a) The simulation focuses on the distribution of milk and milk products.
b) Dairy farm or association is the raw material provider, which serves to Milk Powder, Condensed Milk, and Clarified Butter industries.
c) Unlike the original version, manufacturers are not allowed to produce as per their wish, since there is limited raw material availability.

d) The raw material is supplied in the proportion of orders placed by three manufacturers, so the manufacturers have to compete with the other two industries to receive raw material.

LITERATURE REVIEW

Over the period of time, with the evolution of business environment; the understanding and definition of Supply Chain Management has evolved significantly. Thereby, its understanding and Implications have also broadened. Supply Chain Management (SCM) refers to overall management of the materials; from their procurement to the delivery. Hence, the process involves the functions of acquiring, processing, transporting and delivering the final product to the end user. Essentially, SCM studies the inter-relations and inter- linkages amongst the members involved, processes undertaken and functions performed within the Chain. The implications of the interactions that take place within the Supply Chain; have larger consequences on the profit maximization; due to value addition involved at each stage. New Dimensions such creating and nurturing long-term relationships amongst the members, strategic standpoint with respect to competitiveness, smooth flow of information, developing flexibility in the complex market and the element of e-procurement, have been added to the area of Supply Chain Management; besides its primary function of Cost Reduction. (Shilpa Parkhi, 2015)

Simulation Modeling is to create a model demonstrating the real-world situation, in order to analyse it and predict the results. The (Nance & Sargent, 2002) paper describes various categorizations of various simulation tools and technique that have evolved in around 50 years of simulation study. Out of this discrete event modeling has been greatly used by different areas of work. The method of Simulation has been used in various areas of work such as defence, communication, manufacturing, commerce; etc. (Walker, Giddings, & Armstrong, 2011)

With increasing scope of Supply Chain Management; Simulation is widely used by researchers to unveil complex problems present in the chain. Use of simulation tool helps to critically analyse the complexity involved SCM decisions and is also helpful in understanding the rationale of the decision makers in adopting a policy or a process. (Spagnoletti, D'Atri, & D'Atri, 2013)

Beer Distribution Game was developed by MIT Sloan System Dynamics Group; to illustrate all the hidden problems underlying the supply chain management system. The game demonstrates the ‘bullwhip effect”; commonly known as the over-estimation of demand from a consumer to the vendor. Beer Distribution Game aims at minimizing the cost at each level of the supply chain; by lowering the inventory held and timely delivery of orders.

Simulation modeling in the form of Beer Distribution Game is widely used for experiential learning in the dynamics of the supply chain management. (Jose Costas, 2015) The selfefficacy and Attitudes of the students studying supply chain management with BDG improved positively by role-play simulation. (Juthamas Choomlucksana, 2011).

The Beer Distribution Game - Variations and Extensions

The Beer Distribution Game has also been modified with different variations and extensions by researchers.

The strategies of providing access to Global Information and a shorter lead time, to all the members in the supply chain, reduced costs within the supply change and enabled greater control over inventory. (Muhammad Zeshan Rafique, 2014)

The traditional beer distribution game has certain shortcomings that do not reflect the realistic supply chain environment. Many problems that do arise in SCM can be resolved by shortening the delivery time and centralizing decision making. With this argument, the paper proposes a computerized beer distribution game with different policy options and settings. This new game has only a few exceptions to the traditional BDG; human player assumes only one of the roles of the BDG while the rest are managed by the computer. Three option settings in the game are: A short lead time, Centralize information and the Global Information option. Simulation of the new computerized BDG, offer an enhanced explanation of The Bullwhip Effect, The Centralization Effect and the Short Lead Time Effect. (Levi)

The (Hongliang Liu, 2009) simulates the Beer Distribution Game in four complex demand patterns namely, one step demand (demand changes only once in entire simulation), stationary demand (pattern follows normal distribution), uniform demand (random fluctuations in a specified range) and cyclical demand (seasonal variations). The variations proposed in the traditional game illustrate a more realistic picture of supply chain by the means of the improvised Beer Distribution Game. (Hongliang Liu, 2009)

The (Chen & Samroengraja, 2000) is an improvised version of the Beer Distribution game in terms of the proposed changes in the demand pattern. All the players within the system have prior information about the demand distribution, wherein the demand is independent and identical in different periods. The rationale of the improvisation is in-line with the real-world conditions. Also, the stationary game provides theoretical benchmark.i.e. optimal levels of inventory and costs that can be achieved in the game.

The paper (Sarkara, 2012) aims to enable the decision-making process and risk management wherein the inventory cost is not compromised. It also tries to avoid the bullwhip effect with a location independent model which is capable of acquiring the demand and the supply orders. The outcome of this study was that they were able to eliminate the bullwhip effect without any backlogs with help of the Petri Net Model that they used for their study.

The paper (Norina Szander, 2013) considers that the traditional efficiency method of evaluating the success of a supply chain by just using financial metrics is not complete. It thus takes into account four (financial, customer, internal, innovation and learning) perspectives of Balanced Scorecard performance measurement system. It accomplishes this by playing the beer distribution game. To conclude, while examining the performance of the supply chain, one should always consider the logistics trade-offs as well.

The Bullwhip Effect

The identification of the bullwhip effect helped in making improvements in the supply chain management. The Bullwhip effect has also been studied with different dimensions and measures have been suggested to control it. The bullwhip effect is not only the result of operational shortcomings in the supply chain instead it has also a lot to do with the decision maker’s tendency to undermine the supply chain. (Croson Rachel, 2006) The manager is unable to reduce the new orders considering the ones he has already placed but are yet not delivered. Also, (Croson Rachel, 2006) demonstrates that information sharing; a proposed solution to the bullwhip effect; though successful in doing so, benefits are skewed towards the upstream members of the supply chain.

A study carried out in (Shilpa Parkhi, 2015). presents an agent-based simulation which evaluates the bullwhip effect in SC through a series of numerical experiments in two scenarios. The first is two tiers SC and the other is four tier SC. The results of the experiments were compared with the analytical counterparts in literature. The results show that the assumption of non-negative demand or order quantities should be included for the accurate and realistic measurement of bullwhip effect. The four tier shows how bullwhip effect changes when we move upward in the supply chain.

Four causes of the Bullwhip effect have been identified by various research studies, namely, Demand signal processing (natural demand forecasting) , order batching (order frequency and evenness of order arrivals), price fluctuations (short-term changes in price as a part of sales promotions) and shortage gaming (order cancellation policy). (Hau L. Lee V. P., 2004) The paper (Hau L. Lee V. P., 2004), also suggests the need for more research in the area of “bullwhip effect” in order to develop practical solutions in terms of processes and policies by working with the practitioners.

Measures such as information sharing models, establishing strategic alliance, etc. would help in reducing the bullwhip effect. (Hongchun Wang, 2011) The (Frank Chen, 2000) proposes framework to quantify the bullwhip effect with different number channel members. With sharing demand information at each stage of the supply chain; can reduce the variability in orders placed but does not eliminate the bullwhip effect from the system. (Frank Chen, 2000)

The spreadsheet solution illustrates how the bullwhip effect in the supply chain can be minimized with greater inventory fluctuations and a reduced customer service. (Robert N. Boute,
2009)

In (S. Kamal Chaharsooghi, 2008) Bullwhip effect is studied by altering the forecasting methods used for demand forecast; the core factor to the bullwhip effect. Methods such as Time series forecast, moving average and Exponential Smoothening are used. The results suggest that the accurate the forecast; does not necessarily lower the bullwhip effect. (S. Kamal Chaharsooghi, 2008)

The bullwhip effect paralyzes the supply chain management and the penalties in terms of higher costs are borne by all its members. However, due to various incentives and promotional offers made by upstream members to downstream members (wholesaler’s discounting pricing to retailer, the option to return the unsold goods, etc.) the burden of penalty is borne more by the upstream players. Such outcomes do not provide any incentive to the downstream member (especially the retailer) to not exaggerate demand. (Hau L. Lee V. P., 2016)

The (Saitoh, 2014) uses a knowledge sharing simulation model (Elman Network) that aims to improve the decentralized information sharing and thereby controlling the bullwhip effect within the chain. The data about inventory volumes, demand forecast, etc. at each stage is provided as input, in the form of knowledge, to the model of the next stage. The study proposes a decrease in the magnitude of the bullwhip effect within the supply chain.

The (Wangphanich, Kara, & Kayis, 2007) proposes a model that illustrates The Bullwhip effect in a Supply chain network with three retailers to one manufacturer; three different products and assuming same forecasting technique. The model has been validated using ‘ROO Water’ and hence can be used by decision makers for real life SCM.

All these studies have introduced different dynamics to the problems arising in a supply chain and the corresponding bullwhip effect.

We aim to study the bullwhip effect arising out of introducing a new level in the supply chain and extending the game to multiple products.

SIMULATION DESIGN

The Milk Distribution Game developed during this study consists of 3 supply chains of three different products from the same industry i.e. Dairy, with a common raw material provider, who is the farmer. The three products are Ghee, Condensed milk and milk powder. It has thirteen players, Retailer, Wholesaler, Distributor and Manufacturer for the three supply chains and a farmer.

The participants are asked to sit in columns whereby each column represents a supply chain. There are 3 such columns with 4 levels in each supply chain. There is 1 farmer who caters to the demands of all the manufacturers of supply chains.

Of the total units of milk produced, 54% is used for these three-supply chain and the remaining 46% is utilized as fluid milk for domestic consumption.

The entire supply chain of ghee will have 9 units in their inventory, delay 1 and delay 2. Similarly, condensed milk's supply chain will have 2 units in their inventory, delay 1 and delay 2. Whereas, there will be 1 unit in the inventory, delay 1 and delay 2 of the milk powder's entire supply chain.

The retailer enters the order received manually according to the instructions given by the instructor of the game in the ratio of 9 or 2 or 1 unit(s), depending on the supply chain he/she is in. The demand is catered and the backlog is recorded automatically. According to the order received the participant places a relevant order which is then catered by the wholesaler.

The wholesaler, distributor and the manufacturer follow the suit. The demand is catered and the backlog is recorded automatically. According to the order received the participant places a relevant order to the next member of the supply chain.

The farmer on receiving orders from all the three suppliers; divides the limited supply of milk between them in the ratio of 9:2:1. The number of milk units a farmer can supply is 314 with a standard deviation of 10% (i.e. 31.4 units) in a normal distribution curve. These get converted into respective product units based on milk requirement in each product.

Demand at the retailer was determined on the basis of market share which are 1, 2, 9 for Milk powder, Condensed milk and Ghee respectively (Khamkar, 2014). In this simulation, the demand at the retailer level increases to double once during the 25 weeks' time, at this time the demand at retailer level increases to 2, 4, and 18 respectively. Though the percentage change is same, the impact on Ghee at the time of increase in demand will be higher because the change in absolute terms is very high.

Also, the conversion of 1 unit of product in terms of the milk units required is 8,3,16 for Milk powder, Condensed milk and Ghee respectively. So the milk required for production of Ghee is very high, this also contributes to the high bullwhip. Since the milk unit's requirement of Milk powder and Ghee is higher than that of condensed milk, condensed milk's supply chain suffers when the demand is high, and supply is limited and done in proportion to the demand received by the farmer.

FINAL ANALYSIS

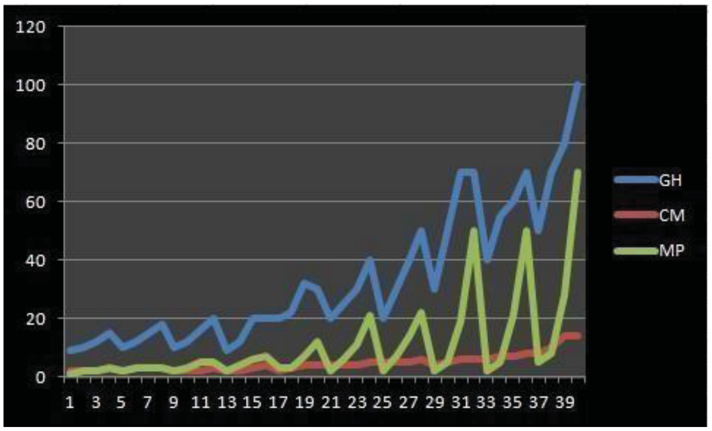
The objectives of the research are to analyze:

1. Observing the change that occurs in the bullwhip effect when a level is added.
2. Examining the cross-section impact on companies which depend on same raw material.

Bullwhip Graphs

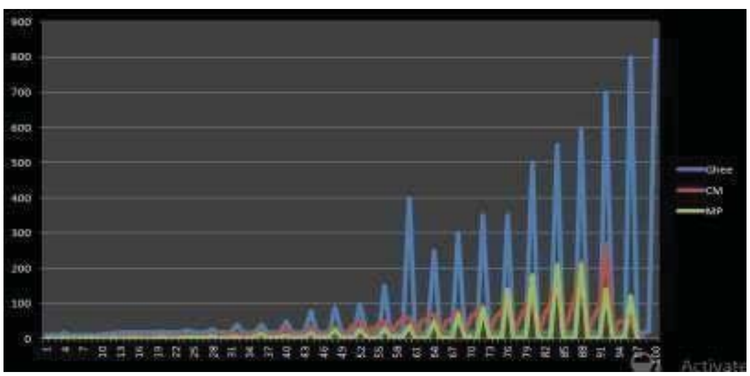
At first, we conducted the experiment with single supply chains of Ghee, Milk Powder and Condensed Milk by adding a raw material (limited supply) to the chain.

Figure 1: Single Product Bullwhip



Secondly, the experiment was conducted with all the three chains simultaneously working with milk as their common raw material. The following graphs show the same: -

Figure 2: Multiple Products Bullwhip Graph 1



Bhatt, Sharma, Kothari, Bhagchandani Is Milk as good as Beer?

Figure 3: Multiple Products Bullwhip Graph 2

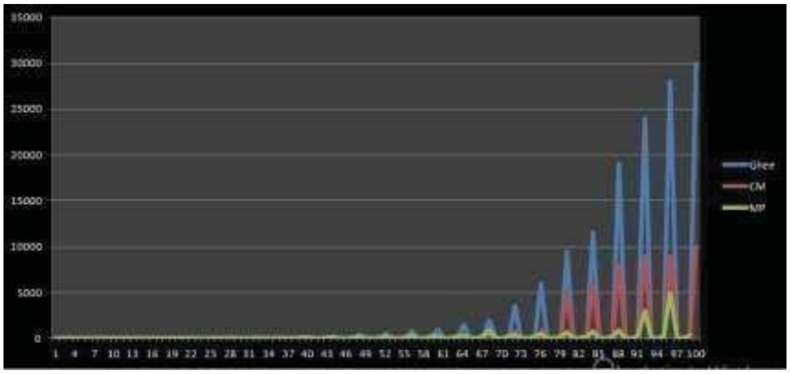
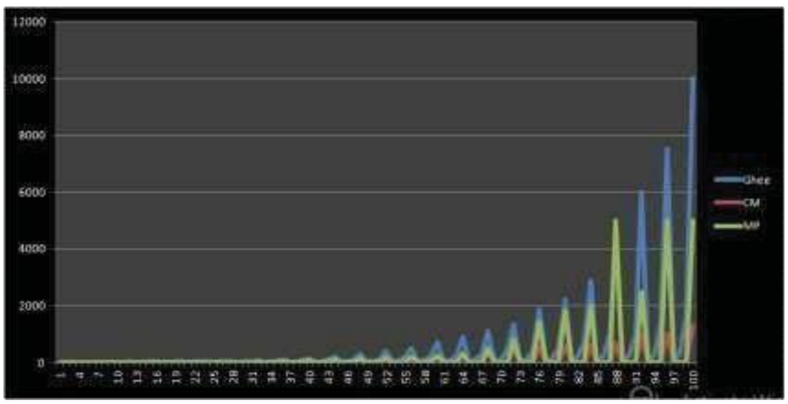


Figure 4: Multiple Products Bullwhip Graph 3



The bullwhip graphs suggest a significant rise in all the three supply chains when compared to the one wherein they were functioning independently. The intensity of the bullwhip significantly rises when there is more than on product in the market.

The demands' variation was tested for significance of difference with difference of variance test, and we found the following results, for three experiments.

Experiment 1

Table 1: F--Test for Condensed Milk Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	CM	*CM*
Mean	4.65	79.58
Variance	8.643589744	51634.24606
Observations	40	100
df	39	99
F	0.0001674	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 2: F-Test for Ghee (Clarified Butter) Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	GH	*GH*
Mean	33.1	526.4
Variance	546.1948718	2173279.616
Observations	40	100
df	39	99
F	0.000251323	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 3: F-Test for Milk Powder Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	MP	*MP*
Mean	10.7	256.52
Variance	226.0615385	858861.4844

Observations	40	100
df	39	99
F	0.000263211	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 4: F-Test for Condensed Milk Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	CM	*CM*
Mean	4.65	38.07
Variance	8.643589744	2872.712222
Observations	40	100
df	39	99
F	0.00300886	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 5: F-Test for Ghee (Clarified Butter) Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	GH	GH
Mean	33.1	76.12
Variance	546.1948718	28493.70263
Observations	40	100
df	39	99
F	0.019168968	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 6: F-Test for Milk Powder Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	MP	*MP*
Mean	10.7	16.86
Variance	226.0615385	1745.131717
Observations	40	100
df	39	99
F	0.129538382	
P(F<=f) one-tail	1.29367E-10	
F Critical one-tail	0.625719655	

Table 7: F-Test for Condensed Milk Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	CM	*CM*
Mean	4.65	505.99
Variance	8.643589744	3670058.293
Observations	40	100
df	39	99
F	2.35516E-06	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 8: F-Test for Ghee (Clarified Butter) Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	GH	*GH*
Mean	33.1	1395.28
Variance	546.1948718	27328154.43

Observations	40	100
df	39	99
F	1.99865E-05	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

Table 9: F-Test for Milk Powder Demands

F-Test Two-Sample for Variances		
	in isolation	**with other products**
	MP	*MP*
Mean	10.7	239.93
Variance	226.0615385	820883.5809
Observations	40	100
df	39	99
F	0.000275388	
P(F<=f) one-tail	0	
F Critical one-tail	0.625719655	

The results of variance difference show that there occurs a significant impact on the orders placed by the agents of supply chain, when they know that there exist other products in the market and there is a common raw material provider with limited capacity. Raw material is supplied in proportion to the demand received by the farmer

The cross-sectional impact is exponentially high; participants were demanding more than what they require because they knew that the raw material supply is limited, and they will get the supplies proportionate to their demand.

CONCLUSION

The raw material supply is limited and supplied to all the three industries proportionate to the demand received by the farmer. Therefore, each industry demands greater quantities than required, in order to ensure that when divided proportionately; each gets a better share. Since this logic is followed by all the three supply chains, the resultant demand orders rise significantly each subsequent time.

The results of the experiment show a significant increase in the bullwhip effect of each of the supply chains operating within the industry with a limited raw material supply.

REFERENCES

Chen, F., & Samroengraja, R. (2000). The stationary beer game. *Production and Operations Management; Muncie*.

Croson Rachel, D. K. (2006). Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information. *Management Sciences*, 323-336.

Frank Chen, Z. D.-I. (2000, March). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Science, 46*(3), 436-443.

Hau L. Lee, V. P. (2004). Comments on "Information Distortion in a Supply Chain: The Bullwhip Effect": The Bullwhip Effect: Reflections. *Management Sciences*, 1887-1893.

Hau L. Lee, V. P. (2016). Bullwhip Effect in Supply Chains. 1-11. Sloan Management Review.

Hongchun Wang, B. H. (2011). Research on the Reducing Measures of Bullwhip Effect. *2011 International Conference on Software and Computer Applications.* Beijing: IACSIT Press.

Hongliang Liu, E. H. (2009). Optimisation of the Beer Distribution Game with complex customer demand patterns. *2009 IEEE Congress on Evolutionary Computation.* IEEE.

Jose Costas, B. P. (2015). Agents playing the Beer Distribution Game: Solving the Production
Dilemma through the Drum-Buffer-Rope Methodology. *ICIEOM 2015 XXI International Conference on Industrial Engineering and Operations Management, At Aveiro (Portugal)*, (pp. 18). Aveiro, Portugal.

Juthamas Choomlucksana, T. L. (2011). An Investigation of Self-efficacy and Attitudes using a Role-playing Simulation. *Norcross* (p. 8). Corvallas: Institute of Industrial and Systems Engineers (IISE).

Khamkar, S. K. (2014). The Consumption Pattern of Dairy Products by Indian. *Asian Journal of Management Studies*, 170-172.

Levi, P. K. (n.d.). *A New Computerized Beer Game.* Retrieved from kaminsky ieor berkeley: http://kaminsky.ieor.berkeley.edu/Reprints/PK_DSL_98b.pdf

Muhammad Zeshan Rafique, W. S. (2014). The improvements in supply chain management through beer game. *The international journal of scientific and research publication*, 1-4.

Nance, R. E., & Sargent, R. G. (2002, Jan/Feb). Perspectives on the evolution of simulation. *Computers*. Linthicum, United States: Institute for Operations Research and the Management Sciences.

Norina Szander, K. L. (2013). BeerGame reference scenarios for Balanced Scorecard evaluation, highlighting customer perspective. *1st Logistics International Conference, At Belgrade* (pp. 1-2). Belgrade: Logic.

Robert N. Boute, M. R. (2009, July). Exploring the bullwhip effect by means of spreadsheet simulation. *Operations and Technology Management Center*. Leuven, Belgium: INFORMS Transactions on Education.

S. Kamal Chaharsooghi, H. F. (2008). A simulation study on the impact of forecasting methods on the bullwhip effect in the supply chain. *2008 IEEE International Conference on Industrial Engineering and Engineering Management.* Singapore: IEEE.

Saitoh, F. (2014). The Knowledge Sharing Model on Supply Chain Simulation using Recurrent Neural Network. *2014 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 1332-1336). Bandar SunWay: IEEE.

Sarkara, B. B. (2012). A Distributed Retail Beer Game for Decision Support System . *Procedia - Social and Behavioral Sciences*, 278-284.

Shilpa Parkhi, S. J. (2015, may). A Study of Evolution and Future of Supply Chain Management. Nashik, India: AIMS International.

Spagnoletti, P., D'Atri, E., & D'Atri, A. (2013). DECISION MAKING IN SUPPLY CHAINS AND VALUE
NETWORKS: THE BEER GAME EVOLUTION(a). *International Journal of Electronic Commerce Studies,* 63-78.

Walker, W. E., Giddings, J., & Armstrong, S. (2011, Sep). Training and learning for crisis management using a virtual simulation/gaming environment. *Cognition, Technology & Work; London*. Netherland, London: Springer Science & Business Media. Retrieved from https://search.proquest.com/docview/883016553?accountid=37713

Wangphanich, P., Kara, S., & Kayis, B. (2007). A simulation model of bullwhip effect in a multistage supply chain. *2007 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 1853-1857). Singapore: IEEE.

DECISION SCIENCES INSTITUTE
Empirical Analysis of Social Media Effects on Supply Chain Management:
Social Network Perspective

Young Sik Cho
Jackson State University
Email: young_sik.cho@jsums.edu

Nero Edevbie
Jackson State University
Email: nero.edevbie@students.jsums.edu

Antonio K.W. Lau
Kyung Hee University
Email: antoniolau@khu.ac.kr

ABSTRACT

From a network perspective, this study conceptualized the role of social media in forming social networks among a focal company, customers, and suppliers. Data analysis based on samples from 117 U.S. companies demonstrated that social media usage had a significant impact on the operational efficiency of SRM and CRM systems.

KEYWORDS: Social media, Supplier relationship management, Customer relationship management, Empirical research, Social network theory

INTRODUCTION

Over the past few decades, the essence of inter-firm competition has shifted to supply chain management (Croom et al., 2000; Cousins et al., 2006; Gunasekaran et al., 2008). Previous research has shown that the supply base must be managed efficiently and carefully to avoid losing competitive advantage in the marketplace (Choi and Krause, 2006; Nam et al., 2011). Supply chain management is no longer regarded merely as a means of logistics and is seen as a key to improving the core results of customer satisfaction (Hult et al., 2007; Lyons et al., 2013). Today, companies are also faced with a new mission to improve organizational performance by combining existing management systems, processes, and capabilities with their social networking skills and technology (Trainor et al., 2014). Since the Internet revolution of the 1990s, the way we communicate has changed dramatically, and now social media has become an essential concept in most business models (Kaplan and Haenlein, 2010). According to the Nielsen Company's (2010) survey, US consumers spent nearly a quarter of their Internet time on social networking sites. As of January 2018, active social media users are estimated at 3.2 billion and 42% of the world's population (Shaw, 2018). Organizations now use social media platforms as tools to improve profitability, identify problems, promote innovation, interact with customers, respond quickly to what they want, and improve organizational performance (Gallaugher and Ransbotham, 2010; Rodriguez et al., 2012). Social media also enables employees to seamlessly communicate information and knowledge within an organization by facilitating customer interactions with external organizations, as well as providing job updates and best practices through improved business intelligence (Wu, 2013; Lam et al., 2016).

Although "social media" has received considerable attention in a number of academic communities, including Marketing (Fader and Winer, 2012) and Information Systems (Aral et al., 2013), it is still at an early stage in Operations and Supply Chain Management research. Further, it is reported that research on social media still remains incomplete and existing literature does not provide sufficient empirical evidence of the impact of social media on organizational performance (Parveen et al., 2015). Therefore, the purpose of this study is to explore how the use of social media impacts supplier and customer relationship management. In order to empirically investigate this research question, we collected data from 117 companies in the United States through surveys. CFA was performed to assess the validity of the measurement scales. Further, this study adopted the insomnia scale as “marker variable” to minimize the likelihood of a common method variance (CMV) in our data and also performed “common latent factor” analysis to diagnose the potential threat of CMV. Finally, the hypotheses developed in this study were examined using structural equation modeling. The results showed that social media platform was positively associated with social media capability and the usage of social media had a significant positive effect on both supplier relationship management (SRM) and customer relationship management (CRM). The results also indicated that CRM directly affected customer satisfaction, but the impact of SRM practices on organizational performance could be achieved positively when combined with the efficient operation of the CRM system.

This study contributes to the literature on social network perspective (Freeman, 2004; Choi and Kim, 2008; Borgatti et al., 2009) by conceptualizing the role of social media in forming social networks between a focal company, customers and suppliers. This study also contributes to the knowledge-based view of the firm (Grant, 1996; Spender, 1996) by demonstrating that the social media technology embedded in supply chain information systems (i.e., CRM and SRM) has become a strategic resource to maintain superior organizational performance.

THEORY AND HYPOTHESES

The Effect of Social Media on CRM

Customer relationship management (CRM) is described as “an information system that collects data from a number of customer-facing activities to help an organization better understand its customers so that it can better match its products and services to customer needs and thereby increase sales” (Hill, 2017, p. 104). CRM is also viewed as “a core organizational process that focuses on establishing, maintaining, and long-term enhancing associations with customers” (Jayachandran et al., 2005, p. 3). Namely, CRM can be understood as an integrated system of people, processes, technologies and all business activities to improve customer relationships (Anton, 1996; Anton and Hoeck, 2002; Feinberg and Kadam, 2002; Chen and Popovich, 2003; Jayachandran et al., 2005).

Social media is described as “an online resource that people use to share content: video, photos, images, text, ideas, insight, humor, opinion, gossip, news” (Drury, 2008, p. 274). Social media is used for information sharing and seeking, entertainment, and expressing an opinion about a product or an issue (Whiting and Williams, 2013). Examples of popular social media platforms in the industry include Blogger, Twitter, Facebook, LinkedIn, Flickr, YouTube, Ning, and Igloo. Previous research has shown that social media can be used to improve financial performance (Du and Jiang, 2014; Ainin et al., 2015), organizational learning (Vuori and Okkonen, 2012; Wu, 2016), employee performance (Cao et al., 2016; Cetinkaya and Rashid, 2018), and innovation (Lam et al., 2016; Dodokh and Al-Maaitah, 2019). Rodriguez et al. (2015) claimed that social media positively guided customer orientation activities and ultimately affected sales and business performance. Trainor et al. (2014) found that using social media could improve the efficiency of the company by promoting internal and external collaboration. Lam et al. (2016)

found that social media initiatives enhanced operational efficiency and promoted corporate innovativeness. Table 1 summarizes major findings on the impact of social media usage on various organizational performance over the past decade.

Social media has become an essential method for people to communicate with their customers (Sano and Sano, 2016). Corte et al. (2015) showed that companies use social media to track customer feedback and product loyalty. According to Lee et al. (2015), social media is also useful for information dissemination during product recalls. Trusov et al. (2009) showed that social media usage could enhance customer relationships by increasing opportunities for interactive conversations. In a similar vein, Parveen et al. (2015) reported that social media could improve customer relationships by increasing access to information. Acker et al. (2011) argue that if social media is effectively managed and used with CRM, it can contribute positively to customer satisfaction as well as financial rewards. Evidence from previous research also suggests that social media can help improve customer relationships by enhancing customer trust through information sharing (Maecker et al., 2016; Khan et al., 2017). Therefore, based on previous evidence, it is reasonably inferred that the use of social media will have a positive effect on customer relationship management:

H1. *The use of social media has a positive impact on customer relationship management.*

Table 1: Key findings on the effects of social media use on organizational performance over the past decade

Article	Purpose	Discipline	Sample	Findings
Luo *et al.*, 2013	Explore how social media affects firms' equity value.	Information System	9 hardware and software companies	Social media is a primary predictor of a company's equity value.
Abzari *et al.*, 2014	Investigate the effects of social media on consumers' attitude towards company brand.	Marketing	210 automaker customers	Social media has a positive influence on customers' buying intentions.
Du and Jiang, 2014	Examine the relationship between social media and firm performance.	Information System	S&P 1500 companies	Companies using social media achieve better financial results.
Parveen *et al.*, 2015	Investigate the impact of social media usage on organizational performance.	Marketing	Interviews with senior managers of 6 companies	Social media improves customer relationships, improves access to information, and reduces marketing costs.
Ainin *et al.*, 2015	Study the effects of social media on the performance of small and medium enterprises (SMEs).	Information System	259 SMEs owners in Malaysia	Social media has a positive effect on customer service, access to information, and financial performance.

Lam *et al.*, 2016	Explore how firms' social media initiatives affect operational improvement.	Operations Management	Longitudinal data between 2006 and 2012 from Compustat and Fortune; 281 firms' social media initiatives between 2006 and 2011.	Social media initiatives have a positive impact on operational efficiency and innovativeness by enhancing firms' knowledge-based advantages (e.g., information flow and knowledge sharing).
Wu, 2016	Study the role of social media strategies in organizational performance.	Marketing	327 chain and franchise stores in Taiwan.	Organizational culture, leadership, and social networks relate to social media strategies that have a positive impact on performance.

The Effect of Social Media on SRM

Strategic supply management has become a critical requirement in today's business environment, especially with multiple purchasing transactions with various suppliers (Wu and Choi, 2005; Moeller et al., 2006; Yeung, 2008). In particular, supplier relationship management (SRM) focuses on strengthening relationships with suppliers to improve collaboration, processes, quality and forecasting while reducing inventory, risk and costs (Hill, 2017). The primary goal of SRM is to ensure that information and materials flow smoothly throughout the supply chain in an effective manner (Feldmann and Müller, 2003; Childerhouse and Towill, 2003). The benefits of using SRM are tremendous. Prior research shows that through active collaboration with suppliers, SRM helps companies reduce unnecessary inventories and costs, enhance operational efficiency, minimize price volatility, and improve information flow (Jap, 1999; Shin et al., 2000; Swink et al., 2005; Li et al., 2006). As such, an effective SRM implementation is considered a prerequisite for optimizing the company's operations to maintain a competitive advantage (Song et al., 2014; Handfield, and Bozarth, 2015). According to the literature, in the early stages of product development, companies can benefit from using virtual communities by engaging users with similar interests in the development process and sharing experiences and knowledge (Hagel, 1999; Kozinets, 1999). Swain and Cao (2013) claim that partners in the supply chain can improve operational productivity and performance as they use more social media. Lam et al. (2016) also argue that social media enables companies to improve intelligence across their supply chain networks through a visual approach to external knowledge and information. In short, previous research points out the important role of social media in strengthening partnerships with suppliers through sharing knowledge and information. Previous research also argues that social media can improve customer relationships by enhancing customer confidence through information sharing (Munar and Jacobsen, 2013; Elena 2016).

This claim can be better understood by the lens of social network perspective (Freeman, 2004; Choi and Kim, 2008). In network analysis, network systems are represented by a set of nodes representing interrelated actors (such as individuals, organizations, and communities) and by lines representing exchange relationships (Borgatti et al., 2009; Kim et al., 2015). Thus, as shown in Figure 1, the structural relationships between a focal company, suppliers, and customers in CRM and SRM systems can be diagrammed from the networks perspective (Choi and Krause, 2006). Here, it is likely that social networking already exists between the focal firm and individual customers and suppliers through the operations of CRM and SRM systems. However, the relationship among individual suppliers as well as among individual customers is likely to still remain as two separate "clusters" based on similarities rather than social networks because there are no open channels for mutual information exchange and interaction. On the other hand, Figure 2 depicts social networks that have been expanded and enhanced due to social media effects on CRM and SRM systems. Here, customers and suppliers can share and interact with information and knowledge through the social media platform, which leads to form "social networks" based on "soft" ties (e.g., information sharing and favoritism) rather than "hard" ties (e.g., money and materials flow) (Borgatti and Li, 2009). As a result, enhanced social networks among the customers and the suppliers are expected to benefit to the focal company because the actor at the top of the network chain has the greatest information advantage when the type of tie is "soft" according to social network theory (Borgatti and Li, 2009). Therefore, from the network perspective, social media platforms enable customers and suppliers to form social networks, ultimately enhancing the efficiency of supply chain information systems such as SRM and CRM. Considering all the above arguments, the following hypothesis is proposed:

H2. *The use of social media has a positive impact on supplier relationship management.*

Figure 1: Social Networks in CRM and SRM Systems

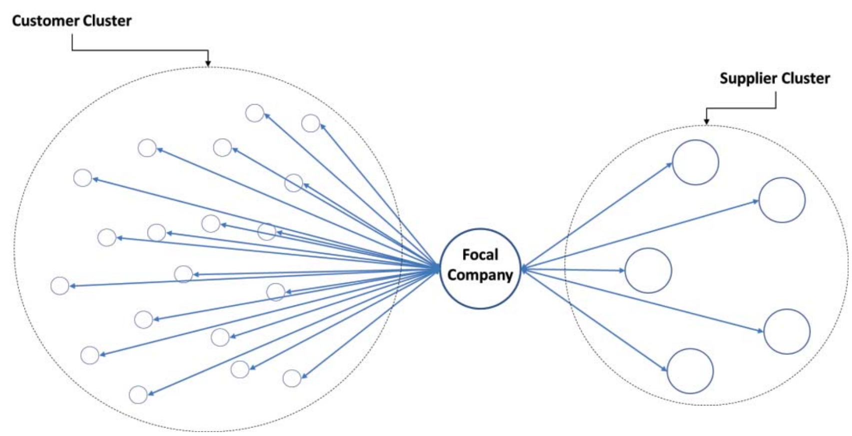
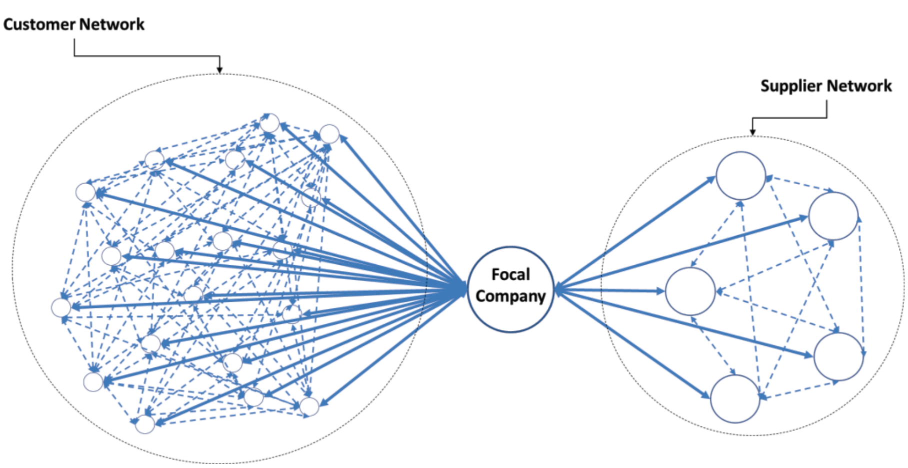


Figure 2: Adapted Social Networks based on Social Media Effects on CRM and SRM Systems



CRM, SRM, and Organizational Performance

Evidence supports that CRM motivates employees to remove functional barriers and encourage information sharing to enhance customer satisfaction (Chen and Popovich, 2003; Day, 2003; Cooper *et al*., 2008; Trainor *et al*., 2014). By implementing CRM, firms can create and maintain long-standing relationships with customers based on complete customer satisfaction (Reinartz *et al.*, 2004). This customer-centric approach is likely to improve customer interaction, increase the value of products and services, and ultimately improve customer satisfaction and business performance (Ryals, 2005; Kasim and Minai, 2009). Kuei *et al*. (2001) pointed out that improvements in organizational performance are related to improved supplier relationships. Previous research has also shown that SRM can help companies stay competitive and improve profitability (Swink *et al.*, 2005; Kosgei and Gitau, 2016**;** Kiarie, 2017). Table 2 summarizes key research on the relationship between CRM, SRM, and organizational performance over the past decade. In summary, considering the previous studies, a firm's CRM and SRM practices are expected to be positively associated with customer satisfaction, ultimately contributing to market performance. Therefore, the following hypotheses are suggested:

H3. *CRM is positively related to customer satisfaction.*
H4. *SRM is positively related to customer satisfaction.*
H5. *Customer satisfaction is positively related to market performance.*

Table 2: Major studies on CRM, SRM and organizational performance over the past decade

Study	Objective	Data	Results
Park *et al.*, 2010	Develop an integrated framework for SRM system.	Case study based on semiconductor manufacturing in South Korea	With SRM, companies can lower their purchasing costs, develop products on time, and maintain high quality.
Coltman *et al*., 2011	Study the effect of CRM on firm performance.	50 firms from industry sectors that use CRM	CRM capability is positively associated with human analytics and business architecture.
Shang and Lu, 2012	Explore the impact of CRM on financial performance in the context of freight forwarder services.	144 freight forwarding companies	CRM has a significant positive impact on financial performance
Bhat and Darzi, 2016	Investigate the effect of CRM on customer loyalty.	278 customers of a private bank	CRM has a positive impact on competitive advantage by improving customer loyalty.
Bashir, 2017	Explore how CRM impact on customer retention.	316 private bank customers	CRM positively impacts customer retention. CRM also helps build long term relationship with prospective customers.
Kiarie, 2017	Investigate the effect of SRM practices on operational performance.	60 large manufacturing firms with over 500 employees	SRM is significantly related to operational performance and productivity of manufacturing firms

Pozza *et al.,* 2018	Investigate the impact of CRM relative implementation time on performance.	142 European and 208 US top managers	Implementing a comprehensive CRM strategy can increase growth and loyalty.
Soltani *et al.,* 2018	Investigate how CRM affect organizational performance.	150 tax employees	The success of CRM is heavily influenced by information technology use and customer knowledge management.

RESEARCH METHODS

Measures

All of the measuring scales used in this study were prepared based on previous studies; however, scale items have been slightly modified to suit our study objectives. The survey items estimate the extent to which the respondent has agreed or disagreed with the given statement based on the 7-point Likert scale. Specifically, social media-related items were developed based on the studies of Jayachandran et al. (2005) and Trainor et al. (2014). Respondents are asked if they are using social media for customer and vendor management. The study of Reinartz et al. (2004) has been used to evaluate the company's CRM practices; the questionnaire measures how companies use CRM to initiate, acquire and maintain customer information to better respond to customer needs and expectations. The work of Swink et al. (2005) was used to assess the company's SRM practices. These measures capture cost, technology, and long-term relationships with suppliers. Various prior studies such as Das et al. (2000), Cho et al. (2017), Kaynak (2003), and Cho and Linderman (2019) were adopted to develop items that measure customer satisfaction and market performance.

In addition, this study adopted three control variables; (i) the number of full-time employees was used to measure company's size (ii) the length of time in business was used to measure the age of the company, and (iii) two-digit SIC codes were applied to categorize the company's industry groups (Cho and Linderman, 2019). Further, to minimize the likelihood of common method variance (CMV), this study used a "marker variable" that measures the severity of the respondent's insomnia problem (Bastien et al., 2001). Details of all measure items used in this study are shown in the Appendix.

Sample

Based on the work of Dillman et al. (2014), we designed our survey to collect primary data. In particular, the following criteria were established to obtain the appropriate sample data from eligible survey respondents; (i) companies participating in the survey must run at least one Social Media platform for their customer and vendor management, (ii) survey participants must have full-time positions in their companies, and (iii) If there are two or more respondents in the same organization, the highest ranked respondents are selected (Miller and Roth, 1994; Cho and Linderman, 2019). Survey invitations were sent out to 241 operations managers in the United States. Through three rounds of the survey, a total of 117 responses were collected, showing a response rate of 48.52%. The demographic profile of the sample is summarized in Table 3.

Table 3: Sample demographics

Category	Count

Company's Age (Length of Time in Business)	< 5 years	3	2.6%
	5 ~ 10 years	15	12.8%
	11 ~ 20 years	15	12.8%
	21 ~ 30 years	7	6.0%
	31 ~ 50 years	18	15.4%
	51 ~ 100 years	34	29.1%
	> 100 years	18	15.4%
	Missing responses	7	6.0%
Company's Size (Number of Employees)	< 10 employees	13	11.1%
	11 ~ 50 employees	9	7.7%
	51 ~ 100 employees	9	7.7%
	101 ~ 500 employees	17	14.5%
	501 ~ 1,000 employees	14	12.0%
	1,001 ~ 10,000 employees	24	20.5%
	> 10,000 employees	24	20.5%
	Missing responses	7	6.0%
Industry [2-digit SIC codes]	Agriculture, Forestry & Fishing [01-09]	2	1.7%
	Construction [15-17]	5	4.3%
	Apparel & Fabricated Textile Products [23]	1	0.9%
	Pharmaceuticals [28]	1	0.9%
	Transportation Services [47]	2	1.7%
	Communications Services [48]	7	6.0%
	Retail Trade [52-59]	11	9.4%
	Financial Services [60-64]	7	6.0%
	Hotels & Other Lodging Places [70]	2	1.7%
	Prepackaged Software [73]	3	2.6%
	Healthcare [80]	26	22.2%
	Education [82]	16	13.7%
	Accounting & Business Consulting Services [87]	2	1.7%
	Others	18	15.4%
	Missing responses	14	12.0%

ANALYSIS AND RESULTS

Testing of Scale Reliability and Validity

First, the scale reliability was examined by Cronbach's alpha. The alpha coefficients for all variables were between .710 and .901, which met the suggested threshold of .70 or higher (Cronbach, 1951; Hair et al., 2010). Next, Confirmatory Factor Analysis (CFA) was conducted to test the construct validity (Lee Park and Paiva, 2018). The CFA results (Chi-square = 381.183; *d.f.* = 215; Normed Chi-square = 1.773; CFI = .895; PNFI = .673; RMSEA = .082; RMSEA 90% confidence interval: .068 ~ .095) indicated that all measure items satisfied the suggested threshold of .50 or higher (Hair et al., 2010). The results of these measure assessments are summarized in Table 4. Descriptive statistics and correlations between variables are also reported in Table 5.

Table 4: Results of CFA

Factor and Items	Loading[a]	S.E.	t	Sig.	α
Social Media Platform					
Conversation support (e.g., Blogger, Twitter)	.616			***	.710
Relationship support (e.g., Facebook, LinkedIn)	.556	.183	4.533	***	

Sharing support (e.g., Flickr, YouTube)	.752	.236	5.285	***	
Groups and community support (e.g., NING, IGLOO)	.572	.202	4.621	***	
Social Media Capability					
Our company uses Social Media to detect changes in customers' product or service preferences.	.835			***	.833
Our company uses Social Media to collect customer complaints or supplier feedback.	.855	.127	8.088	***	
Customer Relationship Management					
We invest in technology to acquire and manage "real time" customer information and feedback.	.700				.893
We have a formal system to assess the value of our customers.	.818	.152	8.242	***	
We have a formal system to identify potential customers.	.783	.148	7.911	***	
We have a formal system to differentiate our communication objectives based on the value of our prospects.	.775	.136	7.833	***	
We continually update customer information to assess the true value of our customers.	.888	.146	8.866	***	
Supplier Relationship Management					
We provide technical support to our suppliers.	.650				.817
We share our cost information with our suppliers.	.744	.153	6.589	***	
We require cost information sharing by our suppliers.	.786	.148	6.852	***	
We make long-term contracts with our suppliers.	.575	.135	5.347	***	
We seek joint investments with our suppliers.	.740	.151	6.566	***	
Customer Satisfaction					
Customer satisfaction with our products/services has increased over the past 3 years.	.783			***	.901
Customer retention in our products/services has increased over the past 3 years.	.910	.089	11.050	***	
The best description of your company's customer satisfaction level.	.775	.110	9.029	***	
The best description of your company's customer retention level.	.892	.094	10.793	***	
Market Performance					
Sales growth	.948			***	.895
Market share growth	.821	.068	12.229	***	
Net profit margin	.824	.061	12.306	***	

Note: n = 117; [a] Standardized regression weights; [b] Standard error (not estimated when loading set to fixed value: i.e., 1.0); AVE = Average variances extracted; *** $p < 0.001$.

Table 5: Correlations and descriptive statistics

	Variables	1	2	3	4	5	6	7	8	9	Mean	S.D.
1.	Social media platform										3.906	1.420
2.	Social media capability	.423**									3.594	1.873
3.	CRM	.283**	.556**								4.260	1.615
4.	SRM	.233*	.433**	.670**							3.826	1.305
5.	Customer satisfaction	.306**	.303**	.510**	.333**						4.915	1.217
6.	Market performance	.186*	.261**	.383**	.277**	.645**					4.593	1.230

7.	Insomnia[a]	.083	-.035	-.001	.063	-.016	-.151				2.432	1.053
8.	Firm age[b]	.094	-.022	-.136	-.073	-.233*	-.116	-.064			4.782	1.794
9.	Firm size[c]	.125	.193*	.179	.235*	-.017	.134	-.166	.541**		4.618	2.023
10.	Industry	-.007	.033	-.046	-.078	.077	.036	-.036	.192*	.128	1.909	0.289

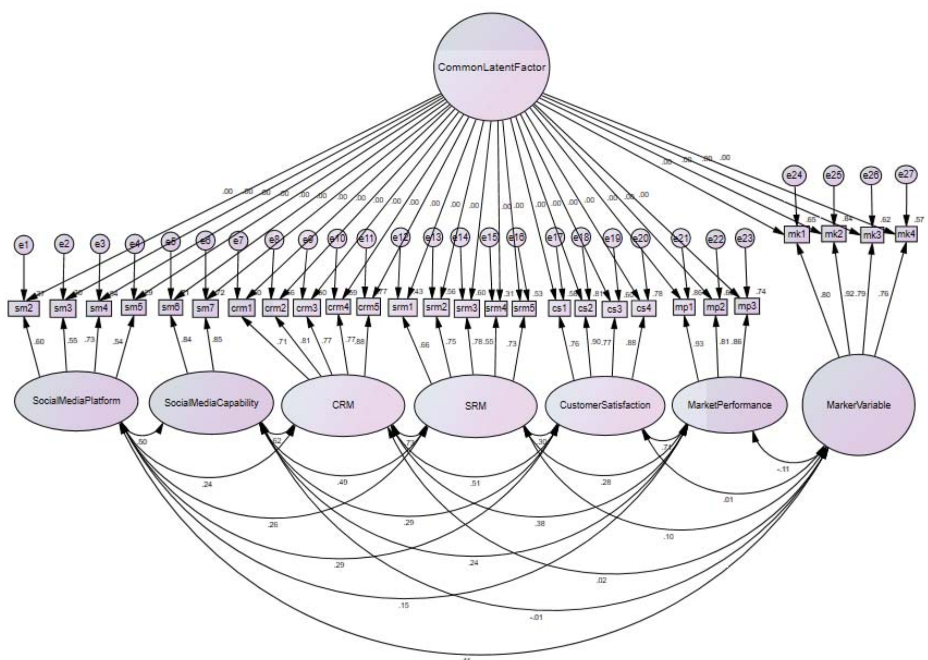
Note: *n* = 117; * significant at the .05 level (2-tailed). ** significant at the .01 level (2-tailed); [a] marker variable (to assess CMV); [b] length of time in business; [c] number of full-time employees.

Testing of Common Method Variance

Since the data of this study were obtained from a single source from each company, this study may not be free from common method variance (CMV) (Podsakoff et al., 2003). Hence, as an ex-ante remedy for minimizing the CMV threat, we included "marker variable" in our survey questionnaire (Podsakoff et al., 2003; Williams et al., 2010; Craighead et al., 2011). According to the guidelines of Simmering et al. (2015), "insomnia" was chosen as a marker variable for this study because the variable is not theoretically related to other variables in this study; the smallest correlation with other variables was also reported, as shown in Table 5. The insomnia measures used in this study are listed in the Appendix.

As an ex-post approach, we also investigated the potential threats of CMV in our data by performing a "common latent factor" analysis after completing data collection (Podsakoff et al., 2003; Richardson et al., 2009; Cho et al., 2017). Accordingly, we examined changes in the structural parameters when adding the common latent factor (CLF) to the measurement model. The test results demonstrated that the changes in parameters were minimal and CMV accounted for less than 1%, indicating that the CMV threat is not a pervasive problem in this study. Figure 4 shows the results of this CMV test performed by IBM AMOS 24.

Figure 4: Results of Common Latent Factor Analysis for CMV Assessment



Note: *n* = 108 (9 missing responses in the marker variable); standardized regression weights; Chi-square = 507.392; *d.f.* = 302; Normed Chi-square = 1.680; CFI = .879; PNFI = .647; RMSEA = .080 (90% confidence interval: .067 ~ .092).

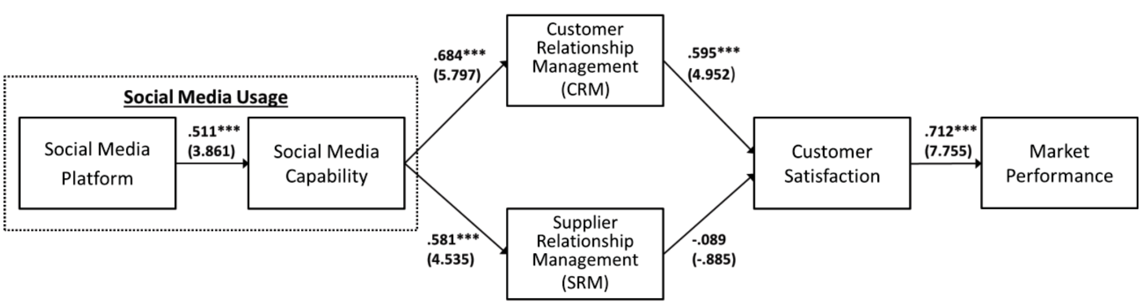
Results of SEM Analysis

To examine the hypotheses proposed in this study, we conducted a structural equaling modeling (SEM) analysis. The model fit was assessed by several different fit indices. The fit statistics reached or were very close to the desired threshold for each fit index (Chi-square = 415.688; *d.f.* = 224; Normed Chi-square = 1.856; CFI = .878; PNFI = .684; RMSEA = .086). Figure 5 illustrates the test results of the structural model hypothesized in this study. SEM test results demonstrate that social media platform is positively associated with social media capability (β = .511, t = 3.861, p < .001) and the usage of social media has a significant positive effect on both CRM (β = .684, t = 5.797, p < .001) and SRM (β = .581, t = 4.535, p < .001), supporting *H1* and *H2*. In addition, SEM test results indicate that CRM has a significant impact on customer satisfaction (β = .595, t = 4.952, p < .001) while there is no significant relationship between SRM and customer satisfaction (β = -.089, t = -.885, p = .376); therefore, *H3* is supported, but *H4* is rejected. Finally, the results also show that customer satisfaction significantly contributes to the company's market performance (β = .712, t = 7.755, p < .001), supporting *H5.*

Post Hoc Study

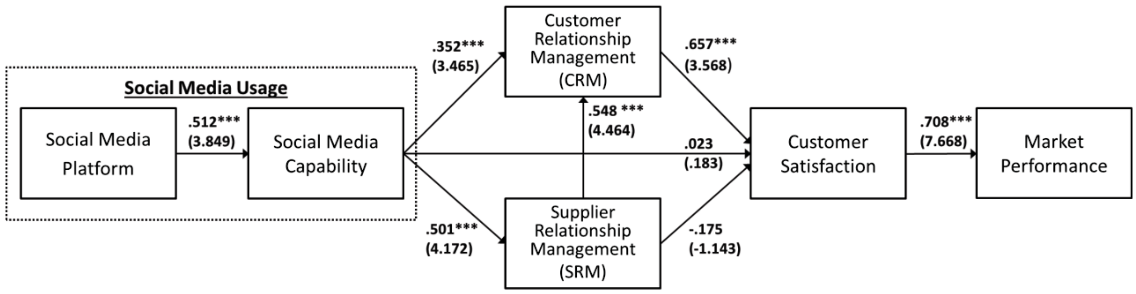
Although our sample did not support the direct effect of SRM on customer satisfaction (*H4*), SRM had a significant correlation with CRM and customer satisfaction as shown in Table V. Accordingly it is inferred that SRM indirectly affects customer satisfaction through CRM. We first tested this relationship with the modified SEM, as shown in Figure 6 (Chi-square = 387.551; *d.f.* = 222; Normed Chi-square = 1.746; CFI = .895; PNFI = .691; RMSEA = .080). Test results indicated that SRM had a significant positive impact on CRM (β = .548, t = 4.464, p < .001) while there was still no significant relationship between SRM and customer satisfaction (β = -.175, t = -1.143, p = .253). This implies that CRM is mediating the relationship between SRM and customer satisfaction. Therefore, we conducted an additional Sobel (1982) test to examine the significance of the mediation effect of CRM. Hence, z-value was computed by the following equation: $x*y/\mathrm{SQRT}(y^2*\mathrm{S.E.}_x^2 + x^2*\mathrm{S.E.}_y^2)$ where x is the unstandardized coefficient between SRM and customer satisfaction and y is the unstandardized coefficient between CRM and customer satisfaction; $\mathrm{S.E.}_x$ and $\mathrm{S.E.}_y$ are the standard errors. This result demonstrated that CRM had a significant mediation effect between SRM and customer satisfaction (z = 2.783, S.E. = 0.118, p = 0.005). Additionally, as shown in Figure 6, the SEM results of post hoc study also indicated that using social media did not directly affect customer satisfaction (β = .023, t = .183, p = .855).

Figure 5: Results of Hypothesized Structural Model



Note: n = 117; standardized path coefficients; numbers in parentheses indicate t-values; Chi-square = 415.688; *d.f.* = 224; Normed Chi-square = 1.856; CFI = .878; PNFI = .684; RMSEA = .086; *** p < .001.

Figure 6: Results of Post Hoc Analysis



Note: n = 117; standardized path coefficients; numbers in parentheses indicate t-values; Chi-square = 387.551; *d.f.* = 222; Normed Chi-square = 1.746; CFI = .895; PNFI = .691; RMSEA = .080; *** $p < .001$.

DISCUSSION

Implications for Practice and Research

The results of this study provide meaningful implications for practice and research in operational management. First, this study found that using social media has a significant impact on the company's CRM implementation, leading to customer satisfaction and competitive advantage in the marketplace. These results suggest that using social media can improve organizational performance by enhancing CRM capabilities. This finding is consistent with the earlier research by Trainor et al. (2014), which indicated that the use of social media could improve company's efficiency by facilitating internal and external collaboration that leads to internal solution improvements. The study findings also empirically support previous claims that effectively managing social media in conjunction with CRM would positively impact customer satisfaction and firm performances (Acker et al., 2011; Luo et al., 2013; Abzari et al., 2014). However, the post hoc study showed that using social media itself did not have a direct positive impact on organizational performance, indicating that technology alone is not sufficient to gain competitive advantage (Kotha, 1995). In short, the results of this study demonstrate that the combined operations of social media and CRM system enable companies to better respond to customer needs and gain feedback on products or services, thereby providing better products or services to customers. As such, it is advisable that operations managers consider increasing their investment in social media infrastructure and integrate their existing CRM systems with social media to improve operational efficiency and customer satisfaction.

Next, this study also showed that social media had a significant positive effect on firms' SRM implementation. This finding implies that the use of social media can improve SRM capability by allowing companies to quickly deliver their customers' real-time interests and changing needs to their suppliers, thereby reducing unnecessary inventories and improving lead times and operational efficiency. This study also revealed that the company's SRM practices did not directly affect customer satisfaction. However, the post hoc study has shown that the company's CRM system positively mediated the relationship between SRM and customer satisfaction. This finding implies that the impact of a firm's SRM practices on organizational performance can be achieved positively when combined with the efficient operation of the CRM system. Lam et al. (2016) argued that social media initiatives could improve knowledge-based benefits (e.g., information flow and knowledge sharing) within an organization and also enable organizations to access external knowledge and information outside their supply chain networks, ultimately driving operational efficiency and innovativeness. As such, instead of customer satisfaction and market performance used in this study, future research may adopt other criterion variables (e.g., new product development, product quality improvement, demand forecasting, and product and process innovation) in order to better understand how the combined operations of social media and the SRM system contributes to organizational performance.

Contributions to Theory

The present study also provides fundamental contributions to theory. First, by conceptualizing the critical role of social media in forming social networks among a focal firm, customers, and suppliers, this research contributes to the literature on social network perspective, integrating with the literature on operations management (Freeman, 2004; Borgatti et al., 2009; Kim et al., 2011). Second, this research also supports the knowledge-based view of the firm by showing that the combined framework of social media and supply chain information systems (i.e., CRM and SRM) can be a better strategic corporate resource to drive knowledge creation and superior organizational performance (Grant, 1996; Spender, 1996). Moreover, by adopting and validating the “insomnia” scale as a marker variable, this study also makes a methodological contribution to operations research in terms of how to defend against CMV threats arising from survey-based data collection.

Limitations and Future Research

Despite these meaningful contributions to practice and theory, there are some limitations to this study. All samples in this study came from U.S.-based companies operating under state-of-the-art social network infrastructure. However, given that many companies today are doing business globally in diverse social media infrastructure and regulatory environments, a fundamental question arises as to whether the results of this study will be useful for business operations in other countries. Further, previous studies have shown that cultural differences have a significant impact on information sharing practices (Li et al., 2014; Guo et al., 2017). Thus, future research should verify the social media effects on CRM and SRM implementations in various social media environments by using other country samples. Another limitation is that our study was conducted in the context of cross-sectional data. In the current study, we can not observe specific processes of how relationships with customers and suppliers are strengthened as social networks evolve through social media channels. Therefore, it seems particularly valuable to use longitudinal data to explore the evolution of social media effects on CRM and SRM implementations.

CONCLUSION

“To question the power of social media in society is to question the importance of sunlight on earth. It’s no longer a want, but a need," said Atanu Shaw, Vice President of Roosterly (2018 May 11, www.forbes.com). As discussed earlier, research on the role of social media has received a lot of attention in a variety of academic disciplines, but it is still in the early stages of operations and supply chain management. In particular, existing literature does not provide sufficient empirical evidence on how the use of social media affects the company's supplier and customer relationship management. Based on data from 117 US companies, we examined the impact of social media usage on the operational efficiency of CRM and SRM systems. The evidence in this study also supported that the social media technology embedded in the supply chain information systems could be a strategic resource for companies to create competitive advantage. We hope that this study will motivate future researchers to further study the role of social media use in operations and supplier chain management.

APPENDIX

Social Media Usage (Adapted from Jayachandran et al., 2005; Trainor et al., 2014)
Does your company use any forms of Social Media? YES □ NO □ If YES,
Social media platform
- Our company use Social Media for conversation support (e.g., Blogger, Twitter)
- Our company use Social Media for relationship support (e.g., Facebook, LinkedIn)
- Our company use Social Media for sharing support (e.g., Flickr, YouTube)
- Our company use Social Media for groups and community support (e.g., NING, IGLOO)

Social media capability
- Our company uses Social Media to detect changes in customers' product or service preferences.
- Our company uses Social Media to collect customer complaints / supplier feedback.

Customer Relationship Management (Adapted from Reinartz et al., 2004)
- We invest in technology to acquire and manage "real time" customer information and feedback.
- We have a formal system to assess the value of our customers.
- We have a formal system to identify potential customers.
- We have a formal system to differentiate our communication objectives based on the value of our prospects.
- We continually update customer information to assess the true value of our customers.

Supplier Relationship Management (Swink et al., 2005)
- We provide technical support to our suppliers.
- We share our cost information with our suppliers.
- We require cost information sharing by our suppliers.
- We make long-term contracts with our suppliers.
- We seek joint investments with our suppliers.

Customer Satisfaction (Adapted from Das et al., 2000; Cho et al., 2017)
- Customer satisfaction with our products/services has increased over the past 3 years.
- Customer retention in our products/services has increased over the past 3 years.
- The best description of your company’s customer satisfaction level on a scale of 1 to 7 (1 “low end of the industry,” 2 “much worse than average,” 3 “worse than average,” 4 “average,” 5 “better than average,” 6 “much better than average,” 7 “superior”)
- The best description of your company’s customer retention level on a scale of 1 to 7.

Market Performance (Kaynak, 2003; Cho and Linderman, 2019)
Evaluate your company’s market performance for competitive products or services in the industry on a scale of 1 to 7 (1 “low end of the industry,” 2 “much worse than average,” 3 “worse than average,” 4 “average,” 5 “better than average,” 6 “much better than average,” 7 “superior”).
- Sales growth
- Market share growth
- Net profit margin

Marker Variable (Bastien et al., 2001)
Please rate the current severity of your insomnia problem.
- Difficulty falling asleep: (1) None, (2) Mild, (3) Moderate, (4) Severe, (5) Very severe

- Difficulty staying asleep: (1) None, (2) Mild, (3) Moderate, (4) Severe, (5) Very severe
- How satisfied are you with your current sleep patterns? (1) Very satisfied, (2) Satisfied, (3) Moderately satisfied, (4) Dissatisfied, (5) Very dissatisfied
- To what extent do you think there is a sleep disorder that interferes with your daily life? (e.g. ability to function at work/daily chores, mood, concentration, memory)? (1) Not at all Interfering, (2) A little, (3) Somewhat, (4) Much, (5) Very much interfering.

REFERENCES

Abzari, M., Ghassemi, R.A. and Vosta, L.N. (2014), “Analysing the effect of social media on brand attitude and purchase intention: The case of Iran Khodro Company”, *Procedia-Social and Behavioral Sciences,* 143, pp.822-826.

Acker, O., Gröne, F., Akkad, F., Pötscher, F. and Yazbek, R. (2011), “Social CRM: How companies can link into the social web of consumers”, *Journal of Direct, Data and Digital Marketing Practice*, 13(1), pp.3-10.

Ainin, S., Parveen, F., Moghavvemi, S., Jaafar, N. I., and Mohd Shuib, N.L. (2015), “Factors influencing the use of social media by SMEs and its performance outcomes”, *Industrial Management & Data Systems*, 115(3), pp. 570-588.

Ali-Hassan, H., Nevo, D. and Wade, M. (2015), “Linking dimensions of social media use to job performance: The role of social capital”, *The Journal of Strategic Information Systems*, 24(2), pp.65-89.

Anton, J. (1996), *Customer relationship management: making hard decisions with soft numbers,* Upper Saddle River, NJ: Prentice-Hall.

Anton, J. and Hoeck, M. (2002), *E-Business Customer Service*, Santa Monica, CA: The Anton Press.

Aral, S., Dellarocas, C. and Godes, D. (2013). “Introduction to the special issue—social media and business transformation: a framework for research.” *Information Systems Research*, 24(1), pp.3-13.

Bashir, N. (2017), “Impact of Customer Relationship Management On Customer Retention (A Case Of Private Banks Of Sialkot, Punjab)”, *International Journal of Scientific & Technology Research*, 6(8), pp. 293-305.

Bastien, C.H., Vallières, A. and Morin, C.M. (2001), “Validation of the Insomnia Severity Index as an outcome measure for insomnia research”, Sleep Medicine, 2(4), pp.297-307.

Borgatti, S.P., Mehra, A., Brass, D.J. and Labianca, G. (2009), “Network analysis in the social sciences”, *Science,* 323, pp. 892-895.

Borgatti, S.P. and Li, X. (2009). “On social network analysis in a supply chain context”, *Journal of Supply Chain Management,* 45(2), pp. 5-22.

Bhat, S.A. and Darzi, M.A. (2016), “Customer relationship management: An approach to competitive advantage in the banking sector by exploring the mediational role of loyalty”, *International Journal of Bank Marketing*, 34(3), pp. 388-410.

Cao, X., Guo, X., Vogel, D. and Zhang, X. (2016), “Exploring the influence of social media on employee work performance”, *Internet Research*, 26(2), pp. 529-545.

Cetinkaya, A.S. and Rashid, M. (2018), “The Effect of Social Media on Employees’ Job Performance: The mediating Role of Organizational Structure” *Journal of Organizational Psychology* , 18(4), pp. 94-116.

Chen, I.J. and Popovich, K. (2003), “Understanding customer relationship management (CRM) People, process and technology”, *Business Process Management Journal*, 9(5), pp. 672-688.

Childerhouse, P. and Towill, D.R. (2003), “Simplified material flow holds the key to supply chain integration”, *Omega*, 31(1), pp.17-27.

Cho, Y.S., Jung, J.Y. and Linderman, K. (2017), "The QM evolution: Behavioral quality management as a firm's strategic resource", *International Journal of Production Economics*, 191, pp. 233-249.

Cho, Y.S. and Linderman, K. (2019), "Metacognition-based process improvement practices", *International Journal of Production Economics*, 211, pp. 132-144.

Choi, T.Y. and Krause, D.R. (2006), "The supply base and its complexity: Implications for transaction costs, risks, responsiveness, and innovation", *Journal of operations management*, 24(5), pp.637-652.

Choi, T.Y. and Kim, Y. (2008), "Structural embeddedness and supplier management: a network perspective", *Journal of Supply Chain Management*, 44(4), pp. 5-13.

Coltman, T., Devinney, T.M. and Midgley, D.F. (2011), "Customer relationship management and firm performance", *Journal of Information Technology*, 26(3), pp. 205-219.

Cooper, M.J., Gwin, C.F. and Wakefield, K.L. (2008), "Cross-functional interface and disruption in CRM projects: Is marketing from Venus and information systems from Mars?", *Journal of Business Research*, 61(4), pp. 292-299.

Corte, V., Iavazzi, A. and D'Andrea, C. (2015), "Customer involvement through social media: the cases of some telecommunication firms", *Journal of Open Innovation: Technology, Market, and Complexity*, 1(1), p. 10.

Cousins, P.D., Lawson, B. and Squire, B. (2006), "Supply chain management: theory and practice–the emergence of an academic discipline?", *International Journal of Operations & Production Management*, 26(7), pp. 697-702.

Craighead, C.W., Ketchen, D.J., Dunn, K.S. and Hult, G.T.M. (2011), "Addressing common method variance: guidelines for survey research on information technology, operations, and supply chain management", *IEEE Transactions on Engineering Management*, 58(3), pp. 578-588.

Croom, S., Romano, P. and Giannakis, M. (2000), "Supply chain management: an analytical framework for critical literature review", *European Journal of Purchasing & Supply Management*, 6(1), pp.67-83.

Cronbach, L.J. (1951), "Coefficient alpha and the internal structure of tests", *Psychometrika*, 16(3), pp. 297-334.

Das, A., Handfield, R.B., Calantone, R.J. and Ghosh, S. (2000), "A contingent view of quality management-the impact of international competition on quality", *Decision Sciences*, 31(3), pp. 649-690.

Day, G.S. (2003), "Creating a superior customer-relating capability", *MIT Sloan Management Review*, 44(3), p. 77.

Dillman, D.A., Smyth, J.D. and Christian, L.M. (2014), "*Internet, phone, mail, and mixed-mode surveys: the tailored design method",* John Wiley & Sons.

Dodokh, A. and Al-Maaitah, M.A. (2019), "Impact of Social Media Usage on Organizational Performance in the Jordanian Dead Sea Cosmetic Sector", *European Journal of Business and Management,* 11(2), pp. 75-91

Drury, G. (2008), "Opinion piece: Social media: Should marketers engage and how can it be done effectively?", *Journal of Direct, Data and Digital Marketing Practice*, 9(3), pp. 274-277.

Du, H. and Jiang, W. (2014), "Do social media matter? Initial empirical evidence.", *Journal of Information Systems,* 29(2), pp. 51-70.

Elena, C.A. (2016), "Social media–A strategy in developing customer relationship management", *Procedia Economics and Finance*, 39, pp.785-790.

Fader, P.S. and Winer, R.S. (2012), "Introduction to the special issue on the emergence and impact of user-generated content", *Marketing Science*, 31(3), pp. 369-371.

Feinberg, R. and Kadam, R. (2002), “E-CRM Web service attributes as determinants of customer satisfaction with retail Web sites”, *International Journal of Service Industry Management*, 13(5), pp. 432-451.

Feldmann, M. and Müller, S. (2003), “An incentive scheme for true information providing in supply chains”, *Omega*, 31(2), pp. 63-73.

Freeman, L. (2004), *The development of social network analysis. A Study in the Sociology of Science*, Empirical Press, Vancouver, BC, Canada.

Gallaugher, J. and Ransbotham, S. (2010), “Social media and customer dialog management at Starbucks”, *MIS Quarterly Executive*, 9(4), pp. 197-212.

Grant, R.M. (1996), “Toward a knowledge-based theory of the firm”, *Strategic Management Journal,* 17(S2), pp. 109-122.

Gunasekaran, A., Lai, K.H. and Cheng, T.E. (2008), “Responsive supply chain: a competitive strategy in a networked economy”, *Omega*, 36(4), pp. 549-564.

Guo, L., Tobias, J., Bendoly, E. and Hu, Y. (2017), “Different departments, different drivers: Asymmetry in antecedents and outcomes of voluntary knowledge exchange between sales and production functions”, *International Journal of Operations & Production Management,* 37(8), pp. 1031-1053.

Hagel, J. (1999), “Net gain: Expanding markets through virtual communities”, *Journal of Interactive Marketing*, 13(1), pp. 55-65.

Hair, J.F., Black, W.C., Babin, B.J., and Anderson, R.E. (2010,) *Multivariate Data Analysis* (7th ed.), Upper Saddle River, NJ: Pearson.

Handfield, R.B. and Bozarth, C.B. (2015), *Introduction to Operations and Supply Chain Management* (4th ed.), Pearson.

Hill, A.V. (2017), *The Encyclopedia of Operations Management: A field manual and encyclopedic glossary of supply chain and operations management terms*, Eden Prairie, MN: Clamshell Beach Press.

Hult, G.T.M., Ketchen, D.J. and Arrfelt, M. (2007), “Strategic supply chain management: Improving performance through a culture of competitiveness and knowledge development”, *Strategic Management Journal*, 28(10), pp. 1035-1052.

Jap, S.D. (1999), “Pie-expansion efforts: Collaboration processes in buyer–supplier relationships”, *Journal of Marketing Research*, 36(4), pp. 461-475.

Jayachandran, S., Sharma, S., Kaufman, P. and Raman, P. (2005), “The role of relational information processes and technology use in customer relationship management”, *Journal of marketing*, 69(4), pp.177-192.

Kaplan, A.M. and Haenlein, M. (2010), “Users of the world, unite! The challenges and opportunities of Social Media”, *Business Horizons*, 53(1), pp. 59-68.

Kasim, A. and Minai, B. (2009), “Linking CRM strategy, customer performance measures and performance in the hotel industry”, *International Journal of Economics and Management*, 3(2), pp. 297-316.

Kaynak, H. (2003), “The relationship between total quality management practices and their effects on firm performance”, *Journal of operations management*, 21(4), pp. 405-435.

Khan, D., Rahman, F., Al-balushi, H., Algaithi, A. and Al-Shihi, A. (2017), “Impact of Social Media on Customers Satisfaction: Bank Muscat–A Case Study”, *Ahead International Journal of Recent Research Review*, 1(11), pp. 154-163

Kiarie, J.W. (2017), *The Influence of supplier relationship management practices on operational performance of large manufacturing organizations in Kenya* (Doctoral dissertation, Strathmore University).

Kim, Y., Choi, T.Y., Yan, T. and Dooley, K. (2011), “Structural investigation of supply networks: A social network analysis approach”, *Journal of Operations Management*, 29(3), pp. 194-211.

Kim, Y., Chen, Y.S. and Linderman, K. (2015), "Supply network disruption and resilience: A network structural perspective", *Journal of Operations Management*, 33, pp. 43-59.

Kosgei, R.C. and Gitau, R. (2016), "Effect of supplier relationship management on organizational performance: A case study of Kenya Airways Limited*", International Academic Journal of Procurement and Supply Chain Management,* 2(2), pp. 134-148.

Kotha, S. (1995), "Mass customization: implementing the emerging paradigm for competitive advantage", *Strategic Management Journal*, 16(S1), pp.21-42.

Kozinets, R.V. (1999), "E-tribalized marketing?: The strategic implications of virtual communities of consumption", *European Management Journal*, 17(3), pp. 252-264.

Kuei, C.H., Madu, C.N. and Lin, C. (2001), The relationship between supply chain quality management practices and organizational performance. *International Journal of Quality & Reliability Management*, 18(8), pp. 864-872.

Lam, H.K., Yeung, A.C. and Cheng, T.E. (2016), "The impact of firms' social media initiatives on operational efficiency and innovativeness", *Journal of Operations Management*, 47, pp. 28-43.

Lee, L.F., Hutton, A.P. and Shu, S. (2015), "The role of social media in the capital market: Evidence from consumer product recalls", *Journal of Accounting Research*, 53(2), pp. 367-404.

Lee Park, C. and Paiva, E.L. (2018), "How do national cultures impact the operations strategy process?", *International Journal of Operations & Production Management*, 38(10), pp.1937-1963.

Li, S., Ragu-Nathan, B., Ragu-Nathan, T.S. and Rao, S.S. (2006), "The impact of supply chain management practices on competitive advantage and organizational performance", *Omega,* 34(2), pp. 107-124.

Li, Y., Ye, F. and Sheu, C. (2014), "Social capital, information sharing and performance: Evidence from China", *International Journal of Operations & Production Management*, 34(11), pp. 1440-1462.

Luo, X., Zhang, J. and Duan, W. (2013), "Social media and firm equity value", *Information Systems Research,* 24(1), pp.146-163.

Lyons, A.C., Everington, L., Hernandez, J., Li, D., Michaelides, R. and Um, J. (2013), "The application of a knowledge-based reference framework to support the provision of requisite variety and customisation across collaborative networks", *International Journal of Production Research*, 51(7), pp. 2019-2033.

Maecker, O., Barrot, C. and Becker, J.U. (2016), "The effect of social media interactions on customer relationship management", *Business Research,* 9(1), pp.133-155.

Miller, J.G. and Roth, A.V. (1994), A taxonomy of manufacturing strategies. *Management Science,* 40(3), pp. 285-304.

Moeller, S., Fassnacht, M. and Klose, S. (2006), "A framework for supplier relationship management (SRM)", *Journal of Business-to-Business Marketing*, 13(4), pp. 69-94.

Munar, A.M. and Jacobsen, J.K.S. (2013), "Trust and involvement in tourism social media and web-based travel information sources", *Scandinavian Journal of Hospitality and Tourism*, 13(1), pp. 1-19.

Nam, S.H., Vitton, J. and Kurata, H. (2011), "Robust supply base management: Determining the optimal number of suppliers utilized by contractors*", International Journal of Production Economics,* 134(2), pp. 333-343.

Nielsen Company (2010, August 2), "*What Americans do online social media and games dominate activity"* Retrieved from https://www.nielsen.com/us/en/insights/news/2010/what-americans-do-online-social-media-and-games-dominate-activity.html

Nyambu, E.M. (2013), "Influence of Social Media Marketing on Performance of Telecommunication Firms in Kenya", *International Journal of Innovative Research and Development*, 2(9), pp. 185-190.

Park, J., Shin, K., Chang, T.W. and Park, J. (2010), "An integrative framework for supplier relationship management", *Industrial Management & Data Systems*, 110(4), pp. 495-515.

Parveen, F., Jaafar, N.I. and Ainin, S. (2015), "Social media usage and organizational performance: Reflections of Malaysian social media managers", *Telematics and Informatics*, *32*(1), pp.67-78.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: A critical review of the literature and recommended remedies", *Journal of applied psychology*, 88(5), p. 879.

Pozza, I.D., Goetz, O. and Sahut, J.M. (2018), Implementation effects in the relationship between CRM and its performance. *Journal of Business Research*, 89, pp. 391-403.

Reinartz, W., Krafft, M. and Hoyer, W.D. (2004), "The customer relationship management process: Its measurement and impact on performance", *Journal of marketing research*, 41(3), pp. 293-305.

Richardson, H. A., Simmering, M. J., and Sturman, M. C. (2009), "A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance", *Organizational Research Methods*, 12(4), pp. 762-800.

Rodriguez, M., Peterson, R.M. and Krishnan, V. (2012), "Social media's influence on business-to-business sales performance", *Journal of Personal Selling & Sales Management*, 32(3), pp. 365-378.

Rodriguez, M., Peterson, R.M. and Ajjan, H. (2015), CRM/social media technology: impact on customer orientation process and organizational sales performance. In *Ideas in Marketing: Finding the New and Polishing the Old* (pp. 636-638). Springer, Cham.

Ryals, L. (2005), "Making customer relationship management work: the measurement and profitable management of customer relationships", *Journal of Marketing,* 69(4), pp. 252-261.

Sano, K. and Sano, H. (2016), The effect of social media on customer satisfaction and relationship strength in a service with high perceived risk in Japan. In *Celebrating America's Pastimes: Baseball, Hot Dogs, Apple Pie and Marketing?* (pp. 435-439). Springer, Cham.

Shang, K.C. and Lu, C.S. (2012), "Customer relationship management and firm performance: an empirical study of freight forwarder services", *Journal of Marine Science and Technology,* 20(1), pp. 64-72.

Shaw, A. (2018, May 11). "How Social Media Can Mover Your Business Forward" Retrieved from https://www.forbes.com/sites/forbescommunicationscouncil/2018/05/11/how-social-media-can-move-your-business-forward/#6ba48ac4cf2b

Shin, H., Collier, D.A. and Wilson, D.D. (2000), "Supply management orientation and supplier/buyer performance", *Journal of Operations Management*, 18(3), pp. 317-333.

Simmering, M.J., Fuller, C.M., Richardson, H.A., Ocal, Y. and Atinc, G.M. (2015), "Marker variable choice, reporting, and interpretation in the detection of common method variance: A review and demonstration", *Organizational Research Methods*, 18(3), pp. 473-511.

Soltani, Z., Zareie, B., Milani, F.S. and Navimipour, N.J. (2018), The impact of the customer relationship management on the organization performance. *The Journal of High Technology Management Research,* 29(2), pp. 237-246.

Song, D.P., Dong, J.X. and Xu, J. (2014), "Integrated inventory management and supplier base reduction in a supply chain with multiple uncertainties", *European Journal of Operational Research*, 232(3), pp. 522-536.

Sobel, M.E. (1982), Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, pp. 290-312.

Spender, J.C. (1996), "Making knowledge the basis of a dynamic theory of the firm", *Strategic Management Journal,* 17, pp. 45-62.

Swain, A. and Cao, Q. (2013) Exploring the impact of social media on supply chain performance: A sentiment analysis. In *the proceedings of 44th Annual Meeting, Decision Sciences Institute* (pp. 16-19).

Swink, M., Narasimhan, R. and Kim, S.W. (2005), "Manufacturing practices and strategy integration: effects on cost efficiency, flexibility, and market-based performance", *Decision Sciences*, 36(3), pp. 427-457.

Trainor, K.J., Andzulis, J.M., Rapp, A. and Agnihotri, R. (2014), "Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM", *Journal of Business Research*, 67(6), pp.1201-1208.

Trusov, M., Bucklin, R.E. and Pauwels, K. (2009), "Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site", *Journal of Marketing*, 73(5), pp. 90-102.

Vuori, V. and Okkonen, J. (2012), "Knowledge sharing motivational factors of using an intra-organizational social media platform", *Journal of Knowledge Management*, 16(4), pp. 592-603.

Wang, W.Y., Pauleen, D.J. and Zhang, T. (2016), How social media applications affect B2B communication and improve business performance in SMEs. *Industrial Marketing Management*, 54, pp. 4-14.

Whiting, A. and Williams, D. (2013), "Why people use social media: a uses and gratifications approach. *Qualitative Market Research", An International Journal*, 16(4), pp. 362-369.

Williams, L.J., Hartman, N. and Cavazotte, F. (2010), "Method variance and marker variables: A review and comprehensive CFA marker technique", *Organizational Research Methods*, 13(3), pp. 477-514.

Wu, Z. and Choi, T.Y. (2005), "Supplier–supplier relationships in the buyer–supplier triad: Building theories from eight case studies", *Journal of Operations management*, 24(1), pp. 27-52.

Wu, L. (2013), "Social network effects on productivity and job security: Evidence from the adoption of a social networking tool", *Information Systems Research*, 24(1), pp. 30-51.

Wu, C.W. (2016), "The performance impact of social media in the chain store industry", *Journal of Business Research*, 69(11), pp. 5310-5316.

Yeung, A.C. (2008), "Strategic supply management, quality initiatives, and organizational performance", *Journal of Operations Management*, 26(4), pp. 490-502.

Ylimaula, K.A. and Ulkuniemi, P. (2013), Exploring the potential of social media in supplier relationship management. In *IMP 2013 Conference*.

DECISION SCIENCES INSTITUTE

CEO Sentiment in Crisis and Subsequent Performance: A Two-Country Study

Supavich (Fone) Pengnate
North Dakota State University
Email: fone.pengnate@ndsu.edu

Derek G. Lehmberg
North Dakota State University
Email: derek.lehmberg@ndsu.edu

Chanchai Tangpong
North Dakota State University
Email: charnchai.tangpong@ndsu.edu

ABSTRACT

Evaluating the organization's performance and communicating to shareholders are central roles of top management. This study extends the literature by investigating sentiment in the letters of shareholders as an indicator of top managers' interpretation of the firms' situation, which consequently influences the firms' post crisis profitability performance. Sentiment analysis was conducted to extract sentiment from the letters of shareholders. Overall, the results suggest that sentiment is a viable indicator of the top managers' reflections on upcoming performance, contingent on the degree of management discretion.

KEYWORDS: Sentiment analysis, Letter to shareholders, organizational performance, economic crisis

INTRODUCTION

Top management plays an important role in interpreting the current situation of the company and communicating the future direction and expectations to organizational stakeholders (Hambrick & Mason, 1984). In economic crisis, where tensions create anxiety and test the emotions of the organization's stakeholders, positive communication from top management can mitigate anxiety and positively affect performance (Lehmberg & Tangpong, 2018). Managerial communications from top management can take several forms and reflect various substances, including managerial sentiment. Sentiment – positive or negative – may indicate what will be coming for the company in the future.

Recently, developments in sentiment analysis have facilitated new research examining sentiment in managerial communications such as letters to shareholders and organizational performance (Boudt & Thewissen, 2019). However, these studies have yet to examine sentiment in crisis situations, where they arguably are particularly important.

To help explore these new areas, we ask the following research questions: (1) In the midst of economic crisis, are top management sentiments identified in letters to shareholders an effective indicator of the top management' interpretation of the firm's situation? And (2), does the degree of managerial discretion moderate the role of sentiment addressed in the first research question?

THEORETICAL BACKGROUND AND HYPOTHESES

Upper echelons theory posits that top managers make decisions based upon their personal interpretation of the situation they face, and that these decisions affect how the company acts, and ultimately performs (Hambrick, 2007; Hambrick & Mason, 1984). These interpretations are colored by the manager's experience, skills, values, and personalities (Hambrick, 2007).

In the management literature, it has been reported that both cognitive and affective factors influence CEO's strategic choices in particular situations and consequently could be reflected in the firms' performance (Delgado-García & De La Fuente-Sabaté, 2010; Staw & Barsade, 1993).

In this study, we aim to examine how sentiment, which is a belief that has emerged from emotions, influences top management's expectations of the firm's subsequent performance. Positive sentiment in letters to shareholders suggests that top management interprets the company's situation as offering potential for improvement and better performance. In other words, positive sentiment can be considered as an expression of top management's positive expectations of future performance. On an individual level, this suggests that top managers may be more active in identifying and investing in growth opportunities. At the organizational level, the positive atmosphere which is established and maintained through positive sentiment of the top managers may lead to a healthy morale and organizational climate which supports effective implementation of existing strategies as well as development of new ones. These positive actions and healthy climate enable the firm to improve subsequent performance. Based upon this logic, we hypothesize:

H1: The greater the positive sentiment from the letter to shareholders, the better the firms' subsequent performance.

Previous studies have suggested that the relationship between sentiment and managers' strategic decision making can be moderated by certain managers' characteristics, such as managerial discretion (Espedal, 2007). Hambrick (2007) notes that the degree of managerial discretion or "latitude of action" (p. 335) present affects the relationship between the top management's interpretations and subsequent performance. Where managerial discretion is high, top managers have discretion to select and implement a large variety of strategic actions, whereas where discretion is low, top management has access to a much narrower variety of actions, and is restricted in how they may implement these actions (Crossland & Hambrick, 2011). In this study, we hypothesize that the relationship between sentiment and subsequent performance may be moderated by the degree of managerial discretion present.

H2: Management discretion moderates the relationship between sentiment and performance.

When research investigates the characteristics of managers across different countries, home country appears to be a reasonable proxy for managerial discretion at the national cluster level. This is because the degree of managerial discretion is systematically influenced by national institutions of the organization's home country.

For this research, we study firms from the U.S. and Japan. This allows us to examine the effect of managerial discretion systematically at the national cluster level as the two countries are arguably the extreme cases at the opposite ends of the spectrum of managerial discretion. Managerial discretion is very high in the U.S. and very low in Japan (Crossland & Hambrick, 2007, 2011). Accordingly, we expect to find a strong relationship between sentiment and subsequent performance in U.S. based firms, and a weak (or no) relationship in the Japan based firms.

METHODOLOGY

Sentiment Analysis

This research employs sentiment analysis to extract sentiments from the letters of shareholders. Sentiment analysis is considered as a text mining technique that classifies the polarity of the text into positive, negative, or neutral (Pang & Lee, 2008). Sentiment analysis can be conducted at various levels of text (e.g., word, sentence, or document) and involve other concepts. In this study, sentiment analysis was conducted at the document level to extract the overall sentiment from letters to shareholders.

A sentiment analysis tool developed by IBM, Watson Natural Language Understanding (NLU), was used to extract sentiments in the letters to shareholders. Watson NLU is a cloud-based application that provides a set of services to analyze text information. In general, the sentiment classification feature in Watson NLU enables users to identify a given piece of text expresses positive or negative opinion. The sentiment classification relies a supervised machine learning technique to create a sentiment classifier that uses a lexicon-based approach to process text. It is considered as a domain independent application, which generate sentiment scores from the correlations between the embedding pre-trained sentiment words and the documents (Li, Zhang, & Sindhwani, 2009). Specifically, Watson NLU relies on the Matrix Factorization with Lexical (MFLK) approach developed based on a non-negative matrix factorization method to derive the sentiment scores (Li et al., 2009). Studies examining managerial communications have content analyzed the letters to shareholders (e.g., Cho & Hambrick, 2006; McClelland, Liang, & Barker, 2010; Patelli & Pedrini, 2013; Staw, McKechnie, & Puffer, 1983), including sentiment analysis (Boudt & Thewissen, 2019). The letter to shareholder is highly suited to such research because it is a concise, high profile, publicly available document which can readily be compared with those of previous years or other companies. Additionally, because unlike other parts of the annual reports, the letters themselves are not audited or required by the SEC, giving the CEO more freedom to put these together as they wish (Abrahamson & Amir, 1996; Boudt & Thewissen, 2019).

Sample Selection and Data Gathering

We used a sample consisting of U.S. and Japan based companies. As stated above, this allows us to examine the effect of managerial discretion on the relationship between sentiment and subsequent performance because the U.S. is very high in managerial discretion while the Japanese context is very low (e.g., Crossland & Hambrick, 2011). Additionally, this also allows us to maintain comparability with past studies of letters to shareholders, which have typically relied on U.S. samples.

Since crisis is a context where the role of sentiment may be particularly important, our study used a time frame in the midst of economic crisis, following Lehman Brothers' bankruptcy announcement in September 2008. This crisis became worldwide and had strong effects on Japanese as well as U.S. based firms.

To facilitate this research and mitigate several potential biases, we used a sampling strategy that involved both judgmental and stratified aspects (Babbie, 1995). We first identified a number of publicly traded firms in Japan and the U.S. in major, representative industries. We then screened out Japanese firms that did not provide letters to shareholders in English, and finally, we adjusted the makeup of the sample so that both Japan and U.S. sub-samples had similar representations in each of the industries.

Our final sample consisted of 100 firms: 50 from the U.S. and 50 from Japan. The set of industries represented was broad, ranging from heavy machinery to food. For each of the sample firms, we obtained the letters to shareholders published immediately following the Lehman Brothers bankruptcy announcement in September 2008.

Dependent, Independent, and Control Variables

The dependent variable in this study is post-crisis profitability performance (PCPP), which is a measure derived from the firms' return on assets (ROA) and return on sales (ROS) reported a year after the crisis. Both ROA and ROS are common measures used in the literature to assess business firms' performance (e.g., Daniel, Lohrke, Fornaciari, & Turner, 2004). Since ROA and ROS are highly correlated, a principal component analysis (PCA) was used on ROA and ROS to create a composite measure, PCPP, as an indicator of the firms' post-crisis profitability performance and the dependent variable in this study.

Two independent variables are included in the model—sentiment and interaction between sentiment and the level of management discretion. The sentiment score is a continuous variable obtained from the MFLK algorithm. An interaction term is created from the sentiment score and the country dummy variable to observe the moderating effect of management discretion on the relationship between sentiment and PCPP.

In addition, a set of control variables are included in the model to control major confounding effects related to the firms' pre-crisis conditions. These variables are commonly used in the previous studies related to firm performance (e.g., Lahiri & Narayanan, 2013; Tangpong, Abebe, & Li, 2015). These control variables are (a) the firms' profitability performance derived from the PCA factor score composed of ROA and ROS, (b) level of leverage measured by debt-to-asset ratio, (c) the firms' liquidity measured by current ratio, (d) the firm size measured by total sale, and (e) management discretion, a categorical variable represented by country (Japan = 0 and US = 1). Note that leverage, liquidity, and firm size were log transformed to correct the distribution skewness.

DATA ANALYSIS AND RESULTS

The results form a correlation analysis suggested that there was an issue with multicollinearity when the interaction term between sentiment score and country variables entered the model. Therefore, sentiment score was mean centered to remove the multicollinearity produced by interaction.

We used three regression models to test our proposed hypotheses in examining the effects of sentiment and management discretion on PCPP, controlling for the firms' pre-crisis conditions. The first model was developed as a baseline, which contained only the control variables. The results revealed that, among the pre-crisis attributes, the firms' pre-crisis profitability performance and the country appear to be significant variables that influence PCPP.

Consequently, the second model was created to test Hypothesis 1, which proposed that sentiment generates positive influence on PCPP. The results suggested that, while the effects of sentiment were in expected direction, the significant level does not reach the 0.05 level; therefore, Hypothesis 1 is partially supported.

Finally, the third regression model was developed to test Hypothesis 2, which hypothesized a positive relationship between the interaction between sentiment and management discretion and PCPP. The findings indicated that the effect of the interaction term on PCPP is statistically significant at the 0.05 level. Therefore, these results suggested that management discretion plays a significant role in moderating the relationship between sentiment and PCPP and Hypothesis 2 is supported. Note that the firms' profitability performance is the only control variable that produced statistically significant effect on PCPP across the three regression models.

DISCUSSION AND CONCLUSION

Overall, the results of this study indicate that sentiment found in letters to shareholders can be an indicator of management's interpretation of the firm's subsequent performance (Hypothesis 1), especially those with high degrees of management discretion (Hypothesis 2). Therefore, this study adds to the current literature and provide insights regarding how top managements interpret the firms' situation and communicate with their shareholders. In addition to the consistent direction of the main effect of sentiment on the firms' performance with previous studies (e.g., Boudt & Thewissen, 2019), our results also indicate that organizational contexts such as management discretion could moderate such effect.

After incorporating the control variables into the model, while the direction was as expected, the effect of sentiment on PCPP was only significant at the 0.1 level (Hypothesis 1). One possible explanation is that sentiment, as an affective aspect of the strategic decision making communicated in the letters to shareholders, may alone be a less effective indicator of subsequent performance. It is also possible that during certain situations (e.g., economic crisis), the cognitive features in the letters to shareholders tend to make more of a contribution than do the affective features for strategic decision (Swann, Griffin, Predmore, & Gaines, 1987). Therefore, in addition to the writing tone and mood overall, future studies should also investigate features related to the cognitive aspect in the letters to shareholders, such as specific goals and outcomes, expected cash flow and net income, and other major financial measures.

In addition, our results indicate that management discretion appears to be a significant moderator between sentiment and PCPP. It is possible that, where managerial discretion is high, top managers are more willing to disclose their sentiment that better reflects the forward-looking into what the companies may be facing in the near future. As such, under the high management discretion condition, the sentiment in the letters to shareholders becomes a more accurate indicator of the subsequent performance. Alternatively, the finding can also be interpreted that when the top management shares the positive sentiment 'publicly' to shareholders and stakeholders, they may become more committed to making the positive future results happen. With managerial discretion, they are more likely to take action and own up to what they have communicated to shareholders and stakeholders publicly.

Contributions

This study makes two important contributions to the literature. First, our study suggests that whether sentiment extracted from the letters to shareholders is a valid indicator of the top managers' reflections on upcoming performance is contingent to the degree of management discretion. This thus extends beyond the previous research focusing primarily on the sentiment-performance link. Second, our study broadens the scope of top management communication study or the research on letters to shareholders by investigating them at the sentiment level, which is a key departure from the mainstream research focusing the investigating efforts at the strategic and cognitive levels (e.g., Lehmberg & Tangpong, 2018).

In addition to the contributions to the literature, this study provides two practical insights. First, based on our findings, when the managerial discretion is high, firms' investors, partners, and other stakeholders can use the sentiment in top management communications to public to gauge whether the firms' situation would be improving in the near future. This foresight would allow those stakeholders to determine whether or how much they should commit their investment of time, effort, and money to the firms. If the situation is assessed as negative, some active interventions from stakeholders would soon be put in place to mitigate the potential adverse consequences. Second, from the perspective of the firms' top managers, if they plan to use their public communication via letters to shareholders as a mechanism to distribute the positive sentiments to shareholders or to use it as a reinforcing mechanism to compel them to strive for performance improvement, they should ensure that they would have adequate managerial discretion to deliver what they have publicly communicated.

Limitations and Future Research

Despite its contributions, there are important limitations of the study that warrant discussion. First, we restricted our investigation on top managers. In fact, lower-level managers and employers may also influence the development of strategic decisions of the firms. In this study, we speculate that while influence may emerge from numerous parts of the organization, the top managers and the CEO dominate the judgement and interpretation of the firm's situation addressed in the letters to shareholders. Second, our sample was restricted to large firms during a major economic crisis. Therefore, our findings may not be generalized broad base. Third, we used the U.S. and Japan as a proxy for high and low managerial discretion, respectively. While this practice is supported by the literature (e.g., Crossland & Hambrick, 2011), it is a crude proxy, thus limiting us to examine the phenomena at a finer-grained level. Finally, we used a proprietary sentiment analysis tool provided by IBM Watson's Natural Language Understanding (NLU) to generate the sentiment scores. While the algorithm used by Watson NLU is considered as a domain-independent technique that can effectively classify sentiments in text data, further studies may examine additional sentiment analysis techniques which may result in different conclusions.

REFERENCES

Abrahamson, E., & Amir, E. (1996). The information content of the president's letter to shareholders. *Journal of Business Finance & Accounting, 23*(8), 1157-1182. doi:10.1111/j.1468-5957.1996.tb01163.x

Babbie, E. (1995). *The practice of social research* (7th ed.). Harrisonburg, VA: Wadsworth Publishing.

Boudt, K., & Thewissen, J. (2019). Jockeying for position in CEO letters: Impression management and sentiment analytics. *Financial Management, 48*(1), 77-115. doi:10.1111/fima.12219

Cho, T. S., & Hambrick, D. C. (2006). Attention as the mediator between top management team characteristics and strategic change: The case of airline deregulation. *Organization Science, 17*(4), 453-469. doi:10.1287/orsc.1060.0192

Crossland, C., & Hambrick, D. C. (2007). How national systems differ in their constraints on corporate executives: a study of CEO effects in three countries. *Strategic Management Journal, 28*(8), 767-789. doi:10.1002/smj.610

Crossland, C., & Hambrick, D. C. (2011). Differences in managerial discretion across countries: How nation-level institutions affect the degree to which ceos matter. *Strategic Management Journal, 32*(8), 797-819. doi:10.1002/smj.913

Daniel, F., Lohrke, F. T., Fornaciari, C. J., & Turner, R. A. (2004). Slack resources and firm performance: A meta-analysis. *Journal of Business Research, 57*(6), 565-574. doi:https://doi.org/10.1016/S0148-2963(02)00439-3

Delgado-García, J. B., & De La Fuente-Sabaté, J. M. (2010). How do CEO emotions matter? Impact of CEO affective traits on strategic and performance conformity in the spanish banking industry. *Strategic Management Journal, 31*(5), 562-574. doi:10.1002/smj.817

Espedal, B. (2007). Why rules rather than discretion: When the leadership intends to transform a desired policy into reality. *Journal of Organizational Change Management, 20*(1), 95-108. doi:doi:10.1108/09534810710715306

Hambrick, D. C. (2007). Upper echelons theory: An update. *The Academy of Management Review, 32*(2), 334-343. doi:10.2307/20159303

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *The Academy of Management Review, 9*(2), 193-206. doi:10.2307/258434

Lahiri, N., & Narayanan, S. (2013). Vertical integration, innovation, and alliance portfolio size: Implications for firm performance. *Strategic Management Journal, 34*(9), 1042-1064. doi:10.1002/smj.2045

Lehmberg, D., & Tangpong, C. (2018). Do top management performance attribution patterns matter to subsequent organizational outcomes? A two-country study of attribution in economic crisis. *Journal of Management & Organization*, 1-20. doi:10.1017/jmo.2018.32

Li, T., Zhang, Y., & Sindhwani, V. (2009). *A non-negative matrix tri-factorization approach to sentiment classification with lexical prior knowledge*. Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1, Suntec, Singapore.

McClelland, P. L., Liang, X., & Barker, V. L. (2010). CEO commitment to the status quo: Replication and extension using content analysis. *Journal of Management, 36*(5), 1251-1277. doi:10.1177/0149206309345019

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Informaiton Retrieval, 2*(1-2), 1-135.

Patelli, L., & Pedrini, M. (2013). *Is the optimism in CEO's letters to shareholders sincere? Impression management versus communicative action during the economic crisis* (Vol. 124).

Staw, B. M., & Barsade, S. G. (1993). Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses. *Administrative Science Quarterly, 38*(2), 304-331. doi:10.2307/2393415

Staw, B. M., McKechnie, P. I., & Puffer, S. M. (1983). The justification of organizational performance. *Administrative Science Quarterly, 28*(4), 582-600. doi:10.2307/2393010

Swann, W. B., Griffin, J. J., Predmore, S. C., & Gaines, B. (1987). The cognitive–affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology, 52*(5), 881-889. doi:10.1037/0022-3514.52.5.881

Tangpong, C., Abebe, M., & Li, Z. (2015). A temporal approach to retrenchment and successful turnaround in declining firms. *Journal of Management Studies, 52*(5), 647-677. doi:10.1111/joms.12131

DECISION SCIENCES INSTITUTE

A comparison of dispatching procedures to schedule jobs in a no-wait flow shop to minimize total earliness and tardiness with additional idle time allowed

Jeffrey Schaller
Eastern Connecticut State University
Email: schallerj@ecsu.ctstateu.edu

Jorge M. S. Valente
Universidade do Porto
Email: jvalente@fep.up.pt

ABSTRACT

This paper considers the problem of scheduling jobs in a no-wait flow shop with the objective of minimizing total earliness and tardiness. Idle time may be needed on the first machine due to the no-wait restriction. Additional idle time can be used to improve the objective. A model is developed that shows additional idle can be inserted on the first machine to help reduce earliness. Several dispatching heuristics that have been in other environments were modified and tested. The LIN1 and Modified Due Date (MDD) procedures performed best.

KEYWORDS: Scheduling; Heuristics; No-Wait Flow Shop; Earliness and Tardiness

INTRODUCTION

The widespread adoption of lean production methods has caused early delivery of products as well as tardy delivery to be viewed as undesirable. Early deliveries result in unnecessary inventory that ties up cash as well as space and resources needed to maintain and manage inventory. Products that are delivered late can cause penalties such as lost sales and loss of customer good will. This paper addresses this trend by considering an objective that sums the penalties for earliness and tardiness for a set of jobs to be processed in a no-wait flow shop.

A flow shop is a production shop that consists of two or more machines. In a flow shop each of the jobs to be processed uses the machines in the same order. In a no-wait flow shop once a job starts processing it must continue through the flow shop without any intermediate waiting. In some cases, the nature of the product is such that the no-wait restriction is a requirement. In many cases, organizations are adopting a lean production philosophy and want to minimize waste by not having any waiting between the machines. When jobs do not have to wait between the machines there is no in-process inventory and space is not wasted.

Formally, suppose there is a set of n jobs to be processed in a flow shop with M machines. Let d_j be the due date of job j $(j = 1, \ldots, n)$. Let p_{jm} and C_{jm} represent the processing time and completion time of job j $(j = 1, \ldots, n)$ on machine m $(m = 1, \ldots, M)$. The earliness of job j, E_j, is defined as:

$$E_j = \max\{d_j - C_{jM}, 0\}, \text{ for } j = 1, \ldots, n \tag{1}$$

and the tardiness of job j, T_j, is defined as:

$T_j = max\ \{C_{jM} - d_j, 0\}$, for $j = 1, \ldots, n$. (2)

The objective function, Z, can be expressed as:

$$Z = \sum_{j=1}^{n} E_j + T_j. \tag{3}$$

In a traditional flow shop, there could be idle time on any of the machines but in a no-wait flow shop idle time may be needed before a job begins processing on the first machine to ensure that the job does not have wait on machines 2 through M. This is due to the no-wait restriction in which storage of jobs is not allowed in between the intermediate stages. The objective in this problem is non-regular, therefore the insertion of additional idle time on the first machine (beyond what is needed to ensure a job's continuous processing once started) could help to reduce the earliness of some jobs and thus improve the objective. However, to our knowledge, previous research has not considered additional idle time for this problem. In this research schedules with additional inserted idle time in a no-wait flow shop are considered.

LITERATURE REVIEW

Earliness and Tardiness Costs

Many papers have been published on scheduling models with earliness and tardiness costs. (Baker and Scudder, 1990) provide an excellent survey of the initial work on early/tardy scheduling. A more recent survey of multicriteria scheduling which includes problems with earliness and tardiness penalties is given in (Hoogeveen, 2005). Most of the research with an objective based on early and tardy job completion costs deals with single machine environments. Recent research for single machine environments with an early/tardy objective and no idle time is summarized in (Valente, 2009). Scheduling models with inserted idle time, on the other hand, are reviewed in (Kanet and Sridharan, 2000).

Earliness and Tardiness in Flow Shops

Ten papers considered earliness and tardiness costs for flow shop environments without the no-wait restriction. Two early papers that considered the objective were (Zegordi et al.,1995) and (Rajendran, 1999). (Moslehi et al., 2009) considered the objective of minimizing the sum of maximum earliness and tardiness in a two-machine flow shop. (Chandra et al., 2009) consider the flow shop problem with earliness and tardiness penalties when there is a common due date. (Madhushini et al., 2009) developed branch-and-bound algorithms by for a variety of objectives including minimizing earliness and tardiness. (Schaller and Valente, 2013b) developed a genetic algorithm, (M'Hallah, 2014) developed a variable neighborhood search heuristic, and a constructive heuristic, as well as local searches were developed by (Fernandez-Viagas et al., 2016) for the problem. Family setups were incorporated into the problem by (Schaller and Valente, 2013a). (Schaller and Valente, 2019) consider the problem of scheduling a permutation flow shop to minimize total earliness and tardiness when unforced idle time is allowed and develop several heuristics for the problem.

No-Wait Flow Shops

To the best of our knowledge there have been no papers that have addressed the problem of minimizing total earliness and tardiness in a no-wait flow shop with additional idle time allowed. However, there has been a great deal of research on scheduling problems in the no-wait flow shop environment. A review of early work was provided by (Hall and Sriskandarajah, 1996). Much of the research dealt with objectives that were measures of efficiency such as minimizing makespan or flowtime. Recent research with an objective of minimizing makespan includes (Pan et al., 2008) and (Laha and Chakraborty, 2009). For the flowtime objective, (Framinan et al., 2010) and (Gao et al., 2011) are recent examples. There has been increased research in recent years on objectives that measure the shop's ability to meet due dates and include tardiness in the objective. Minimizing total tardiness is similar to our objective but does not penalize the early completion of jobs. (Aldowaisan and Allahverdi, 2012), (Liu et al., 2013) and (Ding et al., 2015) are recent examples of research for minimizing total tardiness in no-wait flow shops.

PROBLEM DEFINITION

In this paper we only consider no-wait flow shops. A permutation flow shop is a flow shop in which the order the jobs are processed remains the same on each machine. Therefore no-wait flow shops are also permutation flow shops so a solution is defined by determining the order in which the jobs will be processed. Also, for each pair of jobs it can be determined how much idle time is needed on the first machine between the finish time of the job sequenced first and the start time of the job sequenced second. This time is the same regardless of where the pair of jobs appears in the sequence so it can be calculated before sequencing jobs. We refer to this time as the delay or offset time and use OS_{jk} to denote the offset time needed on the first machine if job j is sequenced immediately after job k for job j not to wait at any of the machines. A dummy job 0 is used to take into consideration that a job may be first in a sequence and $OS_{j0} = 0$ for $j = 1, \ldots, n$. Once a start time for a job is determined, its completion time on the final machine is easy to determine as we only need to add the sum of the job's processing times to its start time. Let $[j]$ represent the job that is in the jth position of a given sequence. The completion time of the job in position j of the sequence if no additional idle were inserted is:

$$C_{[j]M} = C_{[j-1]1} + OS_{[j][j-1]} + \sum_{m=1}^{M} p_{[j]m} \text{, where } C_{[0]1} = 0. \tag{4}$$

If a job was to be completed early, we may consider delaying the start of the job on the first machine to be greater than what would be needed to ensure it did not have to wait at any machine. We refer to this delay as additional idle time. We use $AI_{[j]}$ to denote the additional idle time inserted before the job sequenced in position j. Therefore the completion time of the job in position j of the sequence is:

$$C_{[j]M} = C_{[j-1]1} + AI_{[j]} + OS_{[j][j-1]} + \sum_{m=1}^{M} p_{[j]m} \text{, where } C_{[0]1} = 0. \tag{5}$$

To calculate the offset time between a pair of jobs j and k when job j is to immediately follow job k (OS_{jk}) we can use the following.

$$OS_{jk} = \max_{m=1,\ldots,M-1} \{ \sum_{l=1}^{m} p_{kl+1} - \sum_{l=1}^{m} p_{jl} \}. \tag{6}$$

The offset times between pairs of jobs can be thought of as a sequence-dependent setup time on the first machine. Using this concept, the problem can be transformed into a single-machine early/tardy problem with sequence-dependent setups. Let d_j' be a modified due date for job j by setting

$$d_j' = d_j - \sum_{m=2}^{M} p_{jm} \ . \qquad (7)$$

This is the time job j must finish on the first machine in order to finish on time on the last machine since the no-wait constraint means that the job is processed on all machines without any idle time between machines. The earliness of job j, E_j, can be equivalently defined as:

$$E_j = \max\{d_j' - C_{j1}, 0\}, \text{ for } j = 1,\ldots, n \qquad (8)$$

and the tardiness of job j, T_j, is defined as:

$$T_j = \max\{C_{j1} - d_j', 0\}, \text{ for } j = 1,\ldots, n. \qquad (9)$$

The problem then becomes:

$$\text{Min } Z = \sum_{j=1}^{n} E_{[j]} + T_{[j]} \qquad (10)$$

Subject to:

$$C_{[j]1} = C_{[j-1]1} + AI_{[j]} + OS_{[j][j-1]} + p_{[j]1}. \qquad \text{For } j = 1, \ldots, n \qquad (11)$$

If we have a specific sequence, then the insertion of additional idle on the first machine is an important element in optimizing the total earliness and tardiness for that sequence. A single-machine timetabling procedure for inserting idle time into a given sequence for minimizing total earliness and tardiness (Fry et al., 1987; Davis and Kanet, 1993; Kim and Yano, 1994) can be used with the modified due dates (dj's) to accomplish this task. In the next section we describe dispatching procedures that are used to develop sequences for the problem and once a sequence is generated a timetabling procedure is used to insert additional idle time on the first machine to obtain an objective value.

DISPATCHING HEURISTICS TESTED

Seven dispatching heuristics that have been used for minimizing total earliness and tardiness in permutation flow shops but without the no-wait restriction and considering additional idle time are tested in this research. These procedures were modified for the no-wait flow shop environment and to consider adding additional idle time on the first machine as described in the previous section.

We use the following notation in the heuristics. S is a current partial schedule. If job j ($j \notin S$) is the next job scheduled after S, $C_{jM}(S)$ is job j's the completion time. We use $s_j(S)$ to represent the slack of job j if job j is the next job scheduled after S, where $s_j(S) = d_j - C_{jM}(S)$. Additionally, let t be the current availability time of machine 1 under schedule S. This is the time the last job in schedule S completes its processing on the first machine. $P_j(S)$ ($P_j(S) = C_{jM}(S) - t$) is used to represent the time the shop is allocated to job j if job j is the next job scheduled after S, which includes the total processing time of job *j* plus the initial idle time on the first machine. Therefore, since the problem deals with a no-wait flow shop environment, *Pj(S)* includes both the job's cumulative processing time and its offset time. Suppose *k* jobs have been selected and we are now selecting the $k+1^{st}$ job. Let *[k]* represent the job in position *k*.

$$P_j\,(S) = OS_{j[k]} + \sum_{m=1}^{M} p_{jm} \quad . \tag{12}$$

Earliest Due Date (EDD)

The earliest due date (EDD) was first proposed by (Jackson, 1955). This rule schedules the jobs in non-decreasing order of their due dates d_j.

Modified Due Date (MDD)

In the modified due date (MDD) heuristic (Baker and Bertrand, 1982; Vepsalainen and Morton, 1987), at each iteration we select the job with the minimum value of the modified due date.

$$MDD_j\,(S) = \max\,\{d_j, C_j\,(S)\} = \max\,\{d_j, t + P_j\,(S)\} = \max\,\{d_j - t, P_j\,(S)\}. \tag{13}$$

Minimum Slack Rule (SLK) and Minimum Slack per Work Rule (SLK/P)

The minimum slack (SLK) rule (Panwalkar & Iskander, 1977; Vepsalainen & Morton, 1987) chooses, at each iteration, the job with the minimum slack

$$s_j\,(S) = d_j - C_j\,(S). \tag{14}$$

The minimum slack per required time (SLK/P) (Panwalkar & Iskander, 1977; Vepsalainen & Morton, 1987) selects, at each iteration, the job with the minimum value of the ratio

$$SLK/P_j\,(S) = s_j\,(S)/\,P_j\,(S). \tag{15}$$

LIN-ET Rules (LIN1 and LIN2)

The following two rules are based on the LIN–ET procedure proposed in (Ow & Morton, 1989) for the weighted single machine problem. These rules, which will be denoted by LIN1 and LIN2, choose, at each iteration, the job with the largest value of the following priority indexes:

$$LIN1_j(S) = \begin{cases} \frac{1}{P_j(S)} & if \quad s_j(S) \le 0 \\ \frac{1}{P_j(S)} - \frac{s_j(S)}{slk\_thr} \times \frac{2}{P_j(S)} & if \quad 0 < s_j(S) < slk\_thr \\ -\frac{1}{P_j(S)} & if \quad s_j(S) \ge slk\_thr \end{cases} \tag{16}$$

and

$$LIN2_j(S) = \begin{cases} \frac{1}{P_j(S)} & if \quad s_j(S) \le 0 \\ \frac{1}{P_j(S)} - s_j(S) \times \left(\frac{1}{slk\_thr \times P_j(S)} + \frac{1}{P_j(S)}\right) & if \quad 0 < s_j(S) < slk\_thr, \\ -\frac{s_j(S)}{P_j(S)} & if \quad s_j(S) \ge slk\_thr \end{cases} \tag{17}$$

where slk_thr, which stands for "slack threshold", is a parameter meant to represent a value such that slacks which are greater or equal to that value are considered large.

In the first branch of the priority index, identical in both heuristics, a job is late or on time if scheduled next. When one or more such jobs exist, LIN1 and LIN2 select the job using a shortest time rule, in line with various heuristics for (earliness and) tardiness problems (Baker, 1974; Ow & Morton, 1989; Smith, 1956).

In the third branch, the job has a large slack, and is quite early. If all jobs are quite early, LIN1 chooses the job using a longest time rule, again in line with several heuristics for early/tardy problems (Ow & Morton, 1989; Valente, 2007; Valente & Alves, 2005). LIN2, on the other hand, selects the job with the minimum slack per required time. Finally, the middle branch performs a linear interpolation between the priority values corresponding to $s_j(S) = 0$ and $s_j(S) = slk\_thr$. Such an interpolation was first performed in the LIN–ET procedure (Ow & Morton, 1989).

The *slk_thr* parameter is calculated as follows. At each iteration, representing a partial schedule *S* with *k* jobs, the slack threshold is set equal to

$$slk\_thr = w * (C_{\max}^{\mathrm{LB}}(S) - t), \tag{18}$$

where $C_{\max}^{\mathrm{LB}}(S)$ is a lower bound on the completion time of the last job on the final machine (makespan), given the current schedule *S*, and $0 \leq w \leq 1$ is a user-defined parameter. To calculate this lower bound we first obtain a lower bound on the completion of processing on the first machine of the unscheduled jobs and then add to this time the minimum cumulative processing time on machines two through *M*. To find the lower bound for completion on the first machine we sum the processing times on the first machine of the unscheduled jobs and add this to t ($t + \sum_{j \notin S} p_{[j]1}$). This lower bound is weak because the offset times of the unscheduled jobs is not included. To strengthen this lower bound we find the lowest offset time for each unscheduled job if it is not scheduled next (min{$\mathrm{OS}_{j \notin \mathrm{S}, \mathrm{l} \notin S}$} and sum the $n - k - 1$ of these times and add to this the minimum offset time among the unscheduled jobs if scheduled next. The lower bound on the offset times is then added to the lower bound above to obtain a lower bound on the completion time on the first machine. We then add the minimum among the unscheduled jobs of the cumulative processing time for machines two through *M* to obtain a lower bound on the makespan.

A Heuristic Based on (Fernadez-Viagas et al., 2016)'s Constructive Heuristic

A constructive heuristic for permutation flow shops when waiting is allowed between machines was developed by (Fernandez-Viagas et al., 2016). The heuristic adds one job at a time to a partial sequence based on an index. During each iteration of the procedure the problem is classified based on the due dates of the unselected jobs in order to determine the index to use to select the next job in the sequence. The index used is based on whether the due dates are relative tight, relatively loose or are neither relatively tight or loose. For details about the procedure see (Fernandez-Viagas et al., 2016).

In the original (Fernandez-Viagas et al., 2016) heuristic a variable is used to calculate the weighted idle time of the candidate jobs. The weighted idle variable in the (Fernandez-Viagas et al., 2016) heuristic is based on the idle time that occurs on all the machines if a job is to be scheduled next. Since in our problem the shop is a no-wait flow shop and idle time only occurs on the first machine, we modify this heuristic by using the idle time on the first machine for each unscheduled job if scheduled next. Let IT_{jk} be the idle time on the first machine of the candidate jobs if scheduled next (in the $k+1^{st}$ position): $IT_{jk} = OS_{j[k]}$. This redefined variable is then used in the (Fernandez-Viagas et al., 2016) indexes to select a job during each iteration. This procedure is referred to as FV in this paper.

COMPUTATIONAL TEST

The proposed algorithms are tested on randomly generated problems of various sizes in terms of the number of jobs and number of machines and under various conditions of due date range and tightness.

Data and Performance Measures

The dispatching heuristic procedures described in section three were tested on problems of various sizes in terms of the number of jobs and number of machines for nine sets of distributions of due date range and tightness. Each problem set consists of 10 problems. The problems within a problem set have the same number of jobs and machines, and the due dates for the jobs are generated using the same distribution. Eight levels of number of jobs (n) to be scheduled were tested: $n = 15, 20, 25, 30, 40, 50, 75$ and 100. Three levels of number of machines (M) were tested: $M = 5, 10$ and 20. The processing times of the jobs for each machine were generated using a uniform distribution over the integers 1 and 100. The due dates for the jobs were also randomly generated using a uniform distribution over the integers $MS\ (1 – r – R/2)$ and $MS\ (1 – r + R/2)$, where MS is an estimated makespan found for the problem using the makespan lower bound proposed in (Taillard, 1993), and R and r are two parameters called due date range and tardiness factors. Three levels of due date range (R) were tested: $R = 0.2, 0.6$ and 1.0 and three levels of due date tightness (r) were tested: $r = 0.0, 0.2$ and 0.4. These levels of R and r result in nine sets of due date parameters for each n and M combination.

The procedures were coded in Turbo Pascal and were tested on a Dell Inspiron 1525 GHz Lap Top computer. The measure of performance used to evaluate the dispatching procedures for the problems is percentage deviation ($\%\ Dev$) of the total earliness and tardiness of the solution generated by each procedure from the total earliness and tardiness generated by the EDD procedure.

$$\%\ Dev = [(Z_h - Z_{EDD})/ Z_h] * 100, \qquad (19)$$

where Z_{EDD} = the total earliness and tardiness of the solution generated by the EDD procedure, and Z_h = the total earliness and tardiness of the solutions generated by the dispatching heuristic procedures (MDD, LIN1, LIN2, SLK, SLKP). Since the EDD procedure would always have a $\%\ Dev$ = 0 it is omitted from the tables.

Results

Tables 1, 2 and 3 show the $\%\ Dev$ for each procedure for each level of number of jobs to be sequenced (n) as well as the averages across all the levels of jobs. Table 1 shows the results for $M = 5$, table 2, for $M = 10$ and table 3, for $M = 20$.

Table 1. % deviation from EDD solution among dispatching heuristics for $M = 5$.

	Procedure					
n	MDD	SLK	SLK/P	LIN1	LIN2	FV

15	-14.56	15.71	8.14	-12.74	-6.83	0.00
20	-16.71	15.47	2.19	-16.29	-12.55	-3.76
25	-18.73	16.67	1.91	-18.54	-14.74	-3.96
30	-18.77	16.39	-0.99	-18.70	-15.25	-3.35
40	-20.15	14.59	-4.87	-21.87	-18.58	-7.94
50	-22.34	14.67	-6.32	-23.86	-19.50	-8.34
75	-23.21	14.06	-9.96	-26.80	-21.45	-10.21
100	-23.43	14.13	-11.17	-27.61	-21.80	-10.90
Ave.	-19.74	15.21	-2.63	-20.80	-16.34	-6.06

Table 2. % deviation from EDD solution among dispatching heuristics for $M = 10$.

	Procedure					
n	MDD	SLK	SLK/P	LIN1	LIN2	FV
15	-21.83	18.30	8.40	-20.08	-15.78	-7.31
20	-21.02	18.23	5.26	-19.01	-16.34	-10.33
25	-24.41	18.56	2.52	-22.62	-19.83	-12.22
30	-25.66	19.65	0.01	-22.89	-19.92	-14.01
40	-26.83	17.91	-3.09	-25.17	-21.93	-17.07
50	-27.67	19.53	-4.96	-26.66	-22.76	-17.53
75	-29.77	18.39	-10.96	-30.05	-26.36	-20.15
100	-19.39	17.93	-13.36	-31.79	-27.33	-22.24
Ave.	-24.57	18.56	-2.02	-24.78	-21.28	-15.11

Table 3. % deviation from EDD solution among dispatching heuristics for $M = 20$.

	Procedure					
n	MDD	SLK	SLK/P	LIN1	LIN2	FV
15	-26.17	17.20	7.96	-23.46	-20.29	-15.86
20	-26.18	15.88	4.55	-22.83	-20.82	-16.29
25	-28.44	16.05	2.49	-26.47	-23.12	-21.29
30	-28.79	19.31	2.22	-26.22	-23.81	-20.77
40	-29.20	18.70	-2.38	-27.37	-25.65	-23.02
50	-30.81	19.48	-5.59	-28.80	-27.26	-24.91
75	-26.92	19.33	-10.31	-31.19	-28.78	-26.53
100	-18.40	19.79	-12.83	-32.69	-30.28	-28.57
Ave.	-26.86	18.22	-1.74	-27.38	-25.00	-22.16

The results show that the LIN1 procedure had the lowest average *% Dev* for each level of number of machines followed by the MDD procedure. The MDD procedure had *% Dev*s that were better than those of LIN1 when the number of jobs was lower. For $M = 5$ this was true for $n < 40$, for $M = 10$, $n < 75$, and for $M = 20$, $n < 75$. The LIN2 procedure was generally the third best performing procedure. The SLK and SLK/P performed the worst for all problem sizes. The SLK procedure generated solutions that were worse than the EDD procedure for all problem sizes and the SLK/P procedure had *% Dev*s that were greater than 0 for most of the combinations of *n* and *M*.

In order to show the effect of the due date range (*R*) and tardiness factor (*r*) on the results tables 4 and 5 are presented. Table 4 shows the *% Dev* by due date tardiness factor (*r*) for *n = 50* and *M = 10*.

Table 4. % deviation from EDD solution by *r* for *n = 50* and *M = 10*.

	Procedure					
r	MDD	SLK	SLK/P	LIN1	LIN2	FV
0.0	-22.32	20.25	-0.22	-21.11	-17.02	-13.84
0.2	-28.85	21.11	-4.82	-27.10	-23.42	-16.86
0.4	-31.86	17.24	-9.85	-31.76	-27.85	-21.89

The results by due date tardiness factor (r) show that the MDD procedure was best for for all three levels of *r* for this combination of *n* and *M*, followed by LIN1 and LIN2. The results also show that the performance of these procedures (MDD, LIN1 and LIN2), as well as the FV and SLK/P procedures, improves relative to the EDD procedure as the due dates become tighter.

Table 5 shows the *% Dev* by due date range factor (*R*) for *n = 50* and *M = 10*.

Table 5. % deviation from EDD solution by *R* for *n = 50* and *M = 10*.

	Procedure					
R	MDD	SLK	SLK/P	LIN1	LIN2	FV
0.2	-23.72	23.83	-9.01	-21.70	-19.65	-14.01
0.6	-28.21	15.45	-4.14	-27.21	-22.69	-19.06
1.0	-31.10	19.32	-1.74	-31.07	-25.95	-19.52

The results that the MDD procedure was the best for all three levels of this parameter for this combination of *n* and *M*, followed by the LIN1, LIN2, FV and SLK/P procedures. The SLK procedure was the worst performer for all three values of the parameter. The MDD, LIN1, LIN2 and FV procedures showed improved performance relative to the EDD procedure as due dates became more spread out. The performance of the SLK/P became worse relative to the EDD procedure as due dates became more spread out.

All of the procedures are very efficient and generated solutions quickly for all of the problem sizes tested.

CONCLUSION

In this paper seven dispatching heuristics were tested for minimizing total earliness and tardiness in no-wait flow shops when using additional inserted idle time on the first machine to reduce job earliness is considered. The procedures were tested on problems of various sizes in terms of the number of jobs and machines, and nine sets of distributions that determine the tightness and range of due dates. The procedures are very efficient and were able to generate solutions quickly for all the problem sizes tested.

The results showed that the Modified Due Date (MDD) and LIN1 consistently generated solutions with a lower total earliness and tardiness than the other procedures tested.

Areas of additional research would be the development of a branch-and-bound procedure or a lower bounding technique in order to gauge the accuracy of the tested procedures. Developing additional dispatching procedure that generate better solutions and finding improvement approaches that can be applied to the solutions generated by the dispatching procedures.

REFERENCES

Aldowaisan, T. and A. Allahverdi, (2012). Minimizing total tardiness in no-wait flowshops. *Foundations of Computing and Decision Sciences*, 37, pp. 149 – 162.

Baker, K.R. (1974). Introduction to sequencing and scheduling. New York: Wiley.

Baker, K.R. & Bertrand, J.W.M. (1982). A dynamic priority rule for scheduling against due-dates. Journal of Operations Management, 3(1), 37-42.

Baker, K.R. and Scudder, G.D. (1990) 'Sequencing with earliness and tardiness penalties: a review', *Operations Research*, Vol. 38, pp.22–36.

Chandra, C. Mehta, P. and D. Tirupati, Permutation flow shop scheduling with earliness and tardiness penalties, *International Journal of Production Research*, 2009; 47: 5591 – 5610.

Davis, J. S. and J. J. Kanet, "Single-machine scheduling with Early and Tardy Completion Costs," *Naval Research Logistics,* 40, 85 – 101, 1993.

Ding, J., Song, S., Zhang, R. Gupta, J. N. D., and C. Wu, (2015). Accelerated methods for total tardiness minimization in no-wait flowshops, *International Journal of Production Research*, 53 (4), pp. 1002 – 1018.

Fernandez-Viagas, V., Dios, M. and J. M. Framinan, "Efficient constructive and composit heuristics for the permutation flowshop to minimize total earliness and tardiness," *Computers and Operations Research*, 70, 38 – 68, 2016.

Framinan, J. M., Nagano, M. S., and J. V. Moccellin, (2010). An efficient heuristic for total flowtime minimization in no-wait flowshops, *The International Journal of Advanced Manufacturing Technology*, 46 (9-12), pp. 1049 – 1057.

Fry, T. D., Armstrong, R. D., and J. H. Blackstone, "Minimizing weighted absolute deviation in single machine scheduling," *IIE Transactions,* 9, 445 – 450, 1987.

Gao, K., Pan, Q., and J. Li, (2011). Discrete harmony search algorithm for the no-wait flow shop scheduling problem with total flow time criterion, *The International Journal of Advanced Manufacturing Technology*, 56 (5-8), pp. 683 – 692.

Hall, N. G. and C. Sriskandarajah. (1996). A survey of machine scheduling problems with blocking and no-wait in process. Operations Research, 44 (3): 510 – 525.

Hoogeveen, H. (2005) Multicriteria scheduling, European Journal of Operational Research, Vol. 167, pp.592–623.

Jackson, J.R. (1955). Scheduling a production line to minimize maximum tardiness. In: Management Science Research Project. Los Angeles: University of California.

Kanet, J.J. and Sridharan, V. (2000) 'Scheduling with inserted idle time: problem taxonomy and literature review', Operations Research, Vol. 48, pp.99–110.

Kim, Y. D., and C. A. Yano, "Minimizing mean tardiness and earliness in single-machine scheduling problems with unequal due dates," *Naval Research Logistics*, 41, 913 – 933, 1994.

Laha, D., and U. K. Chakraborty. (2009). A constructive heuristic for minimizing makespan in no-wait flow shop scheduling. *The International Journal of Advanced Manufacturing Technology*, 41 (1-2), pp. 97 – 109.

Liu, G., Song, S., and W. Cheng, (2013). Some heuristics for no-wait flowshops with total tardiness criterion, *Computers and Operations Research*, 40 (2), pp. 521 – 525.

Madhushini, N., Rajendran, C. and Y. Deepa, Branch-and-bound algorithms for scheduling in permutation flowshops to minimize the sum of weighted flowtime/sum of weighted tardiness/sum of weighted flowtime and weighted tardiness/sum of weighted flowtime, weighted tardiness and weighted earliness of jobs, *Journal of the Operational Research Society*, 2009; 40: 991 – 1004.

Moslehi, G., Mirzaee, M., Vasei, M., and A. Azaron, Two-machine flow shop scheduling to minimize the sum of maximum earliness and tardiness, *International Journal of Production Economics*, 2009; 122: 763 – 773.

M'Hallah, R. An iterated local search variable neighborhood descent hybrid heuristic for the total earliness tardiness permutation flow shop. *International Journal of Production Research*, 52 (13); 3802 – 19.

Ow, P.S. & Morton, T.E. (1989). The Single-Machine Early Tardy Problem. Management Science, 35(2), 177-191.

Pan, Q., Wang, L., and B. Zhao. (2008). An improved iterated greedy algorithm for the no-wait flow shop scheduling problem with makespan criterion. *The International Journal of Advanced Manufacturing Technology*, 38 (7-8), pp. 778 – 786.

Panwalkar, S.S. & Iskander, W. (1977). Survey of Scheduling Rules. Operations Research, 25(1), 45-61.

Rajendran, C., Formulations and heuristics for scheduling in a kanban flowshop to minimize the sum of weighted flowtime, weighted tardiness and weighted earliness of containers, *International Journal of Production Research*, 1999; 37: 1137 – 1158.

Schaller, J.E. and J.M. Valente, "An evaluation of heuristics for scheduling a non-delay permutation flow shop with family setups to minimize total earliness and tardiness," *Journal of the Operational Research Society*, vol. 64, no. 6, p. 805, 2013.

Schaller, J.E. and J.M. Valente, "A comparison of metaheuristic procedures to schedule jobs in a permutation flow shop to minimise total earliness and tardiness," *The International Journal of Production Research*, vol. 51, no. 3, p. 772, 2013.

Schaller, J.E. and J.M. Valente, "Heuristics for scheduling jobs in a permutation flow shop to minimize total earliness with unforced idle time allowed," *Expert Systems With Applications*, vol. 119, pp. 76 – 386, 2019.

Smith, W.E. (1956). Various optimizers for single-stage production. Naval Research Logistics Quarterly, 3, 59-66.

Taillard, E., Benchmarks for basic scheduling problems, *European Journal of Operational Research*, 1993, 64: 278-285.

Valente, J.M.S. (2007). Dispatching heuristics for the single machine early/tardy scheduling problem with job-independent penalties. Computers & Industrial Engineering, 52(4), 434-447.

Valente, J. M. S., Beam search heuristics for the single machine scheduling problem with linear earliness and quadratic tardiness costs, *Asia-Pacific Journal of Operational Research*, 2009; 26: 319 – 339.

Valente, J.M.S. & Alves, R.A.F.S. (2005). Improved heuristics for the early/tardy scheduling problem with no idle time. Computers & Operations Research, 32(3), 557-569.

Vepsalainen, A.P.J. & Morton, T.E. (1987). Priority Rules for Job Shops with Weighted Tardiness Costs. Management Science, 33(8), 1035-1047.

Zegordi, S. H., K. Itoh, and T. Enkawa, A knowledgeable simulated annealing scheme for the early/tardy flow shop scheduling problem, *International Journal of Production Research*, 1997; 33: 1449 – 1466.

DECISION SCIENCES INSTITUTE

Inventory Balance Formulations and Extensions of the Deterministic Dynamic Lot Sizing Problem

Hadi Farhangi
Savannah State University
Email: farhangih@savannahstate.edu

Suman Niranjan
University of North Texas
Email: suman.niranjan@unt.edu

ABSTRACT

Dynamic Lot Sizing problem primarily decides on the production and inventory levels. This work studies the deterministic extensions of this problem by considering issues such as productions, inventories, capacities, bill of material, lead time, backorder, return and re-manufacturing, and transportation. The paper focuses on a brief review of different variations of the problem considering all mentioned issues and leaves the extensive literature review out of consideration. The solution methods are also disregarded in this study. The conclusion of paper hints on the possible directions to further explore variations of the dynamic lot sizing problem.

KEYWORDS: Dynamic lot sizing, Lead time, Backorder, Bill of material, Re-manufacturing

INTRODUCTION

Dynamic Lot Sizing (DLS) problem has variety applications in supply chain management. Since its first appearance in the work of (Wagner and Whitin, 1958), DLS problem has been extended and applied to a variety of industrial problems. DLS problem, as stated by (Wagner and Whitin, 1958), aims at finding the production and inventory levels at every period of planning horizon considering the total cost of production and inventory. There are two major streams to extend DLS problem which are the deterministic and stochastic streams. The stochastic demands and service level issues are studied in stochastic DLS problem and they are important for the producers and supply chain managers, for related works see (Bookbinder and Tan, 1988; Sox, 1997; Tarim and Kingsman, 2004; Helber and Sahling, 2010; Tempelmeier, 2013) among others. This paper considers the deterministic DLS and the related extensions.
The presented deterministic extensions are based on the inventory balance constraints. Our study exclude the formulations that are based on shortest path problem, transportation problem, and facility location problem. Interested readers may look at (Stadtler, 1997; Drexl and Kimms, 1996; Jayaraman and Pirkul, 2001; Brahimi et al., 2006; Denizel and Sural, 2006; Robinson et al., 2009; Helber and Sahling, 2010; Agrali, 2012; Schmid et al., 2013; Kian et al., 2014; Matta et al., 2015) for some examples of the alternative formulations. Using the inventory balance constraints, we focus on several areas such as DLS with multiple products, capacitated DLS, DLS and the assembly of sub-products (Bill of Material), DLS with backorder, DLS and lead time, DLS and multi-level productions, and DLS with a multi-level supply chain. Table 1 in the appendix summarizes a portion of the literature in these areas.

Table 1. DLS Extensions and related literature

Type of Extension	Related Literature
DLS	Wagner and Whitin (1958); Eppen and Martin (1987) Karmarkar et al. (1987); Wagelmans et al. (1992) Chang (2001); Karimi et al. (2003) Jaruphongsa et al. (2004); Li et al. (2004) Brahimi et al. (2006); Parveen and Haque (2007) Jans and Degraeve (2008); Li and Meissner (2011) Zhang et al. (2012); Dural-Selcuk et al. (2016)
Multi-Product and DLS	Afentakis et al. (1984); Karimi et al. (2003) Jans and Degraeve (2008); Agrali (2012)
Capacitated DLS	Maes and Wassenhove (1988); Drexl and Kimms (1996) Denizel and Sural (2006); Buschkuhl et al. (2010) Helber and Sahling (2010); Melega et al. (2018)
Bill of Material and DLS	Tempelmeier and Helber (1994); Kimms (1996) Tempelmeier and Derstroff (1996); Drexl and Kimms (1996) Stadtler (1997); Pitakaso et al. (2006); Sahling et al. (2009) Prakaiwichien and Rungreunganun (2018)
Backorder and DLS	Pochet and Wolsey (1988); Hung and Chien (2000) Lee et al. (2001); Staggemeier and Clark (2001) Jans and Degraeve (2004); Brahimi et al. (2006) Absi et al. (2011); Almeder and Almada-Lobo (2011)
Lead Time and DLS	Kimms (1993); Drexl and Kimms (1996) Kimms (1996); Hung and Chien (2000) Almeder et al. (2015)
Re-manufacturing and DLS	Teunter et al. (2006); Schulz (2011) Sifaleras et al. (2015)
Multi-Level Supply Chain and DLS	Diaby and Martel (1993); Barbarosoglu and Ozgur (1999) Staggemeier and Clark (2001); Jaruphongsa et al. (2004) Sambasivan and Yahya (2005); Brahimi et al. (2006) Bard and Nananukul (2009); Adulyasak (2012) Seyedhosseini and Ghoreyshi (2014) Adulyasak et al. (2015); Matta et al. (2015); Darvish (2017) Darvish and Coelho (2018); Wu et al. (2018)

The paper focuses on inventory and production issues commonly faced by firms. A company may produce several products and account for the capacity of inventory and the availability of time and resources. A manufacturer usually deals with the assembly of products and may consider the backlog and lead time in their manufacturing process, which are all the subject of this study. Furthermore, it is a general issue for the manufacturer to deal with the returned products and perform re-manufacturing processes which are another areas of consideration. Finally, producers manage several units of production and a network of supply chain that are the final subject of study.
This paper does not exhaust the complete literature of the DLS problems and their extensions. In addition, we do not cover the solution methods for the DLS problems. For reviews and surveys of the solution methods, we invite readers to look at (Bahl et al., 1987; Maes and Wassenhove, 1988; Karimi et al., 2003; Robinson et al., 2009; Buschkuhl et al., 2010) among others.

PROBLEM VARIATIONS

Dynamic Lot Sizing Problem

DLS problem decides on the inventory levels and production in each period of planning. This problem has been studied in the literature for decades and to present the basic formulation, we are committed to the notations similar to the ones that appear in the literature such as (Wagner and Whitin, 1958; Eppen and Martin, 1987; Wagelmans et al., 1992; Chang, 2001; Karimi et al., 2003). Let the set of periods be $T = \{1, 2, ..., |T|\}$, indexed by $t \in T$. We define inventory holding cost per unit per period as h^t, production cost per unit per period as p^t, and set up cost per unit per period as s^t. In addition, we have three sets of decision variables; x^t, q^t, and y^t. Variable x^t shows the amount of production during period t, variable q^t shows the inventory at the end of period t, and variable y^t is equal to 1 if $x^t > 0$, and 0 otherwise. Finally, demand during period t is shown by d^t and M^t is a big number (we set $M^t = d^t$). Then, the DLS problem can be formulated as follows:

$$\text{DLS:} \quad \min \sum_{t\in T} \left(h^t q^t + p^t x^t + s^t y^t\right) \tag{1}$$

$$\text{s.t.} \quad q^{t-1} + x^t = q^t + d^t \qquad \forall t \in T \tag{2}$$

$$x^t \leq M^t y^t \qquad \forall t \in T \tag{3}$$

$$x^t, q^t \geq 0,\ y^t \in \{0, 1\} \qquad \forall t \in T \tag{4}$$

Equation (1) aims at minimizing the total cost of inventories, productions, and set ups. Constraints (2) are inventory balance constraints. These constraints guarantee the conservation of the material flow within the production-inventory unit; the input to the unit (previous inventory and production) should be equal to the output inventory (demand and current inventory). These constraints are the core of DLS problems. Constraints (3) force the set up variables to be equal to one, if there is production. At last, constraints (4) define the variables.
This problem formulation, expressed in (1) - (4), appear in many works in the literature. Interested readers may look at the works of (Wagner and Whitin, 1958; Eppen and Martin, 1987; Wagelmans et al., 1992; Chang, 2001; Karimi et al., 2003; Jaruphongsa et al., 2004; Li et al., 2004; Brahimi et al., 2006; Parveen and Haque, 2007; Jans and Degraeve, 2008; Li and Meissner, 2011; Zhang et al., 2012; Dural-Selcuk et al., 2016). This formulation, however, does not consider the start up cost of the production unit. Assume the model returns a solution that forces

no production at some periods; once the production resume at some later periods, a start up cost may incur (Karmarkar et al., 1987). We use similar notation to (Karmarkar et al., 1987) to capture this cost. Let o^t be the start up cost during period t and z^t be equal to one, if the unit is off-line during period $t-1$ but operational during period t. Then, the following constraints will be added to the model (Karmarkar et al., 1987):

$$z^t \geq y^t - y^{t-1} \quad \forall t \in T \setminus \{1\} \tag{5}$$

the problem, then, can be stated as follows:

$$\begin{aligned}
\text{DLS:} \quad & \min \sum_{t \in T} \left(h^t q^t + p^t x^t + s^t y^t + o^t z^t \right) && && (6) \\
\text{s.t.} \quad & q^{t-1} + x^t = q^t + d^t && \forall t \in T && (7) \\
& x^t \leq M^t y^t && \forall t \in T && (8) \\
& z^t \geq y^t - y^{t-1} && \forall t \in T \setminus \{1\} && \\
& x^t, q^t \geq 0,\, y^t \in \{0,1\} && \forall t \in T && (9) \\
& z^t \in \{0,1\} && \forall t \in T \setminus \{1\} && (10)
\end{aligned}$$

where total cost is updated as (6); the rest of inequalities is the same as the previous version in (1) - (4). Constraints (7) are inventory balance constraints. These constraints guarantee the conservation of the material flow. Constraints (8) force the set up variables to be equal to one, if there is production. At last, constraints (9)-(10) are defining the variables.

Multi-Product DLS Problems

DLS problem formulated in (6) - (10) accounts only for one product. Extending this problem for multiple products is important as manufacturers may produce more than one product type practice. We refer to this extension as Multi-Product DLS problems or MPDLS problems in short. Let $I = \{1, ..., |I|\}$, indexed by i, be the set of products. Moreover, let the holding cost, set up cost, production cost, and start up cost of product i per unit per period be h_i^t, s_i^t, p_i^t, and o_i^t respectively. We also define four sets of variables x_i^t, q_i^t, y_i^t, and z_i^t. Variable x_i^t shows the production of product i during period t and variable q_i^t shows the inventory of product i at the end of period t. Variable $y_i^t = 1$, if there is a production of product i during period t and $y_i^t = 0$, otherwise. Variable $z_i^t = 1$, if there is a production for product i during period t and not during period $t-1$ and $z_i^t = 0$, otherwise. Finally, the demand of product i during period t is denoted by d_i^t and M^t is a big number (we set $M^t = \sum_{i \in I} d_i^t$). Given these definitions, a MPDLS problem can be formulated as follows:

$$\begin{aligned}
\text{MPDLS:} \quad & \min \sum_{t \in T} \sum_{i \in I} \left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^t + o_i^t z_i^t \right) && && (11) \\
\text{s.t.} \quad & q_i^{t-1} + x_i^t = q_i^t + d_i^t && \forall i \in I, \forall t \in T && (12) \\
& x_i^t \leq M^t y_i^t && \forall i \in I, \forall t \in T && (13) \\
& z_i^t \geq y_i^t - y_i^{t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (14) \\
& x_i^t, q_i^t \geq 0,\, y_i^t \in \{0,1\} && \forall i \in I, \forall t \in T && (15)
\end{aligned}$$

$$z_i^t \in \{0,1\} \qquad \forall i \in I, \forall t \in T \setminus \{1\} \tag{16}$$

Equation (11) minimizes the total cost of inventories, productions, set ups, and start ups. Constraints (12) are the inventory balance constraints, similar to (2). These constraints guarantee the flow conservation of material within production-inventory units for every product at each period. Similar to constraints (3), constraints (13) force set up variable to be equal to one, if there is production. Constraints (14) enforce start up, if there is a set up at the new period. At last, constraints (15)-(16) are defining the variables. Similar formulation appeared in the literature such as (Afentakis et al., 1984; Karimi et al., 2003; Jans and Degraeve, 2008; Agrali, 2012).
There are more issues to consider for MPDLS problems. For example, the production unit may not have enough machinery to produce more than one product at each period, e.g. they may only have one machine/worker. As the result, one product should be produced at each period, at most (Jans and Degraeve, 2008). This restriction may be shown in the model in the following constraints (Jans and Degraeve, 2008):

$$\sum_{i \in I} y_i^t \leq 1 \quad \forall t \in T \tag{17}$$

Another issue is the capacity of the production unit. For example, the amount of time available for the production may be limited as the production unit may not work continuously through all periods, the resources for the production may be limited, the inventory capacity may be limited, and etc. To formulate these issues, we first define appropriate parameters. Let b_i^t be the time required to produce one unit of product i during period t, r_i^t be the amount of resources required to produce one unit of product i during period t, and v_i^t be the volume of one unit of product i during period t. In addition, let τ^t, R^t, and V^t be the total available time, total available resources, and total available inventory space for products during period t. The capacity constraints can now be formulated as follows:

$$\sum_{i \in I} b_i^t x_i^t \leq \tau^t \quad \forall t \in T \tag{18}$$

$$\sum_{i \in I} r_i^t x_i^t \leq R^t \quad \forall t \in T \tag{19}$$

$$\sum_{i \in I} v_i^t q_i^t \leq V^t \quad \forall t \in T \tag{20}$$

Constraints (18) are time restriction constraints. These types of constraints appeared in the work of (Denizel and Sural, 2006; Helber and Sahling, 2010; Melega et al., 2018). Constraints (18) are the resource restriction constraints. These types of constraints are very common in the literature of DLS problems. Interested readers may look at the review papers such as (Maes and Wassenhove, 1988; Drexl and Kimms, 1996; Buschkuhl et al., 2010). Constraints (20) are the inventory capacity constraints.

Bill of Materials and DLS Problems

A production unit may produce several sub-products first and then assemble them to the final product through a complex assembly process. In the production process of such products, manufacturers use Bill of Material (BOM) to show the structure of assembly and manufacturing processes (Stevenson, 2018). Examples of BOMs can be seen in (Tempelmeier and Helber, 1994;

Tempelmeier and Derstroff, 1996; Stevenson, 2018). BOM shows the successor relation, as well as the quantity of product that is needed for the successor product. Let the parameter a_{ij} be the number of product i that is needed to produce one unit of its successor; product j (Tempelmeier and Helber, 1994; Kimms, 1996; Tempelmeier and Derstroff, 1996; Drexl and Kimms, 1996; Stadtler, 1997). Note that $a_{ij} = 0$ means there is no immediate relation between products and the corresponding matrix to a_{ij} values is not symmetric. Using the same variables and parameters introduced previously, we formulate the Capacitated Bill of Material DLS (CBMDLS) problem as follows:

$$
\begin{aligned}
\text{CBMDLS:}\quad \min\quad & \sum_{t\in T}\sum_{i\in I}\left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^t + o_i^t z_i^t\right) && && (21)\\
\text{s.t.}\quad & q_i^{t-1} + x_i^t = q_i^t + d_i^t + \sum_{j\in I} a_{ij} x_j^t && \forall i \in I, \forall t \in T && (22)\\
& x_i^t \le M^t y_i^t && \forall i \in I, \forall t \in T && (23)\\
& z_i^t \ge y_i^t - y_i^{t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (24)\\
& \sum_{i\in I} b_i^t x_i^t \le \tau^t && \forall t \in T && (25)\\
& \sum_{i\in I} r_i^t x_i^t \le R^t && \forall t \in T && (26)\\
& \sum_{i\in I} v_i^t q_i^t \le V^t && \forall t \in T && (27)\\
& x_i^t, q_i^t \ge 0,\ y_i^t \in \{0,1\} && \forall i \in I, \forall t \in T && (28)\\
& z_i^t \in \{0,1\} && \forall i \in I, \forall t \in T \setminus \{1\} && (29)
\end{aligned}
$$

The objective function, equation (21), and all the constraints of CBMDLS problem are similar to the MPDLS problem, except for constraints (22). Constraints (22) are inventory balance constraints and they appeared in the previous works, such as (Kimms, 1993; Tempelmeier and Helber, 1994; Kimms, 1996; Tempelmeier and Derstroff, 1996; Drexl and Kimms, 1996; Pitakaso et al., 2006; Sahling et al., 2009; Prakaiwichien and Rungreunganun, 2018). These constraints have an additional element, $\sum_{j\in I} a_{ij} x_j^t$, that shows the demand that is derived for the product i based on the quantity that its successors need for their own production (internal demand). The external demand, i.e. the demand coming from a external customer, is usually zero for the products with successors and only the final products are subject to external demand (Hung and Chien, 2000). As the result, for most products with successors, we have $d_i^t = 0$. Constraints (22) are inventory balance constraints. Constraints (23) force set up variable to be equal to one, if there is production. Constraints (24) enforce start up, if there is set up at the new period. Constraints (25) are time restriction constraints, constraints (26) are resource restriction constraints, and constraints (27) are inventory capacity constraints. At last, constraints (28)-(29) are defining the variables.

Backorder and DLS

Backorder happens in a production site when there is not enough production to satisfy the demand. As a consequence, the orders backlog to the next periods. In DLS problem, backorder is formulated by the additional variables and the updated inventory balance constraints. For example, the work of (Pochet and Wolsey, 1988; Lee et al., 2001; Staggemeier and Clark, 2001;

Brahimi et al., 2006; Absi et al., 2011) update the inventory balanced constraints as follows to reflect the backorder in the model, where u_i^t is the backorder amount of product i during period t:

$$q_i^{t-1} - u_i^{t-1} + x_i^t = q_i^t - u_i^t + d_i^t \quad \forall i \in I, \forall t \in T \tag{30}$$

By defining the backorder penalty cost g_i^t as the backorder cost of product i during period t and using the same notations and variables as before, we can update the CBMDLS problem as Backorder CBMDLS (BCBMDLS) problem reflected in the following:

$$\begin{aligned}
\text{BCBMDLS:} \quad \min \quad & \sum_{t\in T}\sum_{i\in I} \left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^t + o_i^t z_i^t + g_i^t u_i^t\right) && && (31)\\
\text{s.t.} \quad & q_i^{t-1} - u_i^{t-1} + x_i^t = q_i^t - u_i^t + d_i^t + \sum_{j\in I} a_{ij} x_j^t && \forall i \in I, \forall t \in T && (32)\\
& x_i^t \le M^t y_i^t && \forall i \in I, \forall t \in T && (33)\\
& z_i^t \ge y_i^t - y_i^{t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (34)\\
& \sum_{i\in I} b_i^t x_i^t \le \tau^t && \forall t \in T && (35)\\
& \sum_{i\in I} r_i^t x_i^t \le R^t && \forall t \in T && (36)\\
& \sum_{i\in I} v_i^t q_i^t \le V^t && \forall t \in T && (37)\\
& x_i^t, q_i^t, u_i^t \ge 0,\ y_i^t \in \{0,1\} && \forall i \in I, \forall t \in T && (38)\\
& z_i^t \in \{0,1\} && \forall i \in I, \forall t \in T \setminus \{1\} && (39)
\end{aligned}$$

The only updated parts in BCBMDLS compared to CBMDLS are constraints (32) and the objective function (31). Constraints (32) are the updated inventory balance constraints in the presence of BOM that are also appeared in (Hung and Chien, 2000; Almeder and Almada-Lobo, 2011). Constraints (33)-(39) are similar to constraints (23)-(29) considering the similar restrictions that they impose.

Lead Time and DLS Problem

DLS problems, as described before, assume a zero lead time. Zero lead time means the product is immediately available upon order. It also means that in an assembly line, the assembly operations and motion of the product takes zero time. Failing to incorporate lead time in the DLS problems results in unavailability of products that are needed for the production of their successors in the earlier periods of planning (Almeder et al., 2015). To prevent this issue, several researchers presented some solutions. One of the earliest solutions is to force the inventory at each period to be greater than the sum of what is needed in the future periods. Particularly, the work of (Kimms, 1993; Drexl and Kimms, 1996; Kimms, 1996) add the following constraints to the model to accompany the inventory balance constraints by defining l_i as the lead time of product i:

$$q_i^t \ge \sum_{j\in I} \sum_{r=t}^{\min\{t+l_i-1,|T|\}} a_{ij} x_j^t \quad \forall i \in I, \forall t \in T \tag{40}$$

Alternatively, one can update the inventory balance equation to reflect the lead time as follows (Hung and Chien, 2000):

$$q_i^{t-1} - u_i^{t-1} + x_i^t = q_i^t - u_i^t + d_i^t + \sum_{j \in I} a_{ij} x_j^{t+l_i} \quad \forall i \in I, \forall t \in T \tag{41}$$

As it can be seen, constraints (41) update the time index of the successive products; the demand of successive products of i should shift by l_i. We call these kind of problems Lead Time BCB-MDLS or LTBCBMDLS in short. A decision maker needs to either add constraints (40) to the model or replace inventory balance constraints by constraints (41). We name the problem in the first case as LTBCBMDLS1 and the problem in the second case as LTBCBMDLS2. The LTBCB-MDLS1 is formulated as follows:

$$
\begin{aligned}
\text{LTBCBMDLS1:} \quad \min \quad & \sum_{t \in T} \sum_{i \in I} \left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^t + o_i^t z_i^t + g_i^t u_i^t \right) && && (42)\\
\text{s.t.} \quad & q_i^{t-1} - u_i^{t-1} + x_i^t = q_i^t - u_i^t + d_i^t + \sum_{j \in I} a_{ij} x_j^t && \forall i \in I, \forall t \in T && (43)\\
& q_i^t \geq \sum_{j \in I} \sum_{r=t}^{\min\{t+l_i-1, |T|\}} a_{ij} x_j^t && \forall i \in I, \forall t \in T && (44)\\
& x_i^t \leq M^t y_i^t && \forall i \in I, \forall t \in T && (45)\\
& z_i^t \geq y_i^t - y_i^{t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (46)\\
& \sum_{i \in I} b_i^t x_i^t \leq \tau^t && \forall t \in T && (47)\\
& \sum_{i \in I} r_i^t x_i^t \leq R^t && \forall t \in T && (48)\\
& \sum_{i \in I} v_i^t q_i^t \leq V^t && \forall t \in T && (49)\\
& x_i^t, q_i^t, u_i^t \geq 0,\, y_i^t \in \{0,1\} && \forall i \in I, \forall t \in T && (50)\\
& z_i^t \in \{0,1\} && \forall i \in I, \forall t \in T \setminus \{1\} && (51)
\end{aligned}
$$

Constraints (44) impose the lead time to the model. The rest of constraints (42), (43), and (45)-(51) are the same as constraints (31) and (33)-(39). Alternatively, we can incorporate the lead time by defining LTBCBMDLS2 problem as the following:

$$
\begin{aligned}
\text{LTBCBMDLS2:} \quad \min \quad & \sum_{t \in T} \sum_{i \in I} \left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^t + o_i^t z_i^t + g_i^t u_i^t \right) && && (52)\\
\text{s.t.} \quad & q_i^{t-1} - u_i^{t-1} + x_i^t = q_i^t - u_i^t + d_i^t + \sum_{j \in I} a_{ij} x_j^{t+l_i} && \forall i \in I, \forall t \in T && (53)\\
& x_i^t \leq M^t y_i^t && \forall i \in I, \forall t \in T && (54)\\
& z_i^t \geq y_i^t - y_i^{t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (55)\\
& \sum_{i \in I} b_i^t x_i^t \leq \tau^t && \forall t \in T && (56)\\
& \sum_{i \in I} r_i^t x_i^t \leq R^t && \forall t \in T && (57)\\
& \sum_{i \in I} v_i^t q_i^t \leq V^t && \forall t \in T && (58)\\
& x_i^t, q_i^t, u_i^t \geq 0,\, y_i^t \in \{0,1\} && \forall i \in I, \forall t \in T && (59)\\
& z_i^t \in \{0,1\} && \forall i \in I, \forall t \in T \setminus \{1\} && (60)
\end{aligned}
$$

Constraints (52)-(60) are the same as (31)-(39). The only difference is the inventory balance constraints, (53), that is updated based on (41).
It is worth to note that in real life the lead times are not deterministic; rather, they are stochastic and stochastic methods should be used for formulating lead time in DLS problems. The work of (Huang and KucuKyavuz, 2008) is dedicated to the stochastic lead times, which we do not explore this case.

Re-Manufacturing and DLS Problem

We modify MPDLS problem to account for the returns and re-manufacturing activities. In most of works in the literature, only single item DLS problem, expressed in (1) - (4), is extended for re-manufacturing. For related works, interested readers may look at (Teunter et al., 2006; Schulz, 2011; Sifaleras et al., 2015). In addition to the parameters and variables introduced for MPDLS problem, we define holding cost of the returned product i during period t as Rh_i^t, set up cost of the returned product i during period t as Rs_i^t, re-manufacturing cost of the returned product i during period t as Rp_i^t, start up cost of the returned product i during period t as Ro_i^t, the number of the returned product i during period t as Rd_i^t, the required time to re-manufacture one unit of product i during period t as Rb_i^t, and the amount of resources for re-manufacturing one unit of product i during period t as Rr_i^t. In addition, let ω_i^t be the re-manufacturing quantity of product i during period t and μ_i^t be the inventory of the returned product i at the end of period t. We modify the variable y_i^t in MPDLS to $y_i^{1,t}$ and introduce the variable $y_i^{2,t}$ to show the set up for re-manufacturing: $y_i^{2,t} = 1$ if the machine sets up to re-manufacture product i during period t, $y_i^{2,t} = 0$ otherwise. Similarly, we modify the variable z_i^t in MPDLS to $z_i^{1,t}$ and introduce the variable $z_i^{2,t}$ to show the start up for re-manufacturing: $z_i^{2,t} = 1$ if the machine starts up to re-manufacture product i during period t and while it was not operational for product i in previous periods, $z_i^{2,t} = 0$ otherwise. Then the MPDLS can be reformulated for the Re-Manufacturing MPDLS (RMPDLS) as follows:

$$
\begin{aligned}
\text{RMPDLS:} \quad \min \quad & \sum_{t\in T}\sum_{i\in I}\left(h_i^t q_i^t + p_i^t x_i^t + s_i^t y_i^{1,t} + o_i^t z_i^{1,t} + Rh_i^t \mu_i^t + Rp_i^t \omega_i^t + Rs_i^t y_i^{2,t} + Ro_i^t z_i^{2,t}\right) && && (61)\\
\text{s.t.} \quad & q_i^{t-1} + x_i^t + \omega_i^t = q_i^t + d_i^t && \forall i \in I, \forall t \in T && (62)\\
& \mu_i^{t-1} + Rd_i^t = \mu_i^t + \omega_i^t && \forall i \in I, \forall t \in T && (63)\\
& x_i^t \le M_i^t y_i^{1,t} && \forall i \in I, \forall t \in T && (64)\\
& \omega_i^t \le M^t y_i^{2,t} && \forall i \in I, \forall t \in T && (65)\\
& y_i^{1,t} + y_i^{2,t} \le 1 && \forall i \in I, \forall t \in T && (66)\\
& z_i^{1,t} \ge y_i^{1,t} - y_i^{1,t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (67)\\
& z_i^{2,t} \ge y_i^{2,t} - y_i^{2,t-1} && \forall i \in I, \forall t \in T \setminus \{1\} && (68)\\
& \sum_{i\in I}(b_i^t x_i^t + Rb_i^t \omega_i^t) \le \tau^t && \forall t \in T && (69)\\
& \sum_{i\in I}(r_i^t x_i^t + Rr_i^t \omega_i^t) \le R^t && \forall t \in T && (70)\\
& \sum_{i\in I}(v_i^t q_i^t + v_i^t \mu_i^t) \le V^t && \forall t \in T && (71)\\
& x_i^t, q_i^t, \omega_i^t, \mu_i^t \ge 0,\ y_i^{1,t} \in \{0,1\},\ y_i^{2,t} \in \{0,1\} && \forall i \in I, \forall t \in T && (72)\\
& z_i^{1,t} \in \{0,1\},\ z_i^{2,t} \in \{0,1\} && \forall i \in I, \forall t \in T \setminus \{1\} && (73)
\end{aligned}
$$

Objective function (61) minimizes total cost of inventories, productions, set ups, and start ups costs for both manufactured products and re-manufactured products. Constraints (62) are the updated inventory balance constraints to reflect re-manufacturing. Constraints (63) are the inventory balanced constraints for the returned and re-manufactured products. Constraints (64) - (65) guarantee that at each period if we manufacture or re-manufacture we have to set up the process. Constraints (66) enforce that at each period we either manufacture or re-manufacture. Constraints (67) - (68) force a start up if we manufacture or re-manufacture at each period, but not in previous periods. Constraints (69) - (71) are capacity constraints. Constraints (72) - (73) define the problem variables.

Multi-Level Production and DLS Problem

Having multiple manufacturing units and warehouses requires modifications to the DLS problems. Multiple manufacturing units and warehouses/distribution centers coincide with transportation. Variety of publications in DLS problem account for these transportations by modifying DLS constraints and objective function. In (Staggemeier and Clark, 2001; Wu et al., 2018), authors consider multiple machines/units for production. In their case, every machine/unit can produce every item; hence, the extension of DLS problem only reflects on an update of the inventory balance equations. The work of (Sambasivan and Yahya, 2005; Brahimi et al., 2006) also consider the transportation of products between plants.
To formulate this problem, let $K = \{1, ..., |K|\}$, indexed by $k \in K$, be the set of production units/machines. Furthermore, Let the set up cost, production cost, and start up cost of product i in unit k per period be s_i^{kt}, p_i^{kt}, and o_i^{kt}, respectively. Moreover, let b_i^{kt} be the time required to produce one unit of product i in unit k per period t, r_i^{kt} be the amount of resources required to produce one unit of product i in unit k per period t, and v_i^t be the volume of one unit of product i during period t. In addition, let τ^{kt} and R^{kt} be the total available time and total available resources available for products in unit k per period t. Also we define variable x_i^{kt} as the production of product i by unit k during period t. Variable $y_i^{kt} = 1$, if there is a production of product i during period t by unit k and $y_i^{kt} = 0$, otherwise. Variable $z_i^{kt} = 1$, if there is a production of product i during period t by unit k and not during period $t-1$ and $z_i^{kt} = 0$, otherwise. The rest of parameters and variables are similar to MPDLS problem. If we assume there is no transportation between production units and all units use the same inventory space, the problem is Multi-Unit DLS (MUDLS) and it can be formulated as follows (Staggemeier and Clark, 2001; Wu et al., 2018):

$$\text{MUDLS:} \quad \min \sum_{t\in T}\sum_{i\in I}\sum_{k\in K}\left(h_i^t q_i^t + p_i^{kt}x_i^{kt} + s_i^{kt}y_i^{kt} + o_i^{kt}z_i^{kt}\right) \tag{74}$$

$$\text{s.t.} \quad q_i^{t-1} + \sum_{k\in K} x_i^{kt} = q_i^t + d_i^t \qquad \forall i \in I, \forall t \in T \tag{75}$$

$$\sum_{i\in I}\sum_{k\in K} b_i^t x_i^{kt} \le \tau^t \qquad \forall t \in T \tag{76}$$

$$\sum_{i\in I}\sum_{k\in K} r_i^t x_i^{kt} \le R^t \qquad \forall t \in T \tag{77}$$

$$\sum_{i\in I} v_i^t q_i^t \le V^t \qquad \forall t \in T \tag{78}$$

$$x_i^{kt} \le M^t y_i^{kt} \qquad \forall k \in K, \forall i \in I, \forall t \in T \tag{79}$$

$$z_i^{kt} \ge y_i^{kt} - y_i^{k,t-1} \qquad \forall k \in K, \forall i \in I, \forall t \in T \setminus \{1\} \tag{80}$$

$$x_i^{kt}, q_i^t \ge 0,\ y_i^{kt} \in \{0,1\} \qquad \forall k \in K, \forall i \in I, \forall t \in T \tag{81}$$

$$z_i^{kt} \in \{0,1\} \qquad \forall k \in K, \forall i \in I, \forall t \in T \setminus \{1\} \tag{82}$$

Equation (74) minimizes total cost of inventory, production, set up, and start up. Constraints (75) are the modified inventory balance constraints compared to constraints (12). These constraints guarantee the flow conservation of material within production-inventory units for every product at each period by all production units. Constraints (76) are time restriction constraints, constraints (77) are resource restriction constraints, and constraints (78) are inventory capacity constraints. Constraints (79) force set up variable to be equal to one, if there is a production by a unit. Constraints (80) force the start up variable for every production unit to be equal to 1, if we start production in a period after an idle period. At last, constraints (81) - (82) define the variables.

Multi-Level Supply Chain and DLS Problem

Problem MPDLS does not account for transportation and independent production units. In this section we formulate a multi-level Supply Chain DLS (SCDLS) problem to address multiple production units and transportation through a supply chain. For this purpose, we consider a graph that shows the production units, warehouses, distributions centers, and retailers as nodes. The transportation links between these nodes are the edges of the graph. The study of the DLS in a production/distribution network has a long history. For example, the works of (Jaruphongsa et al., 2004; Bard and Nananukul, 2009; Seyedhosseini and Ghoreyshi, 2014; Adulyasak et al., 2015) study a single production unit and multiple distribution centers. Moreover, the works of (Diaby and Martel, 1993; Barbarosoglu and Ozgur, 1999) present a model for multiple production units and multiple distribution centers. Furthermore, (Darvish, 2017; Darvish and Coelho, 2018) study multiple production units and multi-level of supply chain (distribution centers and retailers). The work of (Adulyasak, 2012) accounts for a supply chain that is presented as a complete graph. Also the graph approach has been used for the previous studies such as (Matta et al., 2015), this work accounts for any given supply chain network. We use similar notation to (Darvish, 2017; Darvish and Coelho, 2018) to formulate the problem.
Let $E = \{1, ..., |E|\}$, indexed by $e, f \in E$, be the set of nodes of the supply chain network. For the edges of this directed and connected graph, let $\gamma^t_{ef} = 1$ if node e is connected to node f, $\gamma^t_{ef} = 0$ otherwise ($\gamma^t_{ee} = 0, \forall t \in T$). In addition, let $E^p \subset E$ be the set of production units, $E^d \subset E$ be the set of distribution centers and warehouses, and $E^r \subset E$ be the set of retailers (customers). Note that $E^p \cup E^d \cup E^r \equiv E$. Furthermore, let c^t_{ef} be the transportation cost of the product i from node e to node f during period t, h^t_{ie} be the holding cost of the product i in node $e \in E$ during period t, s^t_{ie} be the set up cost of the product i in node $e \in E^p$ during period t, o^t_{ie} be the start up cost of the product i in node $e \in E^p$ during period t, p^t_{ie} be the production cost of the product i in node $e \in E^p$ during period t, and d^t_{ie} be the external demand of the product i in node $e \in E^r$ during period t. Moreover, let τ^t_e be the available time for production in $e \in E^p$ during period t, R^t_e be the available resources for production in $e \in E^p$ during period t, and V^t_e be the available inventory space in $e \in E$ during period t. Finally, we use five variables in the formulation; x^t_{ief}, q^t_{ie}, P^t_{ie}, y^t_{ie}, and z^t_{ie}. Variable x^t_{ief} shows the amount of transportation of product i from node e to f during period t, variable q^t_{ie} shows the amount of inventory of product i in node e at the end of period t, and variable P^t_{ie} shows the production of item i in node $e \in E^p$ during period t. The variable $y^t_{ie} = 1$ shows that there is a set up to produce i in node $e \in E^p$ during period t, $y^t_{ie} = 0$ otherwise. And, variable $z^t_{ie} = 1$, if there is a start up to produce i in node $e \in E^p$ during period t if there was no production of i during period $t-1$, $z^t_{ie} = 0$ otherwise. Now, we formulate SCDLS as follows:

SCDLS:

$$\begin{aligned}
\min \quad & \sum_{t\in T}\sum_{i\in I}\left(\sum_{e\in E} h_{ie}^t q_{ie}^t + \sum_{e\in E^p} s_{ie}^t y_{ie}^t + \sum_{e\in E^p} o_{ie}^t z_{ie}^t +\right) \\
& + \sum_{t\in T}\sum_{i\in I}\left(\sum_{e\in E^p}\sum_{f\in E^d\cup E^r} \gamma_{ef}^t p_{ie}^t x_{ief}^t + \sum_{e\in E}\sum_{f\in E} \gamma_{ef}^t c_{ef}^t x_{ief}^t\right) && (83)
\end{aligned}$$

$$\text{s.t.} \quad q_{ie}^{t-1} + P_{ie}^t = \sum_{f\in E^d\cup E^r} \gamma_{ef}^t x_{ief}^t + q_{ie}^t \qquad \forall e\in E^p, \forall i\in I, \forall t\in T \quad (84)$$

$$q_{ie}^{t-1} + \sum_{f\in E^p\cup E^d} \gamma_{fe}^t x_{ife}^t = q_{ie}^t + \sum_{f\in E^r\cup E^d} \gamma_{ef}^t x_{ief}^t \qquad \forall e\in E^d, \forall i\in I, \forall t\in T \quad (85)$$

$$q_{ie}^{t-1} + \sum_{f\in E^p\cup E^d} \gamma_{fe}^t x_{ife}^t = q_{ie}^t + d_{ie}^t \qquad \forall e\in E^r, \forall i\in I, \forall t\in T \quad (86)$$

$$q_{ie}^{t-1} + \sum_{f\in E^p\cup E^d} \gamma_{fe}^t x_{ife}^t = q_{ie}^t + \sum_{f\in E^r\cup E^d} \gamma_{ef}^t x_{ief}^t + d_{ie}^t \qquad \forall e\in E^r\cap E^d, \forall i\in I, \forall t\in T \quad (87)$$

$$\sum_{i\in I}\sum_{f\in E^d\cup E^r} b_i^t P_{ie}^t \le \tau_e^t \qquad \forall e\in E^p, \forall t\in T \quad (88)$$

$$\sum_{i\in I}\sum_{f\in E^d\cup E^r} r_i^t P_{ie}^t \le R_e^t \qquad \forall e\in E^p, \forall t\in T \quad (89)$$

$$\sum_{i\in I} v_i^t q_{ie}^t \le V_e^t \qquad \forall e\in E, \forall t\in T \quad (90)$$

$$P_{ie}^t \le M^t y_{ie}^t \qquad \forall e\in E^p, \forall i\in I, \forall t\in T \quad (91)$$

$$z_{ie}^t \ge y_{ie}^t - y_{ie}^{t-1} \qquad \forall e\in E^p, \forall i\in I, \forall t\in T\setminus\{1\} \quad (92)$$

$$x_{ief}^t, q_{ie}^t, P_{ie}^t \ge 0 \qquad \forall e,f\in E, \forall i\in I, \forall t\in T \quad (93)$$

$$y_{ie}^t \in \{0,1\} \qquad \forall e\in E^p, \forall i\in I, \forall t\in T \quad (94)$$

$$z_{ie}^t \in \{0,1\} \qquad \forall e\in E^p, \forall i\in I, \forall t\in T\setminus\{1\} \quad (95)$$

Equation (83) aims to minimize the sum of holding costs, set up costs, start up costs, production costs, and transportation costs. Constraints (84) are the inventory balance constraints for every production node. Constraints (85) are the inventory balance constraints for every distribution/warehouse node. Constraints (86) are the inventory balance constraints for every retailer node. Constraints (87) are the inventory balance constraints for every node that acts as both distributor and retailer. Constraints (88) - (90) are the capacity constraints. Constraints (91) guarantee that each production node should set up the production, if there is a production for products at each period. Constraints (92) force a production start up for the production nodes, if we produce at each period, but not in previous periods. Constraints (93) - (95) are defining variables. Problem SCDLS is multi-echelon and accounts for the inventory, set up, start up, capacity, multiple items, production, and transportation. This problem, however, does not account for the returned items, backorders, assembly procedures (Bill of materials), lead times, and re-manufacturing.

CONCLUSION AND FUTURE RESEARCH

In this paper some generic DLS problem formulations are presented. Variety of DLS problem extensions on multiple products, capacity, bill of materials, backorder, lead time, and multi levels of supply and productions are discussed. DLS problem can be further extended to combine assembly of components (bill of material) and the supply chain (multi-level). Alternative formulations of all these models along with the appropriate exact and approximation algorithms is another future direction for this research. In addition, incorporating the stochastic nature of demands, service levels, and lead times to the model is another future direction for the research in DLS problems.

APPENDIX 1: NOTATION

List of Sets and Indices	
T	: the set of time periods $T = \{1, 2, \ldots, \lvert T\rvert\}$
t, r	: indices used for time, $t, r \in T$
I	: the set of products $I = \{1, 2, \ldots, \lvert I\rvert\}$
i, j	: indices used for products, $i, j \in I$
K	: the set of production units $K = \{1, 2, \ldots, \lvert K\rvert\}$
k	: index used for production units, $k \in K$
E	: the set of network nodes $E = \{1, 2, \ldots, \lvert E\rvert\}$
E^p	: the set of production nodes
E^d	: the set of warehouses/distribution nodes
E^r	: the set of retailer nodes
e, f	: indices used for network nodes, $e, f \in E$

List of Parameters	
h^t	: holding cost per unit during period t
s^t	: set up cost of production during period t
p^t	: production cost during period t
o^t	: start up cost of production during period t
d^t	: demand during period t
h_i^t	: holding cost of product i per unit during period t
s_i^t	: set up cost to produce i during period t
p_i^t	: production cost of product i per unit during period t
o_i^t	: start up cost to produce i during period t
d_i^t	: demand of product i during period t
h_i^{kt}	: holding cost of product i per unit during period t by unit k
s_i^{kt}	: set up cost to produce i during period t by unit k
p_i^{kt}	: production cost of product i per unit during period t by unit k
o_i^{kt}	: start up cost to produce i during period t by unit k
Rh_i^t	: holding cost of returned product i per unit during period t
Rs_i^t	: set up cost to re-manufacture i during period t
Rp_i^t	: re-manufacture cost of i during period t
Ro_i^t	: start up cost to re-manufacture i during period t
h_{ie}^t	: holding cost of product i in node e during period t
s_{ie}^t	: set up cost to produce i in node e during period t

p_{ie}^t : production cost to produce i in node e during period t
o_{ie}^t : start up cost to produce i in node e during period t
c_{ef}^t : transportation cost from node e to f during period t
d_{ie}^t : demand of product i in node e during period t
b_i^t : required time to produce one unit of i during period t
r_i^t : required resources to produce one unit of i during period t
v_i^t : product i volume during period t
Rd_i^t : number of returned product i during period t
Rb_i^t : required time to re-manufacture one unit of i during period t
Rr_i^t : required resources to re-manufacture one unit of i during period t
b_i^{kt} : required time to produce one unit of i during period t by unit k
r_i^{kt} : required resources to produce one unit of i during period t by unit k
τ^t : available production time during period t
R^t : available resources during period t
V^t : inventory space during period t
τ^{kt} : available production time during period t in unit k
R^{kt} : available resources during period t in unit k
V^{kt} : inventory space during period t in unit k
τ_e^t : available production time during period t in node e
R_e^t : available resources during period t in node e
V_e^t : inventory space during period t in node e
a_{ij}^t : number of product i needed to produce a unit of product j during period t
γ_{ef}^t : $\gamma_{ef}^t = 1$ if node e is linked to f, $\gamma_{ef}^t = 0$ otherwise (directed graph parameter)
g_i^t : backorder cost of product i during period t
L_i : lead time of product i

List of Variables

x^t : production quantity during period t
q^t : inventory at the end of period t
y^t : $y^t = 1$ if there is production during period t, $y^t = 0$ otherwise
z^t : $z^t = 1$ if there is production at t and no production at $t-1$, $z^t = 0$ otherwise
x_i^t : production quantity of product i during period t
q_i^t : inventory of product i at the end of period t
y_i^t : $y_i^t = 1$ if we produce i during period t, $y_i^t = 0$ otherwise
z_i^t : $z_i^t = 1$ if we produce i at t and no production of i at $t-1$, $z_i^t = 0$ otherwise
x_i^{kt} : production quantity of product i during period t by unit k
q_i^{kt} : inventory of product i at the end of period t in unit k
y_i^{kt} : $y_i^{kt} = 1$ if we produce i during period t by unit k, $y_i^{kt} = 0$ otherwise
z_i^{kt} : $z_i^{kt} = 1$ if we produce i at t and no production of i at $t-1$ by unit k, $z_i^{kt} = 0$ otherwise
x_{ief}^t : amount of transported product i from e to f during period t
q_{ie}^t : inventory of product i in node e at the end of period t
y_{ie}^t : $y_{ie}^t = 1$ if we produce i during period t in node e, $y_{ie}^t = 0$ otherwise
z_{ie}^t : $z_{ie}^t = 1$ if we produce i at t and no production of i at $t-1$ in node e, $z_{ie}^t = 0$ otherwise
P_{ie}^t : Production of item i in node e at time t
u_i^t : backorder of product i during period t
ω_i^t : amount of re-manufactured product i during period t
μ_i^t : inventory of re-manufactured product i at the end of period t

REFERENCES

Absi, N., Kedad-Sidhoum, S., and Dauzere-Peres, S. (2011). Uncapacitated lot-sizing problem with production time windows, early productions, backlogs and lost sales. *International Journal of Production Research*, 49(9):2551–2566.

Adulyasak, Y. (2012). *Models and solution algorithms for production routing problems*. phdthesis.

Adulyasak, Y., Cordeau, J.-F., and Jans, R. (2015). The production routing problem: A review of formulations and solution algorithms. *Computers & Operations Research*, 55:141–152.

Afentakis, P., Gavish, B., and Karmarkar, U. (1984). Computationally efficient optimal solutions to the lot-sizing problem in multistage assembly systems. *Management Science*, 30(2):222–239.

Agrali, S. (2012). A dynamic uncapacitated lot-sizing problem with co-production. *Optimization Letters*, 6(6):1051–1061.

Almeder, C. and Almada-Lobo, B. (2011). Synchronisation of scarce resources for a parallel machine lotsizing problem. *International Journal of Production Research*, 49(24):7315–7335.

Almeder, C., Klabjan, D., Traxler, R., and Almada-Lobo, B. (2015). Lead time considerations for the multi-level capacitated lot-sizing problem. *Operational Research*, 241(2):727–738.

Bahl, H. C., Ritzman, L. P., and Gupta, J. N. D. (1987). Or practice - determining lot sizes and resource requirements: A review. *Operations Research*, 35(3):329–345.

Barbarosoglu, G. and Ozgur, D. (1999). Hierarchical design of an integrated production and 2-echelon distribution system. *European Journal of Operational Research*, 118(3):464–484.

Bard, J. F. and Nananukul, N. (2009). The integrated production-inventory-distribution-routing problem. *Journal of Scheduling*, 12(3):257.

Bookbinder, J. H. and Tan, J.-Y. (1988). Strategies for the probabilistic lot-sizing problem with service-level constraints. *Management Science*, 34(9):1096–1108.

Brahimi, N., Dauzere-Peres, S., Najid, N. M., and Nordli, A. (2006). Single item lot sizing problems. *European Journal of Operational Research*, 168(1):1–16.

Buschkuhl, L., Sahling, F., Helber, S., and Tempelmeier, H. (2010). Dynamic capacitated lot-sizing problems: a classification and review of solution approaches. *OR Spectrum*, 32(2):231–261.

Chang, P. (2001). *Uncapacitated and capacitated dynamic lot size models for an integrated manufacturer-buyer production system*. phdthesis.

Darvish, M. (2017). *Supply chain optimization: location, production, inventory and distribution*. phdthesis.

Darvish, M. and Coelho, L. C. (2018). Sequential versus integrated optimization: Production, location, inventory control, and distribution. *European Journal of Operational Research*, 268(1):203–214.

Denizel, M. and Sural, H. (2006). On alternative mixed integer programming formulations and lp-based heuristics for lot-sizing with setup times. *Journal of the Operational Research Society*, 57(4):389–399.

Diaby, M. and Martel, A. (1993). Dynamic lot sizing for multi-echelon distribution systems with purchasing and transportation price discounts. *Operations Research*, 41(1):48–59.

Drexl, A. and Kimms, A. (1996). Lot sizing and scheduling: survey and extensions. *Manuskripte aus den Instituten fur Betriebswirtschaftslehre der Universitat Kiel*, 421.

Dural-Selcuk, G., Kilic, O. A., Tarim, S. A., and Rossi, R. (2016). A comparison of non-stationary stochastic lot-sizing strategies. *arXiv preprint arXiv:1607.08896*.

Eppen, G. D. and Martin, R. K. (1987). Solving multi-item capacitated lot-sizing problems using variable redefinition. *Operations Research*, 35:832–848.

Helber, S. and Sahling, F. (2010). A fix-and-optimize approach for the multi-level capacitated lot sizing problem. *International Journal of Production Economics*, 123(2):247–256.

Huang, K. and KucuKyavuz, S. (2008). On stochastic lot-sizing problems with random lead times. *Operations Research Letters*, 36(3):303–308.

Hung, Y.-F. and Chien, K.-L. (2000). A multi-class multi-level capacitated lot sizing model. *Journal of the Operational Research Society*, 51(11):1309–1318.

Jans, R. and Degraeve, Z. (2004). An industrial extension of the discrete lot-sizing and scheduling problem. *IIE transactions*, 36(1):47–58.

Jans, R. and Degraeve, Z. (2008). Modeling industrial lot sizing problems: a review. *International Journal of Production Research*, 46(6):1619–1643.

Jaruphongsa, W., Cetinkaya, S., and Lee, C.-Y. (2004). Warehouse space capacity and delivery time window considerations in dynamic lot-sizing for a simple supply chain. *International Journal of Production Economics*, 92(2):169–180.

Jayaraman, V. and Pirkul, H. (2001). Planning and coordination of production and distribution facilities for multiple commodities. *European journal of operational research*, 133(2):394–408.

Karimi, B., Ghomi, S. M. T. F., and Wilson, J. M. (2003). The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 31(5):365–378.

Karmarkar, U. S., Kekre, S., and Kekre, S. (1987). The dynamic lot-sizing problem with startup and reservation costs. *Operations Research*, 35(3):389–398.

Kian, R., Gurler, U., and Berk, E. (2014). The dynamic lot-sizing problem with convex economic production costs and setups. *International Journal of Production Economics*, 155:361–379.

Kimms, A. (1993). Multi-level, single-machine lot sizing and scheduling: With initial inventory. *Manuskripte aus den Instituten fur Betriebswirtschaftslehre der Universitat Kiel*, 329.

Kimms, A. (1996). A genetic algorithm for multi-level, multi-machine lot sizing and scheduling. *Manuskripte aus den Instituten fur Betriebswirtschaftslehre der Universitat Kiel*, 415.

Lee, C.-Y., Cetinkaya, S., and Wagelmans, A. P. M. (2001). A dynamic lot-sizing model with demand time windows. *Management Science*, 47(10):1384–1395.

Li, C.-L., Hsu, V. N., and Xiao, W.-Q. (2004). Dynamic lot sizing with batch ordering and truckload discounts. *Operations Research*, 52(4):639–654.

Li, H. and Meissner, J. (2011). Capacitated dynamic lot sizing with capacity acquisition. *International Journal of Production Research*, 49(16):4945–4963.

Maes, J. and Wassenhove, L. V. (1988). Multi-item single-level capacitated dynamic lot-sizing heuristics: A general review. *Journal of the Operational Research Society*, 39(11):991–1004.

Matta, R. E. D., Hsu, V. N., and Li, C.-L. (2015). Coordinated production and delivery for an exporter. *IIE transactions*, 47(4):373–391.

Melega, G. M., de Araujo, S. A., and Jans, R. (2018). Classification and literature review of integrated lot-sizing and cutting stock problems. *European Journal of Operational Research*.

Parveen, S. and Haque, A. F. M. A. (2007). A heuristic solution of multi-item single level capacitated dynamic lot-sizing problem. *Journal of Mechanical Engineering*, 38:1–7.

Pitakaso, R., Almeder, C., Doerner, K. F., and Hartl, R. F. (2006). Combining population-based and exact methods for multi-level capacitated lot-sizing problems. *International Journal of Production Research*, 44(22):4755–4771.

Pochet, Y. and Wolsey, L. A. (1988). Lot-size models with backlogging: Strong reformulations and cutting planes. *Mathematical Programming*, 40(1-3):317–335.

Prakaiwichien, S. and Rungreunganun, V. (2018). Solving dynamic multi-product multi-level capacitated lot-sizing problems with modified part period balancing heuristics method. *International Journal of Applied Engineering Research*, 13(6):3350–3360.

Robinson, P., Narayanan, A., and Sahin, F. (2009). Coordinated deterministic dynamic demand lot-sizing problem: A review of models and algorithms. *Omega*, 37(1):3–15.

Sahling, F., Buschkuhl, L., Tempelmeier, H., and Helber, S. (2009). Solving a multi-level capacitated lot sizing problem with multi-period setup carry-over via a fix-and-optimize heuristic. *Computers & Operations Research*, 36(9):2546–2553.

Sambasivan, M. and Yahya, S. (2005). A lagrangean-based heuristic for multi-plant, multi-item, multi-period capacitated lot-sizing problems with inter-plant transfers. *Computers & Operations Research*, 32(3):537–555.

Schmid, V., Doerner, K. F., and Laporte, G. (2013). Rich routing problems arising in supply chain management. *European Journal of Operational Research*, 224(3):435–448.

Schulz, T. (2011). A new silver-meal based heuristic for the single-item dynamic lot sizing problem with returns and remanufacturing. *International Journal of Production Research*, 49(9):2519–2533.

Seyedhosseini, S. M. and Ghoreyshi, S. M. (2014). An integrated model for production and distribution planning of perishable products with inventory and routing considerations. *Mathematical Problems in Engineering 2014*.

Sifaleras, A., Konstantaras, I., and Mladenovi?, N. (2015). Variable neighborhood search for the economic lot sizing problem with product returns and recovery. *International Journal of Production Economics*, 160:133–143.

Sox, C. R. (1997). Dynamic lot sizing with random demand and non-stationary costs. *Operations Research Letters*, 20(4):155–164.

Stadtler, H. (1997). Reformulations of the shortest route model for dynamic multi-item multi-level capacitated lotsizing. *OR Spektrum*, 19(2):87–96.

Staggemeier, A. T. and Clark, A. R. (2001). A survey of lot-sizing and scheduling models. In *23rd annual symposium of the Brazilian operational research society (SOBRAPO)*, pages 938–947.

Stevenson, W. J. (2018). *Operations management*. McGraw-Hill, 13 edition.

Tarim, S. A. and Kingsman, B. G. (2004). The stochastic dynamic production/inventory lot-sizing problem with service-level constraints. *International Journal of Production Economics*, 88(1):105–119.

Tempelmeier, H. (2013). *Stochastic lot sizing problems*, pages 313–344. Springer.

Tempelmeier, H. and Derstroff, M. (1996). A lagrangean-based heuristic for dynamic multilevel multiitem constrained lotsizing with setup times. *Management Science*, 42(5):738–757.

Tempelmeier, H. and Helber, S. (1994). A heuristic for dynamic multi-item multi-level capacitated lotsizing for general product structures. *European Journal of Operational Research*, 75(2):296–311.

Teunter, R. H., Bayindir, Z. P., and Heuve, W. V. D. (2006). Dynamic lot sizing with product returns and remanufacturing. *International Journal of Production Research*, 44(20):4377–4400.

Wagelmans, A., Hoesel, S. V., and Kolen, A. (1992). Economic lot sizing: an o (n log n) algorithm that runs in linear time in the wagner-whitin case. *Operations Research*, 40(1):S145–S156.

Wagner, H. M. and Whitin, T. M. (1958). Dynamic version of the economic lot size model. *Management science*, 5(1):89–96.

Wu, T., Xiao, F., Zhang, C., He, Y., and Liang, Z. (2018). The green capacitated multi-item lot sizing problem with parallel machines. *Computers & Operations Research*.

Zhang, Z.-H., Jiang, H., and Pan, X. (2012). A lagrangian relaxation based approach for the capacitated lot sizing problem in closed-loop supply chain. *International Journal of Production Economics*, 140(1):249–255.

DECISION SCIENCES INSTITUTE

Digital transformation in logistics and supply chain management: about barriers to organizational change and the importance of new concepts of competence development

Sandra Kaczmarek
TU Dortmund University, Fraunhofer-Institute for Materialflow and Logistics IML
Email: kaczmarek@lfo.tu-dortmund.de

Christoph Besenfelder
TU Dortmund University, Fraunhofer-Institute for Materialflow and Logistics IML
Email: besenfelder@lfo.tu-dortmund.de

Micheal Henke
TU Dortmund University, Fraunhofer-Institute for Materialflow and Logistics IML
Email: michael.henke@tu-dortmund.de

ABSTRACT

Digitalization and Industry 4.0 require a holistic transformation process with focus on human-oriented aspects. This contribution identifies the central human-related barriers of digital transformation in logistics and supply chain management, describes the methodological gap of current personnel development approaches and explains innovative game-based concepts to master these challenges.

KEYWORDS: Managing Decisions, Digital Transformation, Management 4.0, Organizational Change, Competence Management

INTRODUCTION

Digitalization plays an important role in many sections of daily life and also companies are pursuing measures to integrate technologies into business processes (Roth 2016). A transformation of business processes is an important requirement to create transparency in supply chains (Kersten et al. 2016). Especially in German-speaking countries this digitalization is referred to as “Industry 4.0”, which describes a vision of economic production through stronger inter- and intra-company networking. Generally speaking, this is understood as the fourth industrial revolution, which represents a further development of the organization and control of the value chain (Plattform Industrie 4.0 2014). Industry 4.0 describes the vertical (within an enterprise) and the horizontal linkage of these Cyber-Physical Systems (both across multiple business units and across several companies along the supply chain) to the efficient, decentralized and flexible production of products or services.” In the course of this process, companies are increasingly formulating digitalization strategies to coordinate their initiatives and projects on digital transformation (Mertens et al. 2017). “Logistics and supply chain management coordinate and organize the physical and information exchange processes in internal and external processes” (Bischoff 2015) and thus bear a major responsibility for the changes in the Industry 4.0. The necessity to focus especially on changes in logistics and supply chain management in the context of Industry 4.0 results from its high relevance as one of the most important economic sectors in Germany. (Kübler et al. 2015) For instance, in 2015, logistics service providers and logistics services in industrial and commercial enterprises in Germany generated a logistics economic volume of 253 billion euros (Schwemmer 2016). In terms of turnover, the logistics sector is thus comparable with the automotive industry, mechanical engineering or the construction industry (Kübler et al. 2015). The importance of the logistics sector is also reflected in the number of employees. In 2015 there were around 2.97 million logistics employees, more than 8% of all employees subject to social insurance contributions in Germany (Schwemmer 2016).

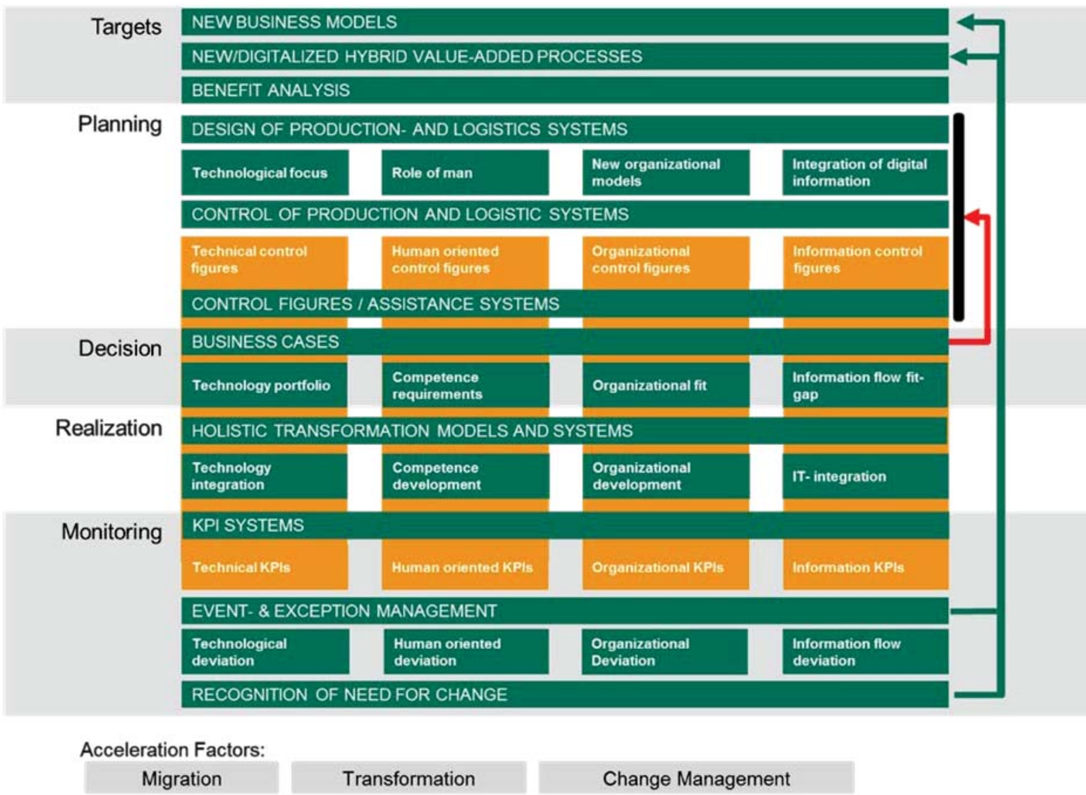
It is important to note that Industry 4.0 is not only the technological development of production but also a vision for the entire organization. The management of Industry 4.0 also has to consider the fact that in the near future new (digital) business models in logistics will emerge (Henke, Hegmanns 2017). As a result, two important key aspects of Industry 4.0 are the transformation of business process and the development of new business models according to company strategy. These developments make it necessary to rethink the entire company and its members in terms of a digital transformation. This involves a rethinking of corporate orientation in the sense of new business models and innovation processes as well as networking along the supply chain in the context of supply chain design and management and a rethinking of management as well as the strengthening and motivation of employees on their way to Industry 4.0. In logistics and SCM practice, companies face the greatest organizational challenges in implementing a digital transformation in the management of increasing complexity in collaboration and business processes as well as new forms of cooperation and networking and the necessary change in leadership culture (HR Report 2017).

This paper deals with the human-centered challenges of digital transformation in logistics and supply chain management. Based on the Dortmund Management Model, the digital transformation will be structured, changes in the working environment and the resulting competence requirements will be highlighted and a methodical gap in current qualification programs for industry 4.0 competences will be outlined. Finally, the paper discusses innovative game-based qualification approaches as a way to close this gap.

Dortmund Management Model

For this reason, the "Dortmund Management Model" (see figure 1) was developed. The Dortmund Management Model is a framework to support the management in the digital transformation of companies and is to be understood as a universal model. The challenge consists in an integrative management process that analyzes the corporate culture, the corporate organization and the value-added processes on the basis of the wealth of information and at the same time promotes process and technology innovations and has to be operationalized on a company-specific basis taking into account the individual needs of people.

Figure 1: Dortmund Management Model



In order to advance the transformation of a company to industry 4.0, the three acceleration factors change management, migration and transformation are integrated into the model as basic factors. Change management encompasses all activities aimed at bringing about far-reaching changes in the implementation of new strategies, structures, systems, processes or behaviors in a company across functions (Berger et al. 2013). Migration describes the step-by-step introduction of technologies in companies and their processes (Bildstein 2014). Transformation focuses on an agile organization, made possible by the use of suitable technologies and organizational learning to adapt to changed framework conditions (Schuh et al. 2017). Furthermore, the developed framework of the Dortmund Management Model structures the change process in two dimensions: a management dimension that describes the relevant (standard) tasks of management ('target', 'planning', 'decision', 'realization' and 'monitoring') (Ant 2018; Esser 2016), and an organizational design dimension in the form of four pillars.

The integrated view of the Dortmund Management Model is reflected in the four pillars 'Technology', 'People', 'Organization' and 'Information', which are equally included: This comprehensive organizational design approach is based on the MTO concept according to Ulich and includes the characteristics 'human', 'technology' and 'organization' (Ulich 2005). The increasing quality and quantity of digital data available in real time offers new opportunities for companies within the framework of Industry 4.0. By merging digital and real processes in the sense of digitization, companies have automatically or autonomously determined amounts of data at their disposal. By using appropriate analysis methods and artificial intelligence, these provide basic information about their own business processes, especially in the form of process, environmental and status data. Thus, on the one hand, increased transparency and valid forecasts can be achieved internally. On the other hand, data and the resulting information, for example from cross-company value creation networks, become a valuable asset for innovative data-supported business models and thus a central asset in industry 4.0. This leads to a self-optimizing organization that enables autonomous and timely adaptation to the changing conditions in the business environment (Schuh et al. 2017). Consequently, a fourth feature is added to the MTO model: Information (I) (MTO plus I).

The design requirements of the digital transformation of companies arise at the interfaces of these classic management tasks and pillars.

CASE STUDIES TO IDENTIFY THE HUMAN FACTORS IN MANAGING DIGITAL TRANSFORMATION

Based on eight concrete transfer- and implementation projects with SME, the special challenges during the transformation process were analyzed and integrated into the structure of the Dortmund management model (Henke et al. 2019). The integration of digital information and the corresponding information flows are identified as essential factors for the projects considered. For example, one metalworking company involved in the project aimed to develop an integrated information system, because the paper-based process that had previously been used increasingly led to intransparent material flows. Once these inefficient processes had been identified, an app was developed specifically for the company that could standardize, process and make available the recorded data via smart devices. In this way, the app ensures a transparent material flow that has a positive effect on delivery service levels, customer satisfaction and cost efficiency. Another company in the steel distribution manufacturing services sector saw the potential in digitizing its information flow. This measure enables the company to maintain its short order throughput time and high flexibility, as setup processes and throughput times can be better planned and integrated into production. In the context of identifying potentials with regard to new business models, the cross-company networking of data to expand the service portfolio was implemented in a pilot project for a wholesaler of technical supplies. (Henke et al. 2019)

The investigation of the cases showed that the central action dimension of most projects was to be located in the pillars of technology and information, since the objective of the digitization projects was usually a gain in information transparency through the use of new technologies in logistics. However, the challenges that arouse during the project can be classified into the pillars of people and organization.

This indicates that the successful digital transformation requires employees to be sensitized to innovations and to take them along on the path of organizational change. New technologies can only develop their full potential if the employees, as process experts and essential interfaces, are integrated into the design of the digitized processes. According to an empirical study by Ziemendorf, the causes of resistance are less to be found in rational decisions than in emotional aspects of "fear", "shyness", "helplessness", "inertia", "anger" and "power", which are increasingly attributable to insufficient participation of people in processes of change (Ziemendorf 2009).

This impairment for the transformation process exists first and foremost when resistance takes place undercover and no appropriate handling of barriers takes place because the causes are not clearly recognizable.
In this context, the emotional acceptance of transformation processes is described as the central problem. This usually leads to resistance on the part of the employees concerned during the transformation process.
The most frequently stated reasons for barriers to change relate both to the emotionality and attitude of the employees and to the type of internal project management, with the emphasis on aspects of a participative project approach and the communication structure.
In this respect, a participatory project approach is almost unanimously considered to promote change. Communication is a decisive factor in ensuring continuous transparency about concrete process changes. This primary importance of communication can also be found as a prerequisite for organisational learning. In theory, this results in optimal conditions for the implementation of change projects as well as the successful implementation of process change in organizations in a logistical environment. However, these theoretically optimal conditions are no guarantee for a smooth course of the project in operational practice. Rather, dealing with a multitude of possible resistances is part of the everyday practice of change projects. This once again shows that people are the central drivers of change.

The identified human factors in managing digital transformation include challenges such as the design of adapted organizational structures for the new digitalized business processes (Organization pillar), the active involvement of employees in the change process and the accompanying and forward-looking qualification for the new technologies (Human Resources pillar). The holistic character of the Dortmund Management Model makes clear that an integrative consideration of all pillars is necessary for a target-oriented and successful design of the change process within the scope of the digitization of companies, especially in the area of logistics.
The excerpt of the results of the case analysis emphasizes in particular the necessity to include human-oriented factors in the goal setting, planning, evaluation, realization and monitoring of digitization processes. On the one hand, it points out that an adaptive management strategy is needed, which regards the people working in the organization as the center of organizational transformation processes. This is supported with the help of the Dortmund Management Model. On the other hand, it becomes clear that the strategic qualification of employees is of particular importance and that companies must anticipate the future working requirements of their employees and proactively develop the necessary competencies.

NEW COMPETENCE REQUIREMENTS IN THE FIELD OF LOGISTICS

The present technological and organizational changes in logistics and SCM are going to transform the employees’ daily working routines. Especially in the field of operational logistics routine activities and manual work, such as the collection, documentation and inspection of incoming goods or the determination of fright weight and volume are going to decrease as part of growing options for automation. In addition, transportation and picking processes as well as packing and sorting orders can be largely automated (Kerner 2016; MHI & Deloitte 2018; Ten Hompel & Henke 2014; Ten Hompel & Kerner 2015). Furthermore, the usage of digital media, technology and mobile and assistance devices is going to increase rapidly as well as the collaborative work between humans and machines in the “Social networked industry”. (Taraj et al. 2007)

That leads to a fast development of new task fields especially in operational logistics and more and more complex processes. E.g. task extension by value creating activities (e.g. assembly activities or the operation of additive manufacturing processes) as well as by fault management related activities and maintenance. For that reasons the deployment of employees will become more flexible and new forms of work, such as collaborative group-work is going to increase, so that the daily tasks shift from manual work to more cognitive tasks (acatech 2016).
Against this background the understanding of processes and problem-solving abilities are going to become important key competences and information competence becomes a requirement for optimized decision making by means of assistance systems. (acatech 2016; Ahrens & Spöttl 2015; bayme vmb 2015; Hirsch-Kreinsen 2014; Ittermann et al. 2015) Moreover, the understanding of processes' interdependencies and the interaction between process, technology and organization will be necessary.
Hence, the development of these competences in the field of operational logistics is crucial to master the logistics system's complexity.

On a more management oriented level the field of logistics is more and more understood as an interdisciplinary field of activity that transcends organizational structures, which is currently increasingly motivating new responsibilities and thus completely new professions and designations, such as "Digital Transformation Manager", "Chief Digital Officer" or "Digital Transformation Architect".
Willingness to change, dealing with complexity, communication and, in particular, lifelong learning are the competencies in which the need for action with regard to digitalization is most significant for 99% of companies (Heise Online 2018, HR Report 2017). In particular, the necessary interdisciplinary cooperation and networking requires further training programs with new methods and approaches that enable interdisciplinary collaboration and communication as well as interpersonal transfer to be experienced and strengthened.

However, more than half of the companies in the field of logistics and SCM surveyed in the HR Report 2017 have no strategy for teaching these digital skills (HR Report 2017). At the same time, over 74% of employees want more training opportunities from their employers in relation to digital change (Heise Online 2018). According to this study, there is currently a gap between the efforts and the need for action on the one hand and the pressure on companies to act on the other. One of the main reasons given by companies for this discrepancy is that they have not found the right offer on the vocational training market.

OVERVIEW OF THE VOCATIONAL TRAINING MARKET

In the following, different types of vocational training programs on the topics of "Digital Transformation", "Digital Management" and "Management 4.0" will be outlined and subsequently examined. The scope of the analysis is initially limited to training courses offered in Germany and was then extended to international programs. For the search in Germany, "Deutsche Bildungsserver" (Bildungsserver für Weiterbildungssuche) provides a database of vocational training providers that was used to help identify suitable further training courses. The analysis of the programs offered by the leading providers of vocational training and management training in Germany revealed a total of almost 40 results for the search terms "Digital Transformation", "Digital Management" and "Management 4.0", which are summarizes in Table 1. These are mainly seminars and workshops. The measures extend over a period from one to five days. Online and certificate courses with a duration between 24 days and 18 months are also considered. The online courses are usually courses with flexible time

management. You can either register for a special course or a subscription and use the corresponding course platform. Additional learning materials are added at regular intervals so that the participant has sufficient resources at the end of the course. Some of these courses can be completed with an exam, while others already provide certificates for full participation.
In the international field, the search for vocational training programs is carried out using a structured Google search with the same search terms. In this way, additional 21 courses are identified that match the initial searches. Here, too, seminars and workshops are largely involved, although some conferences are also offered. The average duration of these measures is two days, but some are limited to two hours, while others last up to five days.

Tab. 1: Overview of the mentioned Methods

<table>
<tr><th colspan="2">Course-Format</th><th>Methods</th><th>Frequency of mentioning</th></tr>
<tr><td colspan="3">Online-Courses/MOOCS</td><td>3</td></tr>
<tr><td rowspan="10">Presence Workshops and Seminars</td><td rowspan="5">Trainer-Centre</td><td>Lectures and Presentations</td><td>12</td></tr>
<tr><td>Use-Cases and Best Practices</td><td>4</td></tr>
<tr><td>(Panel) Discussions</td><td>10</td></tr>
<tr><td>Relexion</td><td>1</td></tr>
<tr><td>Individual and group coaching</td><td>1</td></tr>
<tr><td rowspan="5">Participant-Centred</td><td>Groupwork</td><td>3</td></tr>
<tr><td>Case Study</td><td>5</td></tr>
<tr><td>Networking</td><td>2</td></tr>
<tr><td>Gamification</td><td>1</td></tr>
<tr><td>Prototyping</td><td>1</td></tr>
</table>

While the courses advertise with a variety of methods, these are largely limited to classical training formats. They make use of lectures, individual and group work, group discussions and the exchange of experience. In addition, they work with practical reports and examples and provide qualified trainers. A classical seminar structure therefore always consists of several of the elements mentioned and usually alternates between lectures and interactive elements. For example, ESMT Berlin's seminar on "Digital Transformation: Change in Leadership and Culture" addresses the need for new didactic approaches. On the basis of the available course description, however, this remains on a purely theoretical level. Most international continuing education measures are also based on discussion rounds, case studies and lectures. Debates, question rounds and interviews can also be found here.
The integration of innovative learning environments, methods and approaches, which address the market-oriented competence requirements of a management of digital transformation in logistics and SCM under the necessary didactic approach of experiential learning in an interdisciplinary transfer- and competence-oriented manner, stay unconsidered in the existing offers.

OUTLOOK ON INNOVATIVE GAME-BASED APPROACHES

This demand, however, poses new challenges for current approaches and methods for competence development and for vocational training practice (Pfeiffer et al. 2016). Especially the mechanisms of entertainment games and virtual reality have the potential to interrelate competence development as well as the new requirements of knowledge-intensive technical work and progressive technology development. (Bundesministerium 2014, De Witt 2012, Härtel 2012). Serious Games offers management a tool to continuously improve the qualification of the workforce through new challenges in a protected environment. The didactic advantage lies in the close interlocking of game and learning process, which equates the game mechanics with the learning process.

This can be achieved through a variety of in-game elements, such as point systems, rewards, level ascents and badges. Especially competence-oriented level design, which considers individual user requirements and learning paths, has great potential to cause a state of positive excitement in the learner and thus to support the self-determined confrontation with the learning object for the development of the necessary action routine and the dismantling of empirical knowledge. (Henke et al. 2017) Through the game environment, the user is also able to receive direct feedback on his decisions and actions, which in addition to feedback learning also leads to further motivation through continuous curiosity. (Besenfelder et al. 2018) In addition to the great potential for developing the metacognitive competences necessary for the change of the logistic working world and the promotion of self-motivated learning, Serious Games continue to offer the possibility of automating competence diagnostics processes in the long term due to their structure. Both, the survey of the currently available competences and the comparison with the required competences in a corresponding competence profile are usually carried out subjectively using questionnaires for self-assessment and external assessment. However, the necessary and relevant data for competence measurement and evaluation are still based on personal experience and subjective assessments and cannot be collected autonomized at this point. In particular, a competence-dependent level design in serious games, which results in correspondingly defined key performance indicators for competence assessment, enables competence assessment in real time during the learning process and thus autonomizes the cycle of operational competence management via corresponding interfaces and cross-system data exchange. As a result, the management would always have transparent information available about the currently required competencies and their internal development as decision support. On this basis, management could continue to weigh different change scenarios against each other, anticipate human-oriented risks, and proactively shape the transformation process using risk management approaches and machine learning methods. Relevant topics for further research activities would be, for example, how competence measurement and diagnostic procedures can be autonomized by Big Data Analytics approaches, which data would have to be collected for this and how and which role game based applications can play in the generation of exactly this data. Moreover, a relevant research topic could be how the effect of competence oriented personnel requirements planning can be evaluated with regard to the quality of work and which economic conclusions can be drawn for the entire organization.

CONCLUSION: LIMITATIONS AND FUTURE REASEARCH

The Dortmund management model enables an integrated view of digitization in logistics and supply chain management. As a result, it enables management to merge isolated technological solutions into an integrated management model by using the design principles of supply chain management. For industrial companies, the most important fields of action for a targeted design of the digital transformation are identified and the necessary process steps are outlined. The

model does not favour specific technologies and can be used for various industries. The challenge of digital transformation, taking into account all fields of action, in particular all action dimensions, is addressed directly by the model. In addition to practical support, the management model also enables faster scientific progress. It supports the identification of white spots in science that can be addressed with specific work, leading to a research roadmap for the scientific community in industry. This paper identifies some research requirements related to human-centred factors and, in particular, innovative competence management. New approaches to staff qualification are needed. In particular, game-based Methden can be suitable approaches to address the requirements of the new world of work in logistics and SCM. In the integration of competence-based indicators into game-based qualification measures, the process of competence management and in particular competence diagnostics could be automated in the future and support more human-oriented management decisions.
Due to rapid changes and continuous developments, the Dortmund management model is regarded as a dynamic model that must be critically scrutinized and continuously further developed.

REFERENCES

acatech (2016): Kompetenzentwicklungsstudie Industrie 4.0. Stand April 2016.

Ahrens, D.; Spöttl, G. (2015): Industrie 4.0 und Herausforderungen für die Qualifizierung von Fachkräften. In: Hirsch-Kreinsen, H.; Ittermann, P.; Niehaus, J. (Hrsg.): Digitalisierung industrieller Arbeit. Baden-Baden: edition sigma, S. 185-204.

bayme vbm (2016): Studie Industrie 4.0 – Auswirkungen auf Aus- und Weiterbildung in der M+E Industrie. bayme vbm, Stand April 2016.

Besenfelder, C.; Michalik, A.; Austerjost, M.; Kaczmarek S.: Gamification. Wegbereiter für ein effizientes Management in der Industrie 4.0. In: Jahrbuch Logistik (2018), S. 66-6 Bischoff, J. (2015): „Erschließen der Potenziale der Anwendung von ‘Industrie 4.0‘ im Mittelstand“. Hg. v. agiplan GmbH 2015.

Bildungsserver für Weiterbildungssuche: https://www.iwwb.de/weiterbildung.html?seite=12,zuletzt geprüft am 16.5.2019

BMBF (2014): Die neue Hightech-Strategie. Innovationen für Deutschland. Hg. v. Bundesministerium für Bildung und Forschung (BMBF). Referat Grundsatzfragen der Innovationspolitik. Berlin.

Bundesministerium für Wirtschaft und Energie/Bundesministerium des Innern/Bundesministerium für Verkehr und digitale Infrastruktur (Die Bundesregierung) (Hrsg.) (2014): Digitale Agenda 2014 – 2017. München. http://www.digitaleagenda.de/DA/Redaktion/DE/Publikation/digitaleagenda.pdf?__blob=publicationFile&v=8 Abrufdatum: 16.02.2015

Burke, W. W. (2008): Organization change. Theory and practice. Second edition. Los Angeles: SAGE.

Cacaci, A. (2006): Change Management – Widerstände gegen Wandel. Plädoyer für ein System der Prävention. In: Wüthrich, H. (Hrsg.): Internationalisierung und Management 2006.

de Witt, C. (2012): Neue Lernformen für die berufliche Bildung: Mobile Learning – Social Learning – Game Based Learning. In: Berufsbildung in Wissenschaft und Praxis (BWP) (Hrsg.):

Lernen mit digitalen Medien. Medienkompetenz als neue Kulturtechnik: 41. Jahrgang, Heft 3/2012, S. 6-9. http://www.bibb.de/veroeffentlichungen/de/bwp/show/id/6884 Abrufdatum:16.02.2015

Härtel, M. (2012): Wie viel Didaktik braucht das Lernen in der Berufsbildung mit Web 2.0? Lernprozesse in komplexen und veränderlichen Arbeitsumgebungen gestalten. In: Personalführung 11/2012. Hamburg, S. 22–30.

H. Hays AG, Institut für Beschäftigung und Employability IBE (2017): HR-REPORT 2017. SCHWERPUNKT KOMPETENZEN FÜR EINE DIGITALE WELT. Eine empirische Studie des Instituts für Beschäftigung und Employability IBE im Auftrag von Hays für Deutschland, Österreich und die Schweiz, 2017.

Heise Online (2018): Studie: Unternehmen bieten vermehrt Weiterbildung zur Digitalisierung. Online verfügbar unter https://www.heise.de/newsticker/meldung/Studie-Unternehmen-bieten-vermehrt-Weiterbildung-zur-Digitalisierung-4225076.html, zuletzt geprüft am 03.05.2018.

Henke, M. (2016): Management der Industrie 4.0. In: Hanne Wolf-Kluthausen (Hg.): Jahrbuch Logistik 2016. Korschenbroich: free beratung GmbH, S. 124–125.

Henke, M.; Besenfelder, C.; Kaczmarek, S.; Hetterscheid, E.; Schlüter, Florian (2018): Dortmund Management Model – a Contribution to Digitalization in Logistics and Supply Chain Management. In: Furmans, K.; Wimmer, T. (Hg): Unterstanding future logistics - models, applications, insights. 9th International Scientific Symposium on Logistics. Bremen, BVL International, Bundesvereinigungs Logistiv (BVL) e.V., S. 113-124.

Henke, M.; Hegmanns, T. (2017): Geschäftsmodelle für die Logistik 4.0: Herausforderungen und Handlungsfelder einer grundlegenden Transformation. In: Birgit Vogel-Heuser, Thomas Bauernhansl und Michael ten Hompel (Hg.): Handbuch Industrie 4.0 Bd.3. Berlin, Heidelberg: Springer, S. 335–345.

Henke, M.; Parlings, M.; Besenfelder, C.; Stute, M.; Brandl, T. (2019): Management der Industrie 4.0. In: VWI Fokusthema ¬ Band 1. Verband Deutscher Wirtschaftsingeniuere e.V. (VWI). Manuscript submitted for publication.

Henke, M.; Straub, N.; Kaczmarek S.: Spielbasierte Ansätze. Die digitale Arbeitswelt in der Logistik 4.0. Logistik für Unternehmen. (2017)

Hirsch-Kreinsen, H. (2014): Wandel von Produktionsarbeit - Industrie 4.0. In: WSI-Mitteilungen, 67 (2014) 6, S. 421–429.

Huy, Q. N.; Mintzberg, H. (2003): The rhythm of change. In: MIT Sloan Management Review 44 (4), S. 79-44.

Ittermann, P.; Neuhaus, J.; Hirsch-Kreisen, H. (2015): Arbeiten in der Industrie 4.0. Trendbestimmungen und arbeitspolitische Handlungsfelder. Hrsg. Hans Böckler Stiftung.

Kerner, S. (2016): Pack mit an, Roboter. Autonome Systeme in der Logistik. In: C’T Magazin für Computertechnik 13 (2016), S. 74 76.

Kersten, W.; Seiter, M.; See, B. von; Hackius, N.; Rosentritt, C.; Böhle, C. et al. (2016): Trends und Strategien in Supply Chain Management und Logistik – Chancen der digitalen Transformation. In: Thomas Wimmer und Christian Grotemeier (Hg.): Den Wandel gestalten. 33. Deutscher Logistik-Kongress: Tagungsband Driving change. Hamburg: DVV Media Group GmbH, S. 344–361.

Kostka, C.; Mönch, A. (2009): Change Management. 4. Auflage München 2009.

Kübler, A.; Distel, S.; Veres-Homm, U. (2015): Logistikbeschäftigung in Deutschland. Vermessung, Bedeutung und Struktur, Stuttgart: Fraunhofer Verlag.

Mertens, P.; Bodendorf, F.; König, W.; Schumann, M.; Hess, T.; Buxmann, P. (2017): Grundzüge der Wirtschaftsinformatik. Berlin, Heidelberg: Springer.

MHI; Deloitte (2018.) The 2018 Annual Industry Report: overcoming barriers to nextgen supply chain innovation.

Nonaka, I. (1994): A dynamics theory of organizational knowledge creation. In: Organization Science 5 (1). S. 14-37.

Pfeiffer, S.; Lee, H.; Zirnig, C.; Suphan, A. (2016): Industrie 4.0- Qualifizierung 2025., VDMA (Hrsg.).https://arbeitsmarkt.vdma.org/documents/7974667/7986911/VDMA-Studie%20Qualifizierung%202025/f88fce03-d94e-46cb-a60f-54329236b2b7. Abrufdatum: 10.01.2017.

Plattform Industrie 4.0 (2014): Industrie 4.0. Whitepaper FuE-Themen. Hg. v. Plattform Industrie 4.0.

Roth, A. (2016): Industrie 4.0 – Hype oder Revolution? In: Armin Roth (Hg.): Einführung und Umsetzung von Industrie 4.0. Grundlagen, Vorgehensmodell und Use Cases aus der Praxis. Berlin, Heidelberg: Springer-Verlag, S. 1–15.

Schuh, G.; Kampker, A.; Huesmann, R. (2011): Unternehmensentwicklung. In: Günther Schuh und Achim Kampker (Hg.): Strategie und Management produziernder Unternehmen. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 231–325.

Schwemmer, M. (2016): TOP 100 der Logistik. Marktgrößen, Marktsegmente und Marktführer. Eine Studie der Fraunhofer Arbeitsgruppe für Supply Chain Services, Hamburg: DVV Media Group.

Studt, F. J. (2008): Nachhaltigkeit in der Post Merger Integration. Wiesbaden 2008.

Taraj, S.; Johannesson, M.; Backlund, P. (2007): Serious Games – An Overview. Technical Report HS- IKI -TR-07-00, 2007.

Ten Hompel, M.; Henke, M. (2014): Logistik 4.0. In: T. Bauernhansl, M. ten Hompel, B. Vogel-Heuser (Hrsg.) Industrie 4.0 in Produktion, Automatisierung und Logistik-Anwendung, Technologien, Migration. Wiesbaden 2014, S. 615-624.

Ten Hompel, M.; Kerner, S. (2015): Logistik 4.0. Die Vision vom Internet der autonomen Dinge. In: Informatik Spektrum 38 (2015) 3, S. 176 182.

Wilkesmann, U. (2004): Lernende Organisation, Wissensmanagement und Lernkulturentwicklung – schöne Worte oder mehr?. Überlegungen aus organisationstheoretischerSicht. In: Zeitschrift für Berufs- und Wirtschaftspädagogik ZBW 100 (2004) 3, S. 383-397.

Wilkesmann, U./Georges, A./Romme, L. (2003): Organisationales Lernen, zirkuläres Lernen und die Veränderung der intraorganisatorischen Herrschaftsverhältnisse. In: ARBEIT: Zeitschrift für Arbeitsforschung, Arbeitsgestaltung und Arbeitspolitik 3 (2003), S. 228-241.

Ziemendorf, B. (2009): Emotionale Akzeptanz in Veränderungsprozessen. Entwicklung eines didaktischen Konzeptes. In: Schweizer, K./Korte, P.: Erwachsenenbildung und Weiterbildung. Frankfurt am Main 2009.

DECISION SCIENCES INSTITUTE

Holistic Transformation Model for Additive Manufacturing in existing corporate structures

Stephanie Niehues
Technical University of Dortmund
Email: Stephanie.Niehues@tu-dortmund.de

Ferdinand Deitermann
Technical University of Dortmund
Email: Ferdinand.Deitermann@tu-dortmund.de

ABSTRACT

The central objective of this conceptual paper is the development of a transformation model for the implementation of additive manufacturing in existing corporate structures. Many companies see great advantages in the use of additive manufacturing, but cannot, or only with great effort, integrate the technology into their processes due to the associated complex decision-making structures. The work is based on the leading research question of how to design an appropriate process model taking into account the potentials and challenges of additive manufacturing, which can be seen as a subcategory of advanced manufacturing technologies.

KEYWORDS: Additive Manufacturing, Transformation Model, Advanced Manufacturing Technology (AMT), Production Technology, AMT Implementation

INTRODUCTION

The term additive manufacturing bundles a group of manufacturing technologies which, in contrast to subtractive methods, apply the material layer by layer. This type of manufacturing technology makes it possible to manufacture complex geometries and components with a negligible share of set-up costs, since tool changes are no longer necessary during the manufacturing process. (Möhrle and Emmelmann 2016)

By combining additive manufacturing with conventional manufacturing processes, various advantages can be achieved, especially for manufacturing company structures. For example, bionic designs based on natural models are easy to produce, component adaptations or improvements can be implemented quickly by adapting the data, and spare parts can be manufactured locally at the point of need. As a result, warehousing and capital commitment costs are reduced, as not all spare parts have to be kept in an indefinite quantity. (Zeyn 2017)

Many companies are aware of the advantages of additive manufacturing, especially in the area of customer-specific manufacturing of products and the flexible adaptability of the technology. Nevertheless, many companies have a certain inhibition threshold, since the location and above all the integration of additive manufacturing processes are associated with a high planning effort. The uncertainty as to whether the technology can really be used economically also prevents many companies from integrating it into their own production systems.

The manufacturing technology of Additive Manufacturing (AM) is allocated in this paper as an Advanced Manufacturing Technology.
Advanced Manufacturing Technologies (AMT) in particular are increasingly being used as strategic weapons in global competition (Singh et al. 2007). An important competitive advantage can be achieved by integrating additive manufacturing (Besenfelder et al. 2018) which is a subcategory of AMTs (Chen et al. 2017).
The implementation of AMTs is generally regarded as one of the longest, most expensive and most complex projects a company can undertake (Gunawardana 2006). Moreover, such projects to create efficient production structures are not supported by systematic measures, so that it is particularly necessary to introduce efficient management for process changes (Lerch et al. 2012).

This paper aims to answer the following research questions:

(1) What requirements does AM as an AMT demand from a structured integration into existing corporate structures?
(2) Which AM and AMT transformation models are already available and to what extent do they need to be adapted on the basis of the specific requirements?
(3) How can a holistic transformation model for additive manufacturing be described and structured?
In order to answer these questions, the integration requirements of AMTs will first be determined by a literature search, as well as in a next step the already existing models for the technology integration of AMT will be identified. In the next section, the specific requirements of additive manufacturing, especially in the area of direct manufacturing, will be examined and the most relevant models for the integration of AM will be shown. The final result of the paper is the development and explanation of a holistic transformation model for additive manufacturing processes based on literature research. This model will be finally summarized and future needs for further development and research will be identified.

LITERATURE REVIEW

Additive Manufacturing - an Advanced Manufacturing Technology

Additive manufacturing (AM) is a category of manufacturing technologies besides subtractive manufacturing and formative manufacturing. In contrast to the latter, AM technologies shape the workpiece by adding portions of material. If these portions of material are added in the form of layers, the term Additive Layer Manufacturing can also be used. Most AM processes are automated and feature layer technology, indicating that a workpiece is manufactured layer by layer. Therefore, the geometry of the workpiece is sliced virtually into layers and then built physically applying the virtual layer information. (Gebhardt and Hötter 2016)

The additive manufacturing process can be divided into the three applications Rapid Prototyping, Rapid Tooling and Direct Manufacturing (application level). Rapid prototyping is used to create models and prototypes to quickly implement first test applications and to make designs haptically perceptible. Direct Manufacturing (DM) describes the process by which saleable products are manufactured using AM. Tools and tool inserts are produced within rapid tooling. Rapid Repair is a fourth possible application that complements the process and describes the maintenance and repair of worn products. (Lachmayer et al. 2016)
If these applications of AM are located in the product development process, it becomes clear that rapid prototyping is used for product development phases. This is followed by rapid tooling

in the product development process, with which direct and indirect tools are manufactured. The DM is used for the production of small series or series. (VDI 2013) In principle, the term "series" must be viewed in a differentiated way in the AF sector, as the number of components can vary considerably between industries (Klemp and Pottebaum 2016). The Rapid Repair application can be assigned to the Service area after the actual product creation.

This results in a multitude of application possibilities for the technology and poses a challenge for many companies of identifying economically viable business models. A further challenge is the multitude of technical procedures, which are summarized under the heading of additive manufacturing. Many companies lack the appropriate know-how to select the right technology for the application.

In this section, DM is described in more detail, since the present paper focuses on the additive manufacturing of end products. Studies predict that DM, between 2030 and 2050, will even become the main application of AM (Fischer and Rommel 2018). DM originated from rapid prototyping and, according to Gebhardt and Hötter (2016), still uses the same machines (Gebhardt and Hötter 2016). Due to significant progress in quality, reliability and material selection as well as the increasing digitalisation of companies, DM is becoming increasingly relevant (Fischer and Rommel 2018, Grund 2015, Lutter-Günther et al. 2015b). DM is characterised by the fact that it is possible to produce individual pieces, small series and theoretically also mass goods (Grund 2015). The unit costs remain almost constant due to the tool-free and wear-free production with minimum material input, independent of the batch size. This type of production is therefore particularly interesting for small quantities and individual pieces (Grund 2015, Hague 2006, Hokinson and Dickens 2003, Ruffo and Hague 2007, Schmutzler et al. 2016)

DM as a specific application of the additive manufacturing process can be counted in the category of Advanced Manufacturing Technologies (AMT). (Chen et al. 2017)

AMT can be defined as a modern manufacturing technology and improvement in conventional manufacturing which represent excellent quality, cost efficiency and high productivity with a low environmental footprint (Gupta 2017). Further potential for AMTs comes from faster market access and faster response times to fluctuating customer needs (Small and Yasin 1997). In addition to the manufacturing operations, the overall competitiveness of the company is improved by AMTs (Pagell et al. 2000, Small and Yasin 1997).

Requirements of AMT Integration Projects

For the successful implementation of AMTs and the associated realization of potentials, the consideration of critical factors is elementary (Rahardjo and bin Yahya 2010, Singh et al. 2007) In the previous section it also became clear that differentiation into different phases during implementation is useful. Therefore, the study by Rahardjo & bin Yaha, which refers to individual phases of the implementation, is particularly relevant. The authors identify 15 decisive success factors and determine the five most important factors for each of four different implementation phases (cf.

Figure 1). Among other things, they orient themselves on the model presented by Frohlich (Frohlich 1998).

Figure 1: The five most important success factors of the respective implementation phase adapted from (Rahardjo and bin Yahya 2010)

Rank	Pre-Installation	Installation	Growth	Fully Developed
1	- Good Leadership	- Good Leadership	- Availability of Capital	- Support & Integration of Top Management
2	- Support & Integration of top management	- Availability of Capital	- Teamwork	- Cross-Departmental Coordination & Cooperation
3	- Availability of Capital	- Teamwork	- Availability of Expert Knowledge	- Enhancing Communications Technology
4	- Company Culture	- Support & Integration of top Management	- Control over the AMT process	- Collection and analysis of AMT data
5	- Training & Education of AMT	- Cross-departmental Coordination & Cooperation	- Good Leadership	- Communiction with Supplier

In the pre-installation phase and the installation phase, good management is the most important factor. In the growth phase, however, the availability of capital is the most important. In the final phase, the maturity phase, the support and involvement of top management is essential. (Rahardjo and bin Yahya 2010)

Further factors that are important for the implementation of AMTs are presented by the authors Saberi et al. and Darbanhosseiniamirkhiz & Wan Ismail. The two studies are very similar in their results.

In the area of technological factors, both studies describe that justification before installation is important (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010) as many AMTs are associated with high costs for hardware and software, as well as complex organizational and operational problems. For a comprehensive view, which makes it possible to implement as many advantages of AMTs as possible, it is necessary to use an integrated approach that uses economic, strategic and analytical methods in parallel.

The degree of integration of AMTs into the company should also be considered, as a higher degree of integration requires more intensive planning. In the area of technological factors, the integration of a technology also offers more advantages than the automation of individual processes (Saberi et al. 2010) and the knowledge of potential advantages of AMT can positively influence the adoption of these (Darbanhosseiniamirkhiz and Wan Ismail 2012).

A key factor for a successful implementation is the structure of the organization. The factors of the two studies are very similar, especially in this area. For the corporate structure, as a factor, less hierarchical, bureaucratic structures are better suited (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010).

In addition, employees should be prepared for new structures even before the implementation process, as in many industrial companies it takes several years until the structure and the technology match (Darbanhosseiniamirkhiz and Wan Ismail 2012). Furthermore, the corporate culture must also be taken into account when implementing AMTs. In principle, a flexible corporate culture is an advantage for the implementation of an AMT (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010).

The third factor is the operational strategy or the manufacturing strategy. Within the framework of the manufacturing strategy, managers should see investments as an opportunity to gain lasting competitive advantages. A corresponding manufacturing strategy consists of the four

most important competitive factors: costs, quality, flexibility and reliability/delivery reliability. (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010)
For a successful implementation of AMTs in the future, appropriate staff training is required (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010).
At the same time, management practices must be adapted so that employees are intrinsically motivated and satisfied with their work. Successful implementation projects also benefit from "champions" who, as individuals, drive implementation efforts (Saberi et al. 2010). Top management plays a crucial role in motivating employees for the new technology and the associated challenges (Darbanhosseiniamirkhiz and Wan Ismail 2012).

Integration of Advanced Manufacturing Technologies
In principle, AMTs are seen as the key to improvement in many industries (Frohlich 1998). However, these expectations cannot be met in many implementation projects (Chen et al. 2017, Rahardjo and bin Yahya 2010). In U.S. companies, for example, implementation fails in 50-75% of cases (Chew et al. 1991).
A large part of the existing studies shows that above all, the organization of a company is relevant for the full potential exploitation of a technology and not exclusively the technological aspect (Friedli 2006, Rahardjo and bin Yahya 2010, Voss 1988). Traditional roles and responsibilities, work content, remuneration systems and teaching approaches in companies should be revised accordingly (Rahardjo and bin Yahya 2010).
On the one hand, the theory for the successful implementation of AMTs is therefore relevant in order to create a well thought-out implementation concept. On the other hand, the findings of AMT research can be used as a basis to introduce new industry 4.0 technologies, including AM (Maghazei and Netland 2017).

For the implementation of AMTs, some far-reaching procedural models already exist, the most important of which will be briefly summarized in the table below.

Table 1 Implementation Models for AMT

	Model Phases	Main emphasis
Voss 1988	Pre-Installation Installation and Commissioning After-Commissioning	Taking up special factors that influence the success and failure of the implementation.
Chen & Small 1994	Planning Phase Installation and evaluation phase	The focus of their work is on the pre-installation phase or the planning phase, as this phase is important for a successful implementation.
Small & Yasin 1997	Business Environment Corporate Response Company Goals Performance Indicators Technology Types Infrastructural Changes Legitimation Installation of Systems	Holistic framework, which also includes the dependencies between the individual elements or activities during implementation.

	Measurement of Performance	
Frohlich 1998	Unfreezing Changing Refreezing	Depending on the state of development of the technology (Degree of maturity), different actions are taken for implementation.

It should be emphasized that the Voss model was used as a basis for the Chen & Small and Small & Yasin models, but it was optimized by the continuation.

Requirements of DM Integration Projects

DM not only influences production but also the entire business logic and value chain of the company (Muita et al. 2015, Steinwender 2016). AM is therefore not a pure substitution technology for conventional manufacturing processes (Steinwender 2016). In order to utilize its full potential, various corporate divisions must be considered and their activities coordinated (Lutter-Günther et al. 2015b). For this purpose, a suitable business model should be chosen which is also the prerequisite for the successful implementation of additive manufacturing processes (Lutter-Günther et al. 2015a, 2015b).

Fischer et al. have dealt with the decision making for DM. DM is generally the most complex and demanding application (Fischer and Rommel 2018).

Due to the high complexity of the DM, it must be considered more precisely whether the advantages of the process can be exploited. For the corresponding determination, the authors specify influencing factors (product, material, human & environment, machine and method) which decide on success or failure. (Fischer and Rommel 2018)

For the product, at least one of the features should apply, such as small quantities, high degree of individualization desired, production on demand, high complexity of the component, and integration of functions or lightweight construction. With regard to the material, the following question must be clarified: "Can the material of which my product is made of be printed or can I replace the material sensibly and with added value with a printable one?. Well-trained employees who can handle a high degree of digitisation are necessary for the human and environmental factor. In addition, the introduction of AM is causing major restructuring, so that good change management up to the management level is necessary. (Fischer and Rommel 2018)

The following must be clarified: "Do I have the necessary personnel resources and the quality of employees or can I acquire them? Is my change management good enough for changes of this magnitude? (Fischer and Rommel 2018)

For the last success factor, the method, it is evaluated whether the surface quality and the quality of the mechanical properties of additively manufactured components are sufficient. Basically, the influencing factors must be considered before making a decision. (Fischer and Rommel 2018)

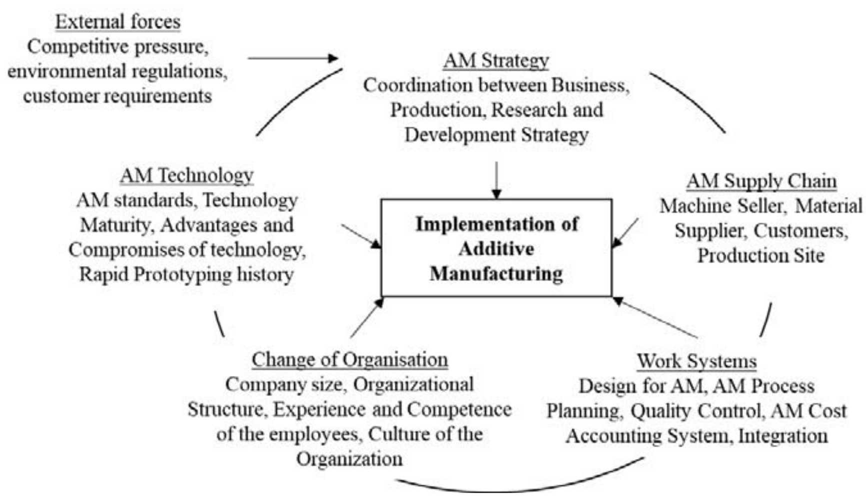
Integration of AM

This section presents research that has developed an approach for implementing AM in production structures, especially the DM. The most important models in this area are listed and illustrated below.

As part of Mellor's research, a framework for the implementation of additive manufacturing in production applications is being developed (cf.

Figure 2). The research is based on the theory of implementing new technologies or AMTs and builds on the findings of Voss. The presented framework considers external factors and the internal strategy as driver of AM as a production method. In the area of external factors, competitive pressure, environmental regulations and customer requirements must be considered. The approach to implementing AM is influenced by five overall constructs. (Mellor et al. 2014)

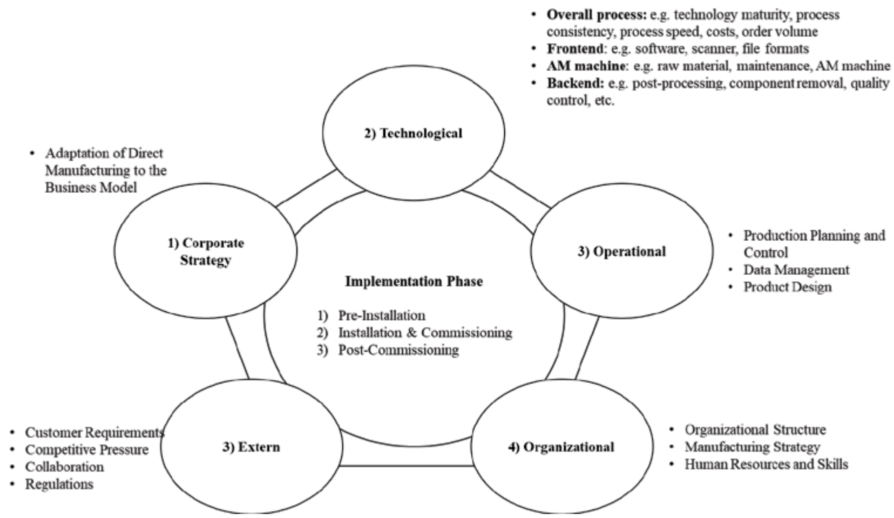
Figure 2: The proposed framework of AM implementation adapted from (Mellor et al. 2014)



As Deradjat & Minshall have already stated in their research, Mellor's model has decisive gaps, which can lead to the failure of the project in the event of implementation. Important technical factors such as necessary software components and necessary post-processing processes are not considered in the model. In addition, the changes to the framework parts during the implementation process are not further taken into account. The three steps along the implementation process (pre-installation, installation and commissioning, post-commissioning), which was already developed at Voss, are not considered in the existing framework. (Deradjat and Minshall 2017)

For this reason, Deradjat & Minshall have developed their own process model (cf.*Figure 3*), which takes the above mentioned aspects into account, but specializes in the field of mass customization.

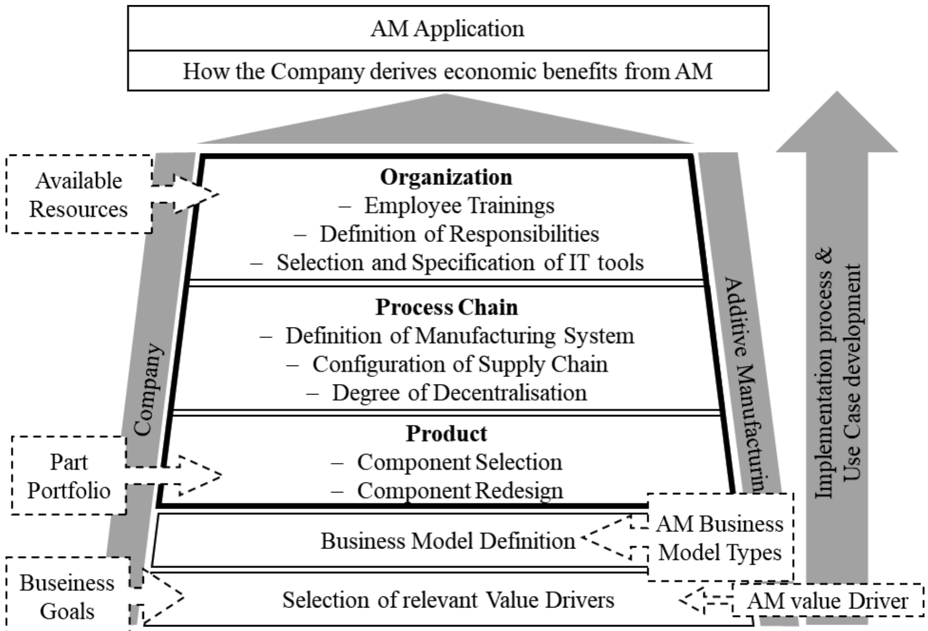
Figure 3: Framework of Rapid Manufacturing implementation for Mass Customization adapted from (Deradjat and Minshall 2017)



Their framework deals with technological, operational, organizational and internal and external factors with a focus on technological aspects. The implementation phase is based on the structure of the Voss model. (Deradjat and Minshall 2017)
The model focuses on the introduction of mass customization and is strongly tailored to the dental sector, and too little attention is paid to the necessary organizational changes. Only in the fourth step the employees and the necessary teams for implementation are taken into account.

The process model of Lutter-Günther et al. aims to provide a systematic support tool for the development of a company-specific AM application case (Lutter-Günther et al. 2015a).
The research is explained in more detail here, as the authors develop a process model that applies to all AF technologies. In addition, the process model has a special focus on business models that are important for the implementation of additive manufacturing processes.
The process model consists of three fields of action: product, process chain and organization (cf. Figure *4*). In addition, the typology of the AM business models is included in the model. The phases of the model are run through one after another. In the first step, the conceptual business model is defined, which in some cases can be a combination of different business models. Value drivers are selected on the basis of the business objectives and challenges using a pair comparison or a weighted multi-criteria evaluation. To gain knowledge about operational and strategic business goals and challenges, workshops can be conducted with AM experts and company representatives. (Lutter-Günther et al. 2015a)

Figure 4: Process model for AM implementation adapted from (Lutter-Günther et al. 2015a)

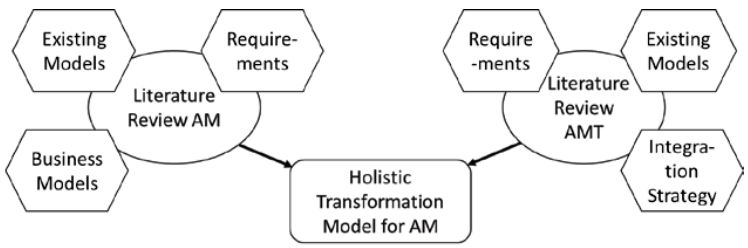


In the model of Lutter-Günther et al., the representation and description of the procedure in the text is partially inconsistent. However, the model is generally easy to understand and due to fewer elements, the model has a simple structure and, according to the author, is universally applicable. However, the procedure model does not have clear connections between individual elements and steps, since, for example, dependencies between individual elements are not described. Furthermore, the given process model has a roadmap character, but does not provide an optimal sequence for the introduction of AM. In addition, the process model answers only conditionally how the implementation takes place in concrete terms. It is not possible to make a statement on the practicability, since this has not yet been demonstrated. There is no concrete specification of the sequence of work steps and definition of responsibilities. Furthermore, there is no success control between the steps and the authors do not mention any feedback.

METHODOLOGICAL APPROACH

The previous chapters clearly show that none of the approaches presented for the development of a procedure model for the integration of additive manufacturing, especially DM, can be adopted. It is rather necessary to develop a procedure on the basis of the theory for the implementation of AMTs and to adapt this for the concrete case of the use of additive manufacturing processes. For the development of a corresponding procedure model it is necessary to make the resulting complexity controllable. The development of the process model therefore takes place step by step as shown in Figure 5. For this purpose, the facts are considered step by step "from rough to detailed".

Figure 5: Methodical approach for the procedure model

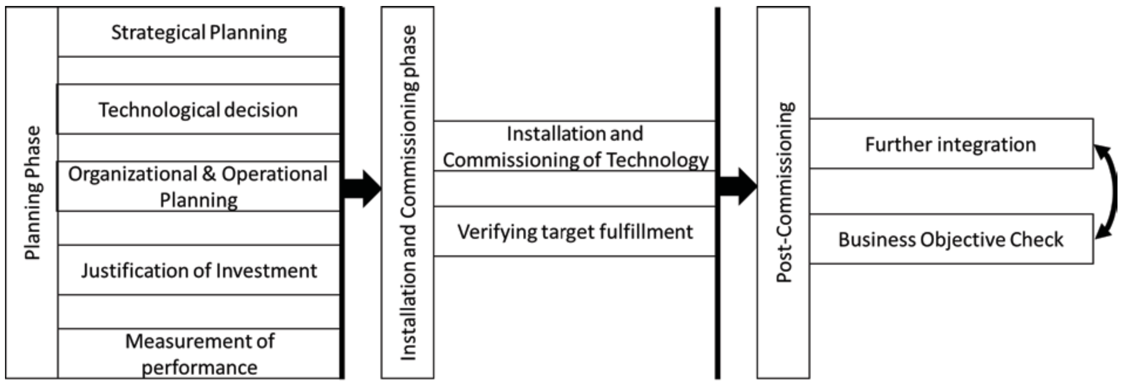


In order to develop the process model, a literature review on AM was first conducted. What are specific requirements during the introduction of the technology and which models were set up in this area? In addition to the technical components of the technology, possible business models were also identified in order to meet the aforementioned important step of the initial strategy.
In addition, a literature review was also conducted in the field of AMTs in order to derive the most important requirements for technology integration, but also to record existing integration models and their underlying integration strategy. Moreover, it became clear that an unstructured introduction of AMTs would lead to considerable blind processes and that therefore a guided process model is necessary for the introduction of AM.
In the next step, the developed process model is explained more closely.

HOLISTIC TRANSFORMATION MODEL FOR ADDITITVE MANUFACTURING

First of all, it must be clarified which structural requirements can be adapted to the model from the requirements of AMT and DM and the previous models introduced for AMTs and DM.
The process model is divided into a planning phase, an installation and commissioning phase and a post-commissioning phase. This structure has already been validated by Voss as suitable for AMTs, and is therefore valid for DM.

Figure 6: Holistic Transformation Model for AM



The planning phase is the most comprehensive phase and includes the determination of the company's objectives, technology monitoring, the planning of infrastructure adjustments, the justification of the investment, as well as the measurement of the current company performance. The installation and commissioning phase deals with the concrete installation of AM and the subsequent technical and operational fulfilment of the objectives. In the final phase, the post-commissioning phase, AM must be fully integrated and it must be checked whether the desired business benefit has been achieved. Each phase will be explained in more detail further on.

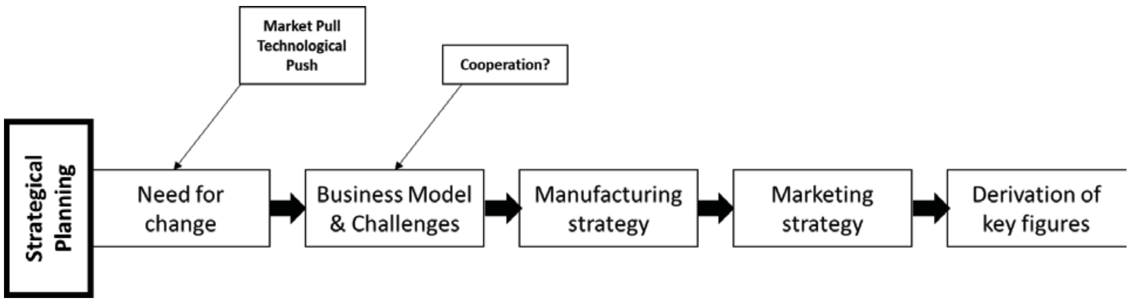
According to Voss and Chen & Small, a technical success measurement must be carried out after installation and commissioning. In addition, both studies provide for validation of the expected business benefits in the post-commissioning phase. Furthermore, the phases in the Voss model are only considered completed after successful testing. Reviews must therefore be integrated into the process model to be developed. A further feature of the Chen & Small and Small & Yasin models are dependencies between individual process steps. Due to these properties, the waterfall model with recesses will be used as the basic structure in the following. The model is characterized by a clear structure and subsequent steps are only carried out after successful verification. In addition, dependencies or feedback can be expressed by the regressions.

Planning Phase

<u>Strategical Planning</u>

The first step of the process begins with the company realizing that the current manufacturing process is no longer sufficient. On the one hand, this insight can result from a market pull in which the needs of the customer or the market have changed (Feldmann et al. 2019, Schallmo 2018). On the other hand, the investment in DM can arise as a technology push and thus for the proactive development of production capacities, as described in Mellor's approach.

Figure 7: Planning Phase

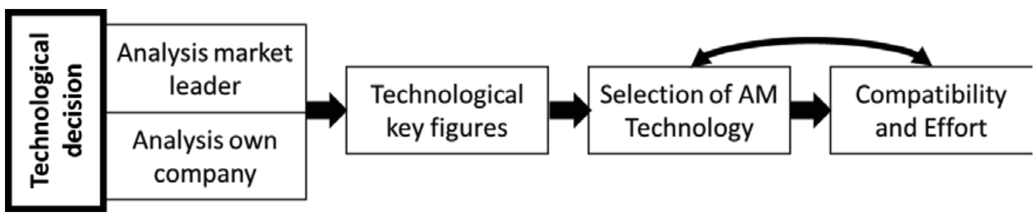


In the second step, the business objectives and challenges are identified.
Within the literature research, it becomes clear that business model development must be included in the framework. On the one hand, it becomes clear that a business model must be selected for the implementation of additive manufacturing processes for the use as DM. On the other hand, the approach of Lutter-Günther et al. also points to this activity as the first step to be carried out (Lutter-Günther et al. 2015a). The development of a business model basically offers the company the possibility to operationalize strategic goals.
Within the development of the model at least one of the following points should apply to the product: Small quantities, high degree of individualization desired, production on demand, high complexity of the component, integration of functions, lightweight construction (Fischer and Rommel 2018).
In principle, the strategic considerations must also determine whether collaborations are to be entered into and if challenges can be overwhelmed by these collaborations. In turn, the decision on this depends on the degree of technological maturity. First movers and early adopters, for example, have high initial costs, which are supported by collaborations. On the other hand, there is the protection of intellectual property. The consideration of partners for a collaboration is basically also a component of the development of the business model (Fueglistaller et al. 2016).
A further partial step of the strategic considerations is the planning of a manufacturing strategy. This strategy is divided into three elements: The manufacturing task, the production structure and the production process. For the manufacturing task, the type and quantity of the products to be produced, the technical specifications and market requirements are defined (Zahn 1988). When selecting a suitable manufacturing strategy for AM, the product is particularly the focus of consideration (Fischer and Rommel 2018). For the second element, the production structure, a differentiation must be made between the basic structure and the infrastructure. For the basic structure, capacity changes, locations and the design of the locations with production facilities are determined. The infrastructure in turn refers to information and communication systems, planning and control systems, organizational processes and personnel skills. The production process determines the vertical integration and handling of logistical processes (Zahn 1988). In principle, the four competitive factors cost, quality, flexibility and reliability/delivery reliability must be considered for a corresponding manufacturing strategy (Darbanhosseiniamirkhiz and Wan Ismail 2012, Saberi et al. 2010).
Subsequently, it is necessary to link the marketing strategy with the manufacturing strategy (Mellor et al. 2014). Focusing on additive manufacturing of products, it may be necessary to convince customers of the quality of the manufactured components (Deradjat and Minshall 2017).

A sub-step deals with the development of measurable variables or key figures, this is important to be able to determine the success of the implementation afterwards. Possible key figures can be found in the Framework of Small & Yasin (Small and Yasin 1997).
The performance review at the end of phase is particularly relevant. On the one hand, it is important to evaluate whether the DM has advantages over other manufacturing technologies. On the other hand, it must be checked whether the strategies are in line with the business objectives. If both questions are answered in the affirmative, the process model will continue to be run through. Otherwise, the implementation case under consideration is rejected.

Technological decision
Once the strategic considerations have been completed, an AM technology will be selected (cf. Figure 8).

Figure 8: Technological decision



First, the technological, organizational and operational profile of the market leaders and their own company are analyzed. The findings from the profile analyses are then brought together and, in the case of the proactive approach, supplemented by profile analyses in related sectors.

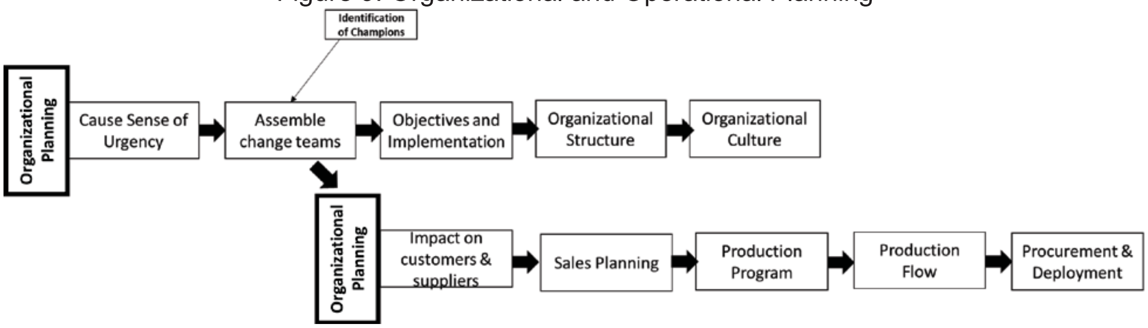
The findings can be used to evaluate the situations in which certain AM technologies can be used in the company itself. The next sub-step is the development of technical variables for the use of AM in the developed business case. The following variables can be used to select an AM technology: Component volume, material, component geometry, necessary quality and accuracy, production speed, costs, maintenance effort, production volume, the rate of introduction of new products, the software and possible software restrictions. (Besenfelder et al. 2019, Deradjat and Minshall 2017, Lutter-Günther et al. 2015b, Mellor et al. 2014)

In the last step, the AM technology is checked for compatibility with the existing production processes. In the special case of AM, the process chain is checked and evaluated to determine which processes are relevant before and after actual construction. This makes it possible to analyze which processes and competencies are available in the company or must be developed (Mellor et al. 2014). If processes and competencies have to be developed, the corresponding effort has to be determined. If the AM technology is not sufficiently compatible with the existing production processes, an accordingly higher expenditure for the integration must be considered (Deradjat and Minshall 2017).
As a result, more resources should be planned or collaborations should be considered. Alternatively, it is possible to select a different technology, as shown by the feedback loop in the figure. At this point, the analysis is less concrete, as concrete organizational and operational planning takes place in the following step. To successfully complete this step, an AM technology must be selected.

Organizational and Operational Planning

Organizational planning and operational planning are fundamentally highly interdependent (Sun and Riis 1994). In order to determine a structured procedure for the respective planning, however, the information in the studies on the implementation of AMTs and DMs is deficient. For this reason, research on operational planning as part of corporate planning, which serves to implement a previously defined strategy is used. (Welge and Al-Laham 1992)

Figure 9: Organizational and Operational Planning



First, organizational planning is described in more detail below. This begins before operative planning (cf.

Figure 9), since teams for cross-departmental planning, which are necessary for operative planning, are first determined in organizational planning.

The first sub-step in organizational planning is the creation of a sense of urgency. To this end, a plan is developed to convince the employees of the need for change, to achieve their active participation and to sharpen their awareness of opportunities and risks. (Rischar 2005, Steiger and Hug 2013)

The second sub-step serves to set up a committed, competent, trustworthy change management team willing to change (Rischar 2005, Steiger and Hug 2013). Implementation projects are particularly successful if "champions" are defined at this point who, as individuals, drive implementation forward to a particularly high degree.

In principle, employees from all affected company divisions must also be involved in the composition of the management team (Saberi et al. 2010). At the end of this sub-step, a team has also been appointed for operational planning, which will now be executed in parallel to the organizational planning

In the next step of organizational planning, top management and the management team derive a vision and goals for the organization depending on the strategic considerations previously made and the selected AM technology. The organizational success factors are processed in parallel in order to create corresponding plans. The organizational structure, as a success factor, must agree with the strategy. (Welge and Al-Laham 1992)

Furthermore, in the third sub-step, the organizational culture is revised. In preparation, the project structure is defined, resources are analyzed and decisions are made on external support. The actual culture is then recorded, whereby environmental factors and strengths and weaknesses of the organizational culture can be analyzed. Subsequently, a target culture is designed. (Homma and Bauschke 2015, Welge and Al-Laham 1992)

Less complex structures with the highest possible decentralization are particularly advantageous for the successful implementation of DM (Mellor et al. 2014).

A further element in the step of organizational culture is personnel. Basically, it is important to employ production staff with a wide range of qualifications and to plan training courses for project participants (Welge and Al-Laham 1992).
Especially for this element, there are strong links with operative planning, in which the planning of the technical components indicates the respective personnel requirements. First the qualitative and quantitative determination of the necessary personnel requirement is accomplished. Requirements profiles are created for this purpose. (Welge and Al-Laham 1992) If it is determined during the comparison of actual and target stock that qualification deficits exist, current employees must then be trained further or new employees with the appropriate knowledge must be hired. The planning for this takes place after the comparison.
The training is carried out before the actual implementation (Darbanhosseiniamirkhiz and Wan Ismail 2012). As a guideline, 25-40% of the project budget should be invested in appropriate continuing training measures (Saberi et al. 2010, Zhao and Co 1997). The training concept causes higher costs, but in some cases enables a faster gain of knowledge. On the one hand, knowledge can be created for larger groups of employees. On the other hand, special knowledge can be imparted. New employments as third possible concept make the fastest knowledge gain possible for the enterprise. This possibility is however most expensive and can lead to a rejecting attitude of the other coworkers. If only limited financial resources are available and the technology has a low degree of maturity, cooperation with machine vendors, material suppliers or research institutions is an option in order to reduce costs.

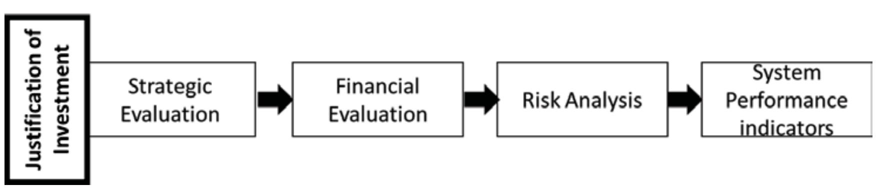
Operational planning transforms the specifications of the strategic planning level into area-related or department-related operational action programs. The responsibility for the operational planning lies in principle with the range and department managers. Besides, a high expenditure must be operated for a successful implementation, observing the consideration of probable effects on the customers and suppliers (Hammer 2015).
Good relationships with customers and AM machine vendors and material suppliers are particularly important when implementing additive manufacturing processes (Darbanhosseiniamirkhiz and Wan Ismail 2012).
Depending on the respective business model, different supply chain configurations must be considered (Lutter-Günther et al. 2015a).
The second step in operative planning is sales planning. Market conditions and the companies own sales are analyzed in order to create a sales forecast (Hammer 2015).
The sales forecast is also part of the approach of Lutter-Günther et al. (Lutter-Günther et al. 2015a) and therefore relevant for the AM. The production capacity from the later sub-step "procurement and supply planning" and the personnel from organizational planning have an influence on sales planning.
The next substep is production program planning, in which the assortment and quantity of products are determined (Hammer 2015).
In the third step, the production flow is planned. The production procedures, lot size, sequence, machine occupancy and production dates must be specified. (Hammer 2015)
The following properties should be taken into account during process planning: Component orientation, maximization of component volume, layer concept, generation of support structures, process speed, laser energy, hatching strategy, temperature control and quality control. (Mellor et al. 2014) Particularly in the context of additive manufacturing processes, it is also necessary to determine the production location (Lutter-Günther et al. 2015a). While a centralized production approach is initially pursued in an early phase of introduction, production can be decentralized at a later point in time (Mellor et al. 2014). The degree of decentralisation also depends on the type of business model (Lutter-Günther et al. 2015a).

As in the approaches for implementing AM, the process chain is also considered in this procedure model for determining the production process. It analyses how the AM can be optimally integrated into the existing production system in order to exploit all the advantages of the technology provided in the business model. For the front-end area it is evaluated under consideration of the available resources, which processes can be accomplished by the company themselves (Mellor et al. 2014). At this point, it must be checked, which resources and competencies are available in the design department in order to include clamping surfaces, reference points and tool paths in the design department that are necessary for possible reworking. If additive structures are to be generated on semi-finished products, resources for material testing and measurement of components must be planned in. It is also necessary to consider software that enables a high degree of automation of the necessary processes. (Manogharan et al. 2015, Möhrle and Emmelmann 2016, Yang et al. 2017)
If sufficient expertise is not available, partnerships can be entered into for the creation of CAD models, for the generation of support structures or for necessary scanners and software (Deradjat and Minshall 2017, Mellor et al. 2014). For the AM machine, it must be checked whether adjustments of the machine parameters are necessary for the respective application. These result, for example, from potential shrinkage and surface roughness (Manogharan et al. 2015).
Similar to the front-end area, the post-processing must also evaluate which processing steps are to be carried out for the specific application.

The last substep is procurement and material staging planning, in which the optimum order quantity is determined. Different methods and models exist to support this planning, which must be selected depending on the respective application. In addition, resource and personnel planning are carried out at this point.
In principle it is possible that strategic projects in the context of the operational planning turn out as unfeasible, so that a feedback is necessary, in order to revise the strategic plan. (Hammer 2015)

<u>Justification of Investment</u>
As part of the justification of the investment, the last review of the investment takes place before the actual implementation of the plans. The justification consists of four sub-steps (cf. *Figure 10*)

Figure 10: Jiustification of Investment



For the strategic evaluation it is necessary to put competition priorities and evaluation criteria in order. The competition priorities result from the classification of the competition factors time, costs, flexibility and quality and depend on the business model. Criteria for the evaluation of the technology are suitable as evaluation criteria. A priority must now be determined for the competitive factors and evaluation criteria. For this purpose, the competition priorities are coordinated with the technology selection criteria and an evaluation is carried out. (Iakymenko et al. 2016)

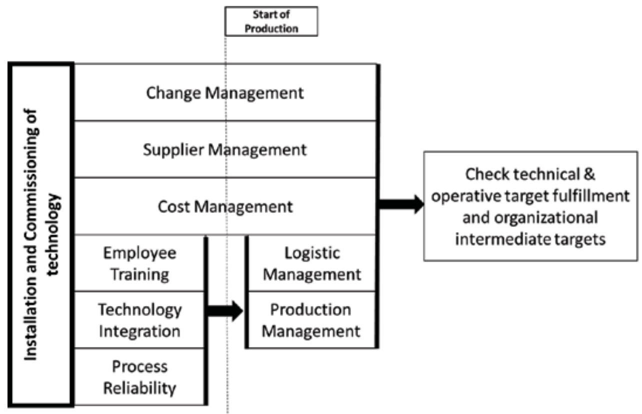
As part of this analysis, the company must commit itself to an AM machine. A feedback for the concretization of the organizational and operational plan, as in the procedure model for AMTs, is omitted, since with only one AM technology, specific planning has already been carried out.
This is followed by a financial evaluation of the integration of the AM technology. In this substep, the team members also define the evaluation requirement. (Iakymenko et al. 2016)
On the one hand, the costs for the technology are determined. Mellor uses in this context the following main factors: machine costs, material costs, personnel costs and the lifespan (Mellor et al. 2014). On the other hand, the organizational expenditure must additionally be calculated for a comprehensive overview of the costs of the integration of DM.
In the third step, a risk analysis is carried out. The decision on the method to be applied must also be made by the team members. Due to the high degree of integration of the DM, a probability and sensitivity analysis can be used. (Iakymenko et al. 2016)

If, after completing these sub-steps, it is determined that the selected AM technology does not yield a positive investment, the decision maker can jump back to the step of selecting the technology at this point to select another technology. However, if the reasons for a poor performance in this step are not due to AM technology, the investment is either discarded or rapid prototyping is introduced first. In case of a positive evaluation, the procedure can be continued.
The last sub-step is the further development of the system performance indicators that were already defined during the strategic considerations. However, the key figures mentioned in this step only allow an evaluation of the technical performance. In order to evaluate the system performance, additional organizational key figures are required that depend on the selected business model (Deradjat and Minshall 2017).

Measurement of performance
The last step of the planning phase is the evaluation of the current company performance on the basis of the previously developed key figures. This step is necessary in order to be able to track changes in company performance at a later point in time.

Installation and Commissioning of Technology
During the installation and commissioning phase, the organizational and operational plan is implemented and the progress is monitored. The existing implementation models for AMTs and the approaches for implementing DM, however, provide only little detail or no procedure for a structured installation and commissioning. Therefore, the procedure model in this paper is based on the AM start-up management model presented by Büsching & Koch, which is divided into two sections. (Büsching and Koch 2017)
Change management is located throughout start-up management and deals with organizational elements (cf.
Figure *11*).

Figure 11: Installation and Commissioning of Technology



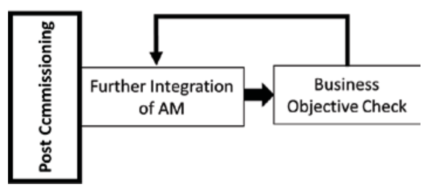
Supplier management, cost management and change management are sub-steps that will be continued in both sections, the phase before and the phase after the start of the actual production. Before the start of the production, change management aims to implement the goals defined in organizational planning for the organizational structure and culture and to adapt unfavorable processes and structures accordingly. Employees must be given room to manoeuvre. Intermediate goals should be celebrated as successes on the way to complete change in the sense of employee motivation. However, the overall goal must not be disregarded. (Schwarz 2019).
For the sub-step supplier management, cooperation with AM machine sellers and material suppliers must be achieved. The type of cooperation depends on how much effort the company has put into adapting the AM technology. The higher the effort for adaptation, the more the company must be involved. Also, for the cooperation with material suppliers, the less tested the material, the stronger the cooperation. Cost management has the task of monitoring the financial situation of the company, as high costs arise from employee training and investments, and only small revenues are generated from the sale of products at that time. The most important variables besides costs are time and quality. (Büsching and Koch 2017)
In the first section of process management, before the start of production, the sub-steps recruitment and/or further training of employees, technology integration and process stability are dealt with in parallel. In the case of employee training, the aspects determined during organizational planning are implemented for the personnel. For this purpose, either new employees are hired or an autodidactic concept or training concept is pursued for the selected employees.
Within the framework of technology integration, the necessary software and hardware are integrated (Chen and Small 1994).
The last partial step before the start of production is process reliability, in which the optimum settings for the AM machine are determined (Büsching and Koch 2017).
A higher degree of automation generally increases the process stability for further operation (Fischer and Rommel 2018).

In the section after the start of production, logistics and production management must take place. Logistics management organizes the flow of materials to and within the production facility. Standards for the use of additive manufacturing processes are established within the

framework of parallel production management in order to reduce complexity and start-up time. (Büsching and Koch 2017)
At the end of the installation and commissioning phase, the technology must work. In the second step of this phase, the technical and operational achievement of the objectives is therefore checked. The system performance indicators further developed when justifying the investment are suitable for this purpose. It is also important to achieve intermediate organizational goals. Here it is important that the employees have fully understood the technology.
Furthermore, progress must be made in the organizational culture and structure as well as in employee qualifications. If the technical and operational goals and the intermediate organizational goals have not been achieved, this phase has not been completed and the corresponding deficits must be improved.

Post-Commissioning
Post-commissioning is the final phase of the integration of additive manufacturing processes and begins with the start of series production. There is no information on this phase in the research on the implementation of DM.

Figure 12: Post-Commissioning



Only change management specifies that changes must be anchored in the last phase (Kostka 2016).
The aim of the post-commissioning phase is the complete development of the marketing and production interface, the further development of technical processes and the achievement of the expected business and marketing benefits.
In addition, organizational changes will be further developed (Sun and Riis 1994).
The phase basically consists of two steps: The further integration of DM and the review of whether the business benefits have been achieved. The steps are performed in a loop so that continuous improvements and the development of routines can be achieved. In theory this phase can last very long or forever (Voss 1988).
It should also be noted that strategic changes may occur over time, such as environmental changes (Sun and Riis 1994), which must be taken into account when integrating DM and adapting the organization. The phase is completed as soon as the desired business benefit has been achieved. If this is not done, at a certain point the company will need to consider whether to continue through the loop of the two steps or stop the implementation.

CONCLUSION AND FUTURE WORK

The procedure model now provides a concrete and clear procedure for the integration of additive manufacturing processes, especially DM, which is elementary for successful implementation. Furthermore, feedback is considered in the procedure model of this work in order to better reflect the real situation of the innovation process. The process model developed here also includes success controls on the respective phase closures and, if possible, specifies responsibilities and methods. Therefore, this paper contributes to the expansion of knowledge in the field of AM research. Another important point in which the process model developed generates an increase in knowledge is organizational and operative planning and implementation. Existing research on the integration of additive manufacturing processes and AMTs shows considerable deficits in the approach.

However, the research carried out also has limitations. Thus, the present work was developed purely theoretically and requires validation in practice. Here it must be specifically examined whether the phases and steps can be separated in each case as clearly as was assumed. Basically, the sharp separation simplifies the success controls. In reality however, a clear separation is difficult to achieve. A further limitation of the results of this work effects from the incomplete specification of responsibilities and methods.

These limitations in turn lead to research areas for future work. One area is the already mentioned validation of the procedure model in practice. The technology of additive manufacturing can be introduced in companies with different product structures with the help of the model, developed in this paper. Based on the results, the structure and the granularity of the model could be adapted and evaluated for its practical suitability. In practical application it would also be interesting to observe what influence a rapid prototyping history has on the integration of the DM and what conclusions can be drawn for a process model, since only a few details can be found in the literature.

Furthermore, there are new technical developments in software and hardware that can considerably facilitate integration. These could simplify the processes in the organizational and operative planning and implementation and make an adjustment of the procedure model necessary.

In addition, the business models can also change, since the profitability of the DM will also be ensured in large series production in the future through faster and hybrid manufacturing processes.

But it is above all the necessary integrated integration of AM that is currently deterring many companies, even with promising business models. So that structured procedure models have to be developed for different business models, for different company sizes and for different technology breakthroughs of companies.

The work on this paper was supported by grants from the Federal State of North Rhine-Westphalia using funds from the European Regional Development Fund (EFRE) 2014-2020, "Investments in Growth and Employment"

REFERENCES

Besenfelder C, Niehues S, Pracht H (2019) Zukunftsperspektiven der Additiven Fertigung. *ZWF*. 114(3):134–136.

Besenfelder C, Niehues S, Wortmann B, Pracht H (2018) Integration additiver Fertigungstechnologien. *ZWF*. 113(1-2):64–66.

Büsching J, Koch R (2017) Ramp-Up-Management in Additive Manufacturing - Technology Integration in existing Business Processes. Bourell D, Crawford R, Seepersad C, Beaman J, Fish S, eds. *Proceedings of the 28th Annual International Solid Freeform Fabrication Symposium. Solid Freeform Fabrication 2017* (Austin, Texas), 2585–2596.

Chen I, Small M (1994) Implementing advanced manufacturing technology: An integrated planning model. *Omega*. 22(1):91–103.

Chen L, He Y, Yang Y, Niu S, Ren H (2017) The research status and development trend of additive manufacturing technology. *Int J Adv Manuf Technol*. 89(9-12):3651–3660.

Chew B, Leonard-Barton D, Bohn R (1991) Beating Murphy's Law, https://sloanreview.mit.edu/article/beating-murphys-law/.

Darbanhosseiniamirkhiz M, Wan Ismail WK (2012) Advanced Manufacturing Technology Adoption in SMEs: an Integrative Model. *Journal of Technology Management & Innovation*. 7(4):112–120.

Deradjat D, Minshall T (2017) Implementation of rapid manufacturing for mass customisation. Forthcoming.

Feldmann C, Schulz C, Fernströning S (2019) Einordnung in das Innovationsmanagement. Feldmann C, Schulz C, Fernströning S, eds. *Digitale Geschäftsmodell-Innovationen mit 3D-Druck. Erfolgreich entwickeln und umsetzen*, 91–92.

Fischer A, Rommel S (2018) *3D-Druck für Unternehmen. Erfolgreich produzieren mit Fused Layer Modeling (FLM): Grundlagen - Anwendungsbeispiele - Best Practices* (Hanser, Carl, München).

Friedli T (2006) *Technologiemanagement. Modelle zur Sicherung der Wettbewerbsfähigkeit* (Springer, Berlin).

Frohlich M (1998) How do you successfully adopt an advanced manufacturing technology? *European Management Journal*. 16(2):151–159.

Fueglistaller U, Müller CA, Müller S, Volery T, eds. (2016) *Entrepreneurship. Modelle - Umsetzung - Perspektiven : it Fallbeispielen aus Deutschland, Österreich und der Schweiz,* 4th ed. (Springer, Wiesbaden).

Gebhardt A, Hötter J-S (2016) *Additive manufacturing. 3D printing for prototyping and manufacturing* (Hanser Publishers; Hanser Publications, Munich, Cincinnati).

Grund M (2015) *Implementierung von schichtadditiven Fertigungsverfahren. Mit Fallbeispielen aus der Luftfahrtindustrie und Medizintechnik*. Light Engineering für die Praxis (Springer Vieweg, Berlin).

Gunawardana K (2006) Introduction of Advanced Manufacturing Technology: A Literature Review. *Sabaragamuwa University Journal*. 6(1):116–134.

Gupta K (2017) *Advanced manufacturing technologies. Modern machining, advanced joining, sustainable manufacturing*. Materials Forming, Machining and Tribology (Springer, Cham, Switzerland).

Hague R (2006) Unlocking the Design Potential of Rapid Manufacturing. Hopkinson N, Hague RJM, Dickens PM, eds. *Rapid manufacturing. An industrial revolution for the digital age* (John Wiley, Chichester, England), 5–18.

Hammer R (2015) *Unternehmensplanung: Planung und Führung* (De Gruyter).

Hokinson N, Dickens PM (2003) Analysis of rapid manufacturing—using layer manufacturing processes for production. *Proceedings of the Institute of Mechanical Engineers, Part C : Journal of Mechanical Engineering Science*. 217(C1):31–39.

Homma N, Bauschke R (2015) *Unternehmenskultur und Führung. Den Wandel gestalten - Methoden, Prozesse, Tools* (Springer Fachmedien Wiesbaden, Wiesbaden [Germany]).

Iakymenko N, Alfnes E, Thomassen MK (2016) A differentiated approach for justification of advanced manufacturing technologies. *Adv. Manuf.* 4(3):257–267.

Klemp E, Pottebaum J (2016) Additive Fertigungsverfahren im Kontext von Industrie 4.0. Vogel-Heuser B, Bauernhansl T, Hompel Mt, eds. *Handbuch Industrie 4.0 Bd.3: Logistik* (Springer Science and Business Media; Springer Vieweg, [Place of publication not identified]), 1–21.

Kostka C (2016) *CHANGE MANAGEMENT* (CARL HANSER Verlag GMBH &, [Place of publication not identified]).

Lachmayer R, Lippert RB, Fahlbusch T (2016) *3D-Druck beleuchtet* (Springer Berlin Heidelberg; Imprint: Springer Vieweg, Berlin, Heidelberg).

Lerch C, Zanker C, Kühlbauch P (2012) Strategische Prozessmodernisierung in KMU. Handlungsfelder und innovative Methode. *wt Werkstatttechnik online*. 102(3):125–129.

Lutter-Günther M, Seidel C, Kamps T, Reinhart G (2015a) Implementation of Additive Manufacturing Business Models. *AMM*. 794:547–554.

Lutter-Günther M, Seidel C, Kamps T, Reinhart G (2015b) Implementation of Additive Manufacturing Business Models. *AMM*. 794:547–554.

Maghazei O, Netland T (2017) Implementation of Industry 4.0 Technologies: What Can We Learn from the Past? Lödding H, Riedel R, Thoben K-D, Cieminski G von, Kiritsis D, eds. *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing. IFIP WG 5.7 International Conference, APMS 2017, Hamburg, Germany, September 3-7, 2017, Proceedings, Part I*, IFIP Advances in Information and Communication Technology, Vol. 513 (Springer International Publishing; Imprint; Springer, Cham), 135–142.

Manogharan G, Wysk R, Harrysson O, Aman R (2015) AIMS – A Metal Additive-hybrid Manufacturing System: System Architecture and Attributes. *Procedia Manufacturing*. 1:273–286.

Mellor S, Hao L, Zhang D (2014) Additive manufacturing: A framework for implementation. *International Journal of Production Economics*. 149:194–201.

Möhrle M, Emmelmann C (2016) Fabrikstrukturen für die additive Fertigung. *ZWF*. 111(9):505–509.

Muita K, Westerlund M, Rajala R (2015) The Evolution of Rapid Production: How to Adopt Novel Manufacturing Technology. *IFAC-PapersOnLine*. 48(3):32–37.

Pagell M, Handfield D RB, Barber AE (2000) EFFECTS OF OPERATIONAL EMPLOYEE SKILLS ON ADVANCED MANUFACTURING TECHNOLOGY PERFORMANCE. *Production and Operations Management*. 9(3):222–238.

Rahardjo J, bin Yahya S (2010) Advanced Manufacturing Technology Implementation Process in SME: Critical Success Factors. *JTI*. 12(2):101–108.
Rischar K (2005) *Veränderungsmanagement. Praxiserprobte Instrumente für Führungskräfte* (Expert-Verl., Renningen).
Ruffo M, Hague R (2007) Cost estimation for rapid manufacturing ' simultaneous production of mixed components using laser sintering. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*. 221(11):1585–1591.
Saberi S, Mohd. Yusu R, Zulkifli N, Megat Ahma MMH (2010) Effective Factors on Advanced Manufacturing Technology Implementation Performance: A Review. *J. of Applied Sciences*. 10(13):1229–1242.
Schallmo DRA (2018) Theoretische Grundlagen. Schallmo DRA, ed. *Geschäftsmodelle erfolgreich entwickeln und implementieren. Mit Aufgaben, Kontrollfragen und Templates,* 2nd ed. (Springer Gabler, Berlin, Heidelberg), 11–45.
Schmutzler C, Teufelhart S, Reinhart G, Zäh MF (2016) Neue Produktionstechnologien am Beispiel der additiven Verfahren. Lindemann U, ed. *Handbuch Produktentwicklung,* Hanser eLibrary (Hanser, München), 953–977.
Schwarz C (2019) Banken im Strukturwandel – von der Filiale zum Multikanalansatz. Muchna C, ed. *Aspekte des Innovations- und Changemanagements. Ein Theorie-Praxis-Transfer* (Springer Gabler, Wiesbaden, Germany), 227–262.
Singh RK, Garg SK, Deshmukh SG, Kumar M (2007) Modelling of critical success factors for implementation of AMTs. *Jnl of Modelling in Management*. 2(3):232–250.
Small MH, Yasin MM (1997) Developing a framework for the effective planning and implementation of advanced manufacturing technology. *Int Jrnl of Op & Prod Mnagemnt*. 17(5):468–489.
Steiger T, Hug B (2013) Psychologische Konsequenzen von Veränderungen. *Handbuch Angewandte Psychologie für Führungskräfte. Führungskompetenz und Führungswissen,* 3rd ed. (Springer, Berlin).
Steinwender A (2016) Frühe Berücksichtigung von Aspekten der generativen Fertigung. Abele T, ed. *Die frühe Phase des Innovationsprozesses. Neue, praxiserprobte Methoden und Ansätze,* FOM-Edition, FOM Hochschule für Oekonomie & Management (Gabler, Wiesbaden), 191–215.
Sun H, Riis JO (1994) Organizational, technical, strategic, and managerial issues along the implementation process of advanced manufacturing technology—a general framework of implementation guide. *Int. J. Hum. Factors Manuf.* 4(1):23–36.
VDI (2013) *Additive manufacturing processes, rapid manufacturing - Laser sintering of polymer parts - Quality control*.
Voss CA (1988) Implementation: A key issue in manufacturing technology: The need for a field of study. *Research Policy*. 17(2):55–63.
Welge MK, Al-Laham A (1992) *Planung. Prozesse - Strategien - Massnahmen* (Gabler, Wiesbaden).
Yang L, Hsu K, Baughman B, Godfrey D, Medina F, Menon M, Wiener S (2017) *Additive manufacturing of metals. The technology, materials, design and production*. Springer series in advanced manufacturing (Springer, Cham).
Zahn E (1988) Produktionsstrategie. *Handbuch Strategische Führung* (Gabler Verlag, Wiesbaden), 515–543.

Zeyn H (2017) *Industrialisierung der Additiven Fertigung. Digitalisierte Prozesskette - von der Entwicklung bis zum einsetzbaren Artikel,* 1st ed. (Beuth Verlag GmbH, Berlin).

Zhao H, Co HC (1997) Adoption and implementation of advanced manufacturing technology in Singapore. *International Journal of Production Economics*. 48(1):7–19.

DECISION SCIENCES INSTITUTE

Lean Approach to Business Model Development for Smart Product Service Systems Using the Example of Maintenance

Dr. Laura Reder
Fraunhofer Institute for Material Flow and Logistics
Email: laura.reder@iml.fraunhofer.de

Dipl.-Logist. Christoph Pott
Fraunhofer Institute for Material Flow and Logistics
Email: Christoph.pott@iml.fraunhofer.de

Alexander Michalik, M.Sc.
Technical University Dortmund, Chair of Corporate Logistics
Email: michalik@lfo.tu-dortmund.de

ABSTRACT

Through increasing digitization, classic Product Service Systems (PSS) are further developed into smart PSS. An early business model perspective is an important success factor of the development of smart PSS. Therefore, in this paper a procedure model for the lean design of business models for smart PSS is presented.

KEYWORDS: Lean Startup, Business Model Development, Smart Product Service Systems

BUSINESS MODEL DEVELOPMENT AS A KEY SUCCESS FACTOR FOR SMART PRODUCT SERVICE SYSTEMS

Increasing digitalization and the widespread availability of innovative technologies are leading to both product and service innovations. The development and use of new products and services for mastering complex processes and securing competitive advantages are essential, especially for industrial companies, which are characterized in many processes by a **combination of human competencies, skills and technologies** (cf. Gorldt et al. 2017, p. 355). Existing benefits in kind (products) are often combined with services to create **Product Service Systems (PSS)**. Through an additional combination with intelligent service modules, classic PSS are further developed into **smart PSS** (cf. Steven 2019, p. 190).

Due to the importance of smart PSS both for the competitiveness of companies and for the design of a future digital and integrative working environment, its development is being promoted within the framework of research programs. The **research project "Innovation Laboratory Hybrid Services in Logistics"** is a lighthouse project that focuses on a **human-centered design** of the future working environment, taking into account the dimensions of people, technology and organization. For such a **"Social Networked Industry"** it is necessary to enable new forms of cooperation and interaction between people and cyber physical systems (vertical networking) (cf. ten Hompel et al. 2017a, p. 3). In order to support and enable these **human-machine interfaces**, the aim of the innovation laboratory is to advance the development of smart PSS for production and logistics and to strengthen acceptance and, associated with this, the industrial introduction of these new technical solutions (cf. ten Hompel et al. 2017b, p. 1 f.).

The investment in appropriate resources for the development and implementation of smart PSS can only be carried out sensibly if, on the one hand, there is a market for corresponding solutions and, on the other hand, a perceptible (customer) benefit exists and can be communicated. Accordingly, an early **business model perspective** is an important success factor and an integral part of the development of smart PSS. For this purpose, a **procedure model for the design of business models** will be developed within the framework of the innovation laboratory in order to provide the smart PSS gained with timely business models. The use of the developed process model is illustrated in the context of an application example from **maintenance**.

SMART PRODUCT SERVICE SYSTEMS

First, the basics of PSS and smart PSS are explained. Following a definition, a thematic classification of smart PSS in the context of the research project "Innovation Laboratory Hybrid Services in Logistics" will follow, as well as an explanation of their significance for both a social networked industry and for the business-to-business application area.

Basics of Smart Product Service Systems

The term **Product Service Systems (PSS)**, also referred to as hybrid service packages or solutions (cf. Michalik et al. 2018a, p.310), was coined in particular in the context of the efforts of mechanical and plant engineering companies to differentiate themselves from increasing international competition. For this purpose, in addition to offering benefits in kind (production plants), greater emphasis is being placed on additional services (e.g. plant maintenance) (cf. Meier/Uhlmann 2012, p. 1 f; Steven 2016, p. 200).

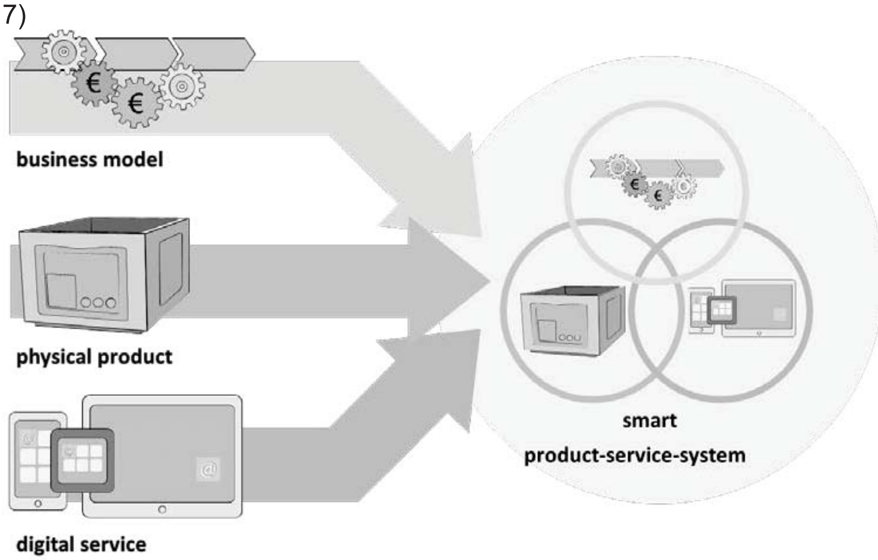
PSS is the term used to describe holistic and individual, customer-oriented problem solutions which are characterized above all by **systematic and integrated** planning, development and provision of the interlinked **material and service components** (cf. Meier et al. 2005, p. 529; Meier/Uhlmann 2012; p. 1 f; Grandjean et al. 2017, p. 18). This transformation from a pure producer of goods and services to a provider of customer-specific combinations of goods and services is also known as **servitization** and is not limited to mechanical engineering and plant construction, but can also be observed in other sectors (cf. Steven 2019, p. 133; Wiesner et al. 2014, p. 287).

Although many companies offer additional services, the competitiveness of purely product-oriented solutions seems to be no longer guaranteed in the digital age (cf. Chowdhury et al. 2018, p. 26). **Digitization** has a major influence on the provision of services and thus on the customer-specific benefit promise of PSS (cf. Grandjean et al. 2017, p. 8; Bruhn/Hadwich 2016, p. 11). In order to remain competitive, companies must therefore integrate digital technologies into the development and deployment of PSS. These PSS equipped with digital technologies are also referred to as **smart PSS** (cf. Grandjean et al. 2017, p. 8; Valencia et al. 2015, p. 11 f; Abramovici, 2018, p. 4). Smart PSS are essentially characterized by a high degree of (partial) autonomy, networking capability, and personalization capability as well as user friendliness and user centricity. Further characteristics are dynamic reconfigurability throughout the product life cycle, real-time reactivity to environmental changes and context sensitivity (cf. Abramovici et al. 2016, p. 18).

The Internet of Things and Services is regarded as an important technological basis for smart PSS. On the basis of the Internet of Things, **cyber-physical systems (CPS)** are used, which lead to the networking of the physical and the digital world through sensors and actuators (cf. Kagerman et al. 2013, p. 84; Abramovici 2018, p. 4). The exponentially increasing connectivity already enabled networking of 8.4 billion devices worldwide in 2017. The turnover with such devices and software services offered via them was estimated at almost two trillion USD. In 2020, the number of networked devices is expected to rise to 20.4 billion (see Gartner 2017). This high degree of networking represents an excellent **application prerequisite for smart PSS**, so that development and deployment activities must be intensified. The BMBF-funded project "Innovation Laboratory Hybrid Services in Logistics" focuses its activities in particular on the application-oriented development of smart PSS for production and logistics.

In order to develop smart PSS in a customer-oriented manner, the requirements for a solution must be derived from the respective customer needs (cf. Gorldt et al. 2017, p. 366). In addition to the technical requirements, this also requires an analysis and, if necessary, **adaptation of established or the development of completely new business models**. The use of CPS thus enables a PSS to provide new disruptive (digital) services (so-called smart services) and also requires the adaptation of existing and/or the development of new business models. Overall, **smart PSS** are thus formed by the interaction of **digital services** through the use of innovative hardware and software [in this context, hardware includes smart devices, cloud computing and augmented reality. Software includes service-oriented architectures, semantic and big data technologies (see Abramovici 2018, p. 3)], **physical products** and a **business model** tailored to them (cf. Figure 1).

Figure 1: Components of Smart Product Service Systems (based on ten Hompel et al. 2017b, p. 7)



Importance of Smart Product Service Systems

<u>Importance of smart PSS for a Social Networked Industry</u>
Interaction and communication between man and machine already play an important role in today's world of work, and this will increase significantly in the future (cf. ten Hompel et al. 2017b, p. 1). Therefore, it is of high relevance to develop and test technological solutions and to include them in the future digitalized world of work. This so-called **Social Networked Industry**

describes industrially oriented forms of social networks that enable a new kind of cooperation between people and cyber physical systems (cf. ten Hompel et al. 2017a, p. 3).

An **integrated view of CPS and PSS** is becoming more and more important for the design of this social networked industry and future working environments. The use of CPS enables an improvement of the human-machine interface and the combination of human performance, physical objects and virtual services provides for the emergence of new smart PSS. Autonomously interacting CPS ensure the functioning connectivity between humans and technology (cf. Drossel et al. 2018, p.199 f.), so that the **smart PSS** can come into contact with its environment and exchange information (cf. Steven 2019, p.191). The use of smart PSS thus enables the realization of the vision of **socially compatible human-machine communication** and interaction in the context of a social networked industry.

The **"Innovation Laboratory Hybrid Services in Logistics"** is currently the largest application-oriented research project for the investigation and design of an innovative human-technology interaction. This is accompanied by the development and prototypical use of smart PSS in production and logistics (cf. ten Hompel et al. 2017b, p. 4). In order to also do justice to financial and economic aspects in the context of application orientation, the smart PSS gained will be provided with timely **business models**.

<u>Importance of smart PSS in the B2B Sector</u>
After smart PSS (such as the iPod as a benefit in kind and iTunes as an associated digital service) have already been offered on the business-to-consumer (B2C) market for some time, smart solutions are also emerging in the **business-to-business (B2B) sector** that are oriented towards industrial users and their much more complex areas of application.

Increasing globalization and **highly competitive markets** in the field of high-quality machines and plants as well as **more individual customer requirements** require a differentiation beyond previous efforts in the field of servitization. The digital services offered complement the material product and thus represent added value both for the company (USP) and for its customers (value proposition) (cf. Steven 2019, pp. 131, 170f.).

In addition to maintaining a competitive position and opening up new markets, **customer loyalty** plays a key role in offering smart PSS in the B2B sector. By offering digital product-related services during the usage phase, the customer relationship can be strengthened and additional, continuous revenues generated.

The offer of such digital services requires an **adaptation or reorientation of the classic business models**, taking into account the individual benefit promise generated by the smart PSS.

Smart PSS therefore have a high significance both for the future vision of a social networked industry and for the B2B sector and will be an integral part of a digital and networked economy, or in some cases already are. This is accompanied by the need to develop new mechanisms and approaches for the way in which value can be created for the actors involved through the use of smart PSS (cf. Boßlau et al. 2017, p. 299f.). Based on this need, a **lean process model for the generation of business model approaches for smart PSS** will be developed in the following. The application will be initiated on the basis of a showcase from the innovation laboratory.

RESEARCH FRAMEWORK

The research methodology follows the principles of Design Science Research (DSR) utilizing a maintenance showcase stemming from the Innovation Laboratory. DSR is a proven methodology for the structured design of artifacts in information systems research. The authors utilize the Design Science Research Methodology, as proposed by Peffers et al., which provides the following steps**: Identify Problem, Define Objectives, Design and Build, Demonstration, Evaluation, and Communication** (Peffers et al. 2007, p. 45-77). Firstly, the authors gathered theoretical and practical insight from literature and experts of the innovation laboratory. Thus, they derive requirements for the targeted solution. For this, the authors chose to conduct multiple workshops with the responsible experts of the different showcases of the innovation laboratory. Workshops are an appropriate measure for collecting data in a domain-specific organizational context (Orngreen and Levinsen 2017, p. 70-81). The requirement is to develop business models in an early phase based on smart PSS. Secondly, the authors mark out the scope of this research, which is to develop a lean approach for sketching out business models for the showcase at hand. Thirdly, the authors present their findings concerning a lean development design of business models for smart PSS. Fourth, the authors present their procedure model for the lean business model design approach and further evaluate it based on a showcase in the field of maintenance.

DEVELOPMENT OF BUSINESS MODELS FOR SMART PRODUCT SERVICE SYSTEMS

It is necessary to integratively adapt, or renew business models in order to develop innovative, smart PSS. In addition to a definition and the dimensions of business models, their significance, especially in the context of digitization, will be discussed below. The procedure model used in the innovation laboratory to develop business models for smart PSS will then be presented.

Definition of Business Model

The concept of the business model was first discussed in the 1950s (cf. Osterwalder et al. 2005, p. 4) and was transferred to the business context with the 'New Economy' (cf. for this and in the following Becker et al. 2012, p. 8 ff.).

Basically, two types of business model can be distinguished: universal and partial business models. While **universal models** are generic business models that can be transferred to different companies and industries, **partial models** describe industry-specific business models. The **definition** of Osterwalder and Pigneur conveys an intuitive and **universal basic understanding** of a business model. They define a business model as "[...] the basic principle according to which an organization creates, imparts and captures value". (Osterwalder/Pigneur 2011, p. 18, for a definition overview see Schallmo 2014, p. 2-6).

Business models represent in an aggregated, conceptual way a considered operational unit or sales performance (cf. Grandjean et al. 2017, p. 70f.). In this context, necessary **value creation mechanisms** such as organizational, logistical and production-related processes for providing this sales performance are presented in simplified form (cf. Steven 2019, p. 187). The focus of business model development is on **customer benefit** as a unique service promise to the customer. The business model also takes into account the costs associated with the provision of services and the expected revenues (see Stähler 2001, p. 41ff; Wirtz 2010, p. 8 ff).

Importance of Business Model Development in the Context of Smart Product Service Systems

The business model is currently experiencing another major upswing in the context of ongoing digitization and Industry 4.0. The offer of digital PSS changes business relationships between companies, raises a new set of strategic choices about how value is created and captured, how companies work with traditional and new partners and how they secure their competitive advantage (cf. Porter/Heppelmann 2014, pp. 67). Hence, the development of smart PSS equally requires the development of new, valuable **business models** in order to tap new business potentials (cf. Gausemeier/Wieseke 2017, p. 23). **Business model development** offers a structured approach for this, fulfilling the following **functions** (cf. Grandjean et al. 2017, p. 70f, Wirtz 2010, p. 18; Osterwalder/Pigneur 2011, p. 19):

- One of the most important functions of a business model is **to better understand** one's own business and **customer benefits**. This aspect is essential for the company's success, so that it can be useful to think about a practicable business model not only after the completion of a product, but already during the development process.
- A further function is the **improvement of the current business and the development of new business segments**. This enables a differentiation from the competition and the development of new customer groups.
- Furthermore, the modelling of the value creation mechanisms of a company or business unit within the framework of business model analysis provides a basis for the **systematic analysis of the decision-making situation**.
- Furthermore, the **complexity is reduced** by the aggregated representation of the value creation processes in a business model.
- In addition, a business model can be used as a **strategic instrument** for the systematic development and design of an operational unit, e.g. in the context of the development of a sales performance.

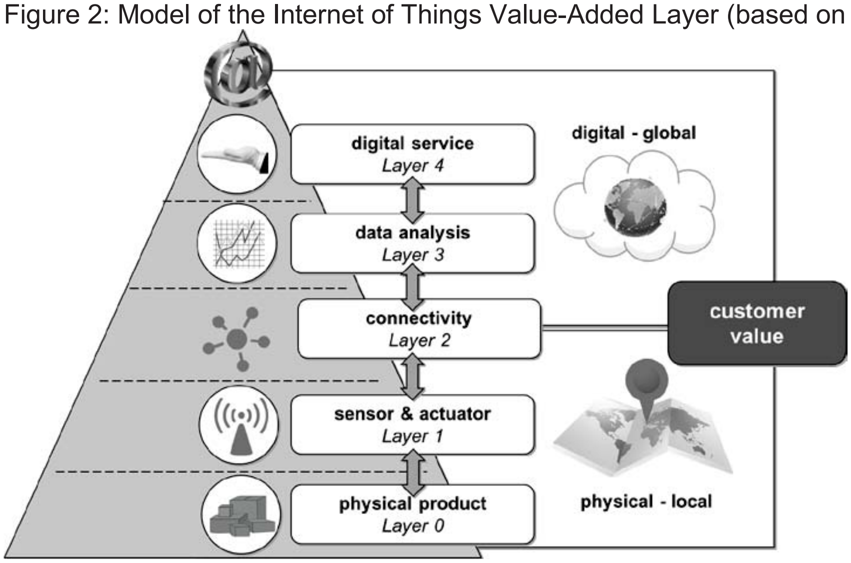
Digitalization in particular influences existing business models and leads to the development of new business model patterns (cf. Buhler/Maas 2017, p. 47 ff.). At present, however, it poses a challenge for companies to keep an overview of the opportunities and risks associated with digitization and industry 4.0 (cf. Geisberger/Broy 2012, p. 175) and to derive new, innovative business models (cf. Gausemeier/Klocke 2016, p. 14). According to a study by McKinsey, 80% of all companies **expect industry 4.0** and related technological innovations **to have an impact on their own business model**. A structured further development of business models rarely takes place in everyday business life, despite their high relevance for corporate success (cf. McKinsey 2015, p. 8; Gausemeier/Wieseke 2017, p. 7). The reason for the **failure** of many companies is therefore often a **failure to adapt the business model** to changing conditions and a lack of focus on the needs of customers. [A well-known example of this is Kodak, which has lost its market share for photography products to its competitor Fujifilm due to a lack of adaptation of its own business model to market changes (cf. Fujifilm; Schultz 2012).]

In order to meet these challenges, it is often necessary to take a holistic perspective from product and business model development, which is often lacking. In addition to the development of smart PSS, within the framework of the **Innovation Laboratory**, the possible exploitation of smart PSS through an initial **systematic and structured development of business model approaches** is also examined. Business model development is of particular importance because it is not sufficient to develop a product: For product success, it is also of great importance to examine marketability and **customer benefit**. Ideally, this test should be carried out **continuously and parallel to the solution development** in order to identify possible weak points or product characteristics that need to be added during the development process and to take them into account in the next development phases. In order to guarantee this integrative development and examination of the marketability, a **lean procedure model** is used in the context of the innovation laboratory, which already examines the fit between product characteristics and requirements of the target customers during the development process in close coordination cycles.

Thanks to very **successful business model developments**, companies have often succeeded in revolutionizing an entire industry **without a physical product** (cf. Gassmann, et al. 2013, p. 5; Gausemeier/Wieseke 2017, p. 24). Originally, physical products such as music CDs, books or video cassettes were replaced by digital products, which are always available on **platform markets** everywhere. However, the disruption potential of business model innovations is not limited to the B2C sector. In **B2B markets**, too, it can be assumed that business model innovations will be similarly revolutionary (cf. Gausemeier/Wieseke 2017, p. 24).

Integrated development of products, services and business models is also gaining in importance for **companies in mechanical and plant engineering**. They are increasingly moving along a product-service continuum, in which service-based value creation takes up an increasingly larger share of the revenue model. The benefit promise is generally based on a benefit in kind, a so-called "installed basis" (Braun 2013, p. 26), whereby the use of CPS in combination with intelligent, Internet-based services is expected to result in a steady decoupling of physical capital goods and thus increased scalability of smart PSS in the future (cf. Abramovici 2018, p. 4). The digitization of this installed base and the design of associated business opportunities have a decisive influence on the success of a company (cf. Henke/Hegmanns 2017, p. 339). Figure 2 shows the model of Bilgeri et al., which is used to generate **customer value** by **linking physical products with digital services (smart PSS)** (cf. Bilgeri et al. 2015, p. 14).

Figure 2: Model of the Internet of Things Value-Added Layer (based on Bilgeri et al. 2015, p. 14)



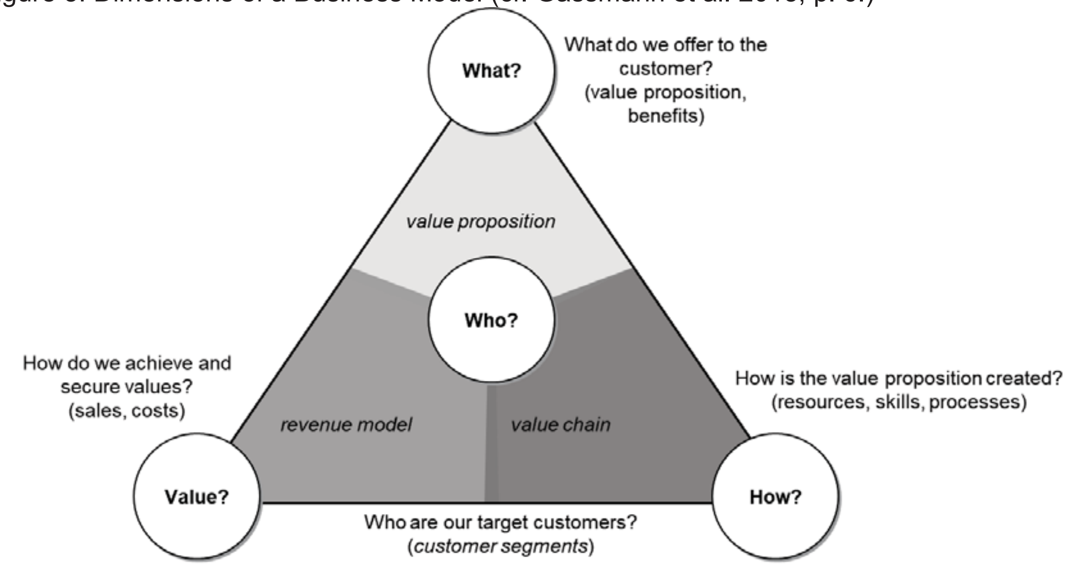
The so-called **model of the Internet of Things (IoT) value creation layer** represents the development from an analog, physical product to a digital service on the basis of successive layers. In the first step, a **purely physical product** is considered. In the second stage, sensors and actuators are added to this product so that data can be stored and shared. With the achievement of communication capability, the third IoT layer is reached and the product is on the boundary between the physical and digital worlds. Already at this point it is useful to initiate the **development of a business model** that goes beyond the mere sale of the physical product

and takes into account the data captured by the **digital component**. If the transmitted data is then transferred to an analysis unit or software, the fourth layer of the model is reached and the business model can be expanded and innovated (cf. Bilgeri et al. 2015, p. 13 ff.).
Current studies show, however, that companies that offer products in the context of industry 4.0 primarily market the **functional components** (cf. layers 0 and 1 in the model) of the product. For example, with intelligent, remote-controlled thermostats ("Smart Home"), the customer pays for the convenience of remotely controlling his heating system via app. At the same time, this results in large amounts of data that could be used in the course of further developing the business model. Other conceivable approaches for new business models include the sale of additional services or the customer's participation in cost savings (see Henke/Hegmanns 2017, p. 339).
The business model development within the framework of the innovation laboratory addresses these current challenges: In order to leverage the often untapped potential, an **early and continuous business model development** takes place here. On the one hand, this enables the early development of smart PSS product features to be evaluated and, on the other hand, potential marketing models to be identified, even beyond the pure sale of the non-cash component. The aim is to sharpen the service offering for the customer, identify customer segments and sales channels as well as the resources required. In addition, costs and potential revenue channels arising during the provision of services are to be identified.
Thus, the development of smart PSS takes a **holistic perspective consisting of physical product, digital service and business model** (see also Figure 1).

Dimensions of Business Model Development

A uniform **description pattern** is required for the analysis of existing business models and for the development of new ones (cf. Schallmo 2014, p. 21). Figure 3 illustrates the **four dimensions** of business model development that address key issues relating to the design of a business model. The four dimensions can be divided into internal and external dimensions. The **external dimensions** include the **customer segments** (**Who** are our target customers?) and the **benefit promise** (**What** do we offer our customers?). The **internal dimensions** address the **value-added processes** required to produce the service (**How** do we produce the service?) and the **earnings mechanisms** (**How** do we achieve and secure value?) (cf. Gassmann et al. 2013, p. 6; Schallmo 2014, p. 23).

Figure 3: Dimensions of a Business Model (cf. Gassmann et al. 2013, p. 6.)



These business model dimensions and associated key questions serve as **description patterns** for the analysis of existing and the design of new business models. The aim is to combine these dimensions in such a way that it is difficult for the competition to imitate a service that simultaneously represents a unique value proposition for the customer (cf. Schallmo, 2013, p. 23). These dimensions represent the fundamental frame of reference for business model development within the framework of the innovation laboratory.

If at least two of these four dimensions are subject to change, this is also referred to as **business model innovation** (cf. Gassmann et al. 2013, p. 6). In contrast to the pure further development of existing business models, a completely new composition and design of individual business model dimensions is aimed at, resulting in previously unavailable products and services (cf. Labbé/Mazet 2005, p. 897 f.; Lindgardth et al. 2009, p. 2; Mitchel/Coles 2003, p. 17).

Integration of the Lean Startup Approach into the Development of Business Models for smart PSS

The Internet of Things and smart PPS will not change business completely, as the rules of competition and competitive advantage remain the same. However companies need to understand these rules better than ever, as smart PSS open a new era of competition (cf. Porter/Heppelmann 2014, pp. 67-68). Due to this new competition and the importance of continuous review and (further) development of business activities, the design of business models must be an **integral part** of the product development of smart PSS. In the following, the **procedure model** used in the innovation laboratory **for the business model development of smart PSS** is presented.

In recent years, the number of approaches to developing business models has increased (cf. Schallmo 2013, p. 47). However, many of the approaches are at such a level of abstraction that no connections between the individual elements of the business model are explained and consequently there is **no procedural model** for developing business models (cf. Schallmo, 2013, p. 48). Other approaches, including those by Buchere (2010), Chesbrough (2007) and

Giesen et al. (2007), provide a process model for the development of business models and are generally applicable, but focus on the **analysis of existing business models**. A further development or business model innovation, as often required with smart PSS, is not achieved. The approaches of Mitchel/Coles (2004), Teece (2010) and Wirtz (2010) concentrate both on the analysis of existing and the development of new business models and propose techniques for the design of business models (cf. Schallmo 2013, pp. 47-113). However, none of the approaches takes into account the integrative customer-oriented **development of products** and **services** as well as an associated **business model** from a **holistic perspective**. A comprehensive overview and description of existing approaches to business model development can be found in Schallmo (2013).

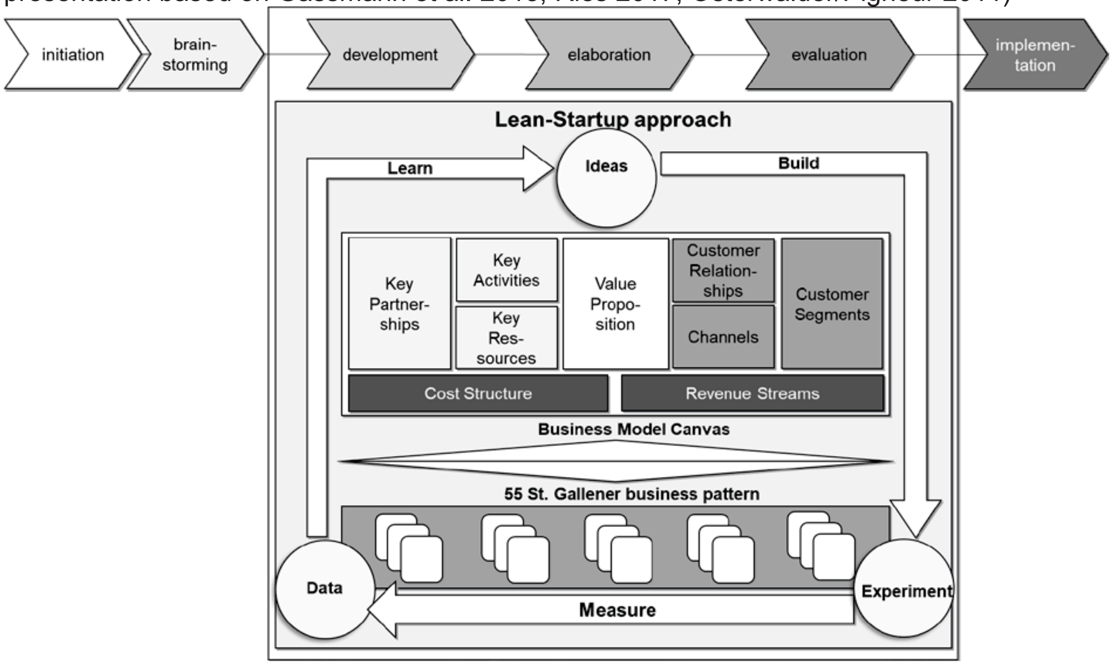
In addition to these classical approaches, the **lean approach**, originally a production philosophy to avoid waste, has recently been transferred to other areas, including business model development (cf. Ghezzi/Cavallo 2018, p. 7). Based on the general understanding of Lean, in the business model context this is understood as the attempt to streamline or even abolish all activities and processes that are not required by the **target customer**. This **customer orientation and centering** is increasingly becoming the focus of new lean approaches. In particular, the **lean startup approach** pursues high customer and market orientation in the development of products. The approach is characterized by short, iterative product development cycles and early marketing, so that customer feedback in the early development stage enables further learning and improvement of the product (cf. Ries 2017, p. 77 ff.).

This strong focus on customer needs is consistent with the understanding of **customer benefit** in business model development: both the lean startup approach and the business model dimensions (see Figure 3) focus on the value proposition to the customer (see Ghezzi/Cavallo 2018, p. 28).

Based on these parallels and first case studies, which classify an application of the approach in business model development as synergetic and promising (cf. Ghezzi/Cavallo 2018, p. 29), the **lean startup approach is applied** in the previous model for the development of business models. This enables a continuous **comparison between product development and customer benefit**. The comparison is achieved through a recurring loop of prototype construction, experimentation & measurement and continuous learning (**build, measure, learn principle**) (see Figure 4). The individual phases of the process model and the lean startup approach are explained below.

Gassmann et al.'s process model used here as a basis represents a proven construction methodology for business model development, which has been developed and applied over five years within the framework of action research and consulting projects with numerous internationally leading companies from a wide range of industries (cf. Gassmann et al. 2013, p. 15). The **six phases** of the model and the integration of the lean startup approach are shown in Figure 4 (cf. Gassmann et al. 2013, p. 16 ff.).

Figure 4: Procedure Model for the Business Model Development of smart PSS (own presentation based on Gassmann et al. 2013; Ries 2017, Osterwalder/Pigneur 2011)



The **initiation phase** lays the foundation for business model development and includes a holistic analysis of the current business model environment. In the first instance, different cause-effect relationships are uncovered, e.g. through the analysis of actors or possible influencing factors.

The second phase involves **finding ideas**. In this phase, business models that have already been successfully implemented are analyzed and supplemented with own ideas or adapted for a specific, individual application context.

Before the new business model idea can survive, it must be integrated into a holistic business model (see Figure 3: Who? - What? - How? - Value?). This is done integratively in the development, elaboration and evaluation phases. In the first instance, further **development and elaboration** of the generated business model idea will be carried out using different methods and workshop formats, before an **assessment** will be carried out in which consistency with the internal requirements as well as with the external corporate environment (influencing factors and actors) will be checked.

The business model development of the innovation laboratory is currently in these phases. The **lean start-up approach** of Eric Ries (cf. Ries 2017) is used to support and structure these phases. This approach can be used throughout the entire product development process, from **development** to **elaboration** and **implementation of the business model**. Individual prototypes in the form of so-called **Minimum Viable Products (MVP)** are used to test and evaluate hypotheses on customer requirements and business model approaches, and to take the findings gained directly into account in the next development stages.

An MVP is the first minimally functional iteration of a product that uncovers customer requirements with minimal effort and derives measures for future development work. MVPs are developed according to the **Build-Measure-Learn principle**, which focuses on fast testing, measuring and agile adaptation. In contrast to classical prototypes, MVPs focus on the validation of hypotheses regarding customer requirements (external), while classical prototypes are primarily used by the development team (internal) to carry out tests on product features (cf. Ries, 2017, p. 76 f.). As part of product development in the innovation laboratory, functionality and market potential are tested in the first **showcases** in the form of MVPs.

On the basis of the test results, there are **three options for action for the business model**: (i) **maintaining** the current business model if the hypotheses are correct; (ii) **adapting** the business model while maintaining the business model elements confirmed by the tests and adapting the others; (iii) **rejecting** the business idea. The process is repeated until all key hypotheses have been confirmed or validated by MVP testing, resulting in "product-market fitness" as the successful completion of the build measure learn loop. (cf. Ries 2017, p. 76 ff.; Ghezzi/Cavallo 2018, p. 7).

The **Business Model Canvas (BMC)** from Osterwalder/Pigneur is used for the structured development and documentation of the integratively developed business model approaches for the showcases. The BMC is a widely used and proven method in practice, which is also used for the development of business models for smart PSS due to its high application orientation and practical relevance (cf. Michalik et al. 2018a, p. 311). Figure 4 schematically shows the BMC at the center of the Build-Measure-Learn loop. The central positioning illustrates that the business model or individual business model dimensions must be reviewed after each MVP test cycle and maintained, adjusted or rejected depending on the test result.

The **BMC** focuses on a component-based view of a business model. **Nine partial models** are considered for the holistic description of business models (see Osterwalder/Pigneur 2011, p. 19). The central component is the **value proposition** to the customer in the middle of the BMC. The right-hand side of the BMC focuses on the **customer perspective**, examining the customer relationships, the sales and communication channels used and the customer groups (customer segments) to which the value proposition is to be submitted. A continuous comparison between the value offering and the target customer segment is important in order to ensure the best possible fit between customer needs and benefit promises. The left side looks at the **internal prerequisites** that must be created to generate the value proposition. It identifies key partnerships, activities and resources required to establish customer relationships, establish sales and communication channels and generate value. The lower part of the BMC shows the **cost structure and the revenue streams**. This includes the type and structure of the costs incurred in implementing the business model, as well as the revenues expected from the value proposition from the respective customer segments. Detailed information on the BMC can be found in Osterwalder/Pigneur 2011.

In addition to the BMC, the **St. Gallen Business Model Navigator** is used as part of the process model (see Figure 4). The basic idea of the Business Model Navigator is that a business model does not have to be completely redeveloped, but that by **recombining** and creatively adapting elements of **existing business models**, it can be transferred to other industries or branches of industry. Numerous studies and consulting projects have shown that 90% of all new business models are not really new, but are based on 55 existing patterns (cf. Gassmann et al. 2013, p. 17).

A **business model pattern** is a specific configuration of the four business model dimensions (see Figure 3) that has proven successful in different companies and industries. An analysis of existing business models shows that the innovation leap from one sector to another is relatively small, whereby it is not a question of pure imitation, but rather of a meaningful combination of business model patterns. Business model patterns thus help to integrate successful patterns into one's own business model. In the process model, these patterns serve as **reference patterns** that are used to generate ideas and sharpen one's own business idea. A detailed explanation of all 55 business model patterns can be found in Gassmann et al. 2013.

Once the development of the new business model has been completed, it is implemented in the sixth phase of the process model, the **implementation phase**. In this phase, a business model innovation should not be implemented in a roll-out on the market. Rather, prototypes must first be developed and tested in order to limit risk and thus create the opportunity for further learning and adaptation.

APPLICATION OF THE LEAN APPROACH MODEL FOR BUSINESS MODEL DEVELOPMENT FOR SMART PRODUCT SERVICE SYSTEMS

The application of the developed procedure model and the integrated methods Lean-Start-Up, BMC and St. Gallen business model pattern is exemplary for a maintenance showcase from the innovation laboratory. The subsequent validation of the process model is based on a comparison with a process model already used in practice. The results obtained are presented below.

Application of the Procedure Model within the Framework of the Innovation Laboratory

Germany is the world's leading exporter of machinery and equipment (see Wiechers 2018). In order to maintain this lead in the future, German industry must also develop into the leading **supplier of smart PSS in the field of maintenance**. The use of sensors in maintenance makes it possible to achieve greater transparency regarding the condition of the machines, to record their continuous maintenance requirements and to identify the triggers of malfunctions. Novel information and knowledge networks between operators, manufacturers and service providers are created, which contribute to increasing maintenance efficiency by exchanging knowledge.

The **maintenance showcase of the innovation laboratory** focuses on the integration of new technologies and the resulting development of smart PSS. The **objectives of the maintenance showcase** are to reduce downtime, shift from unplanned to planned downtime, and ultimately increase system availability with the help of remote services in the form of **smart PSS**.

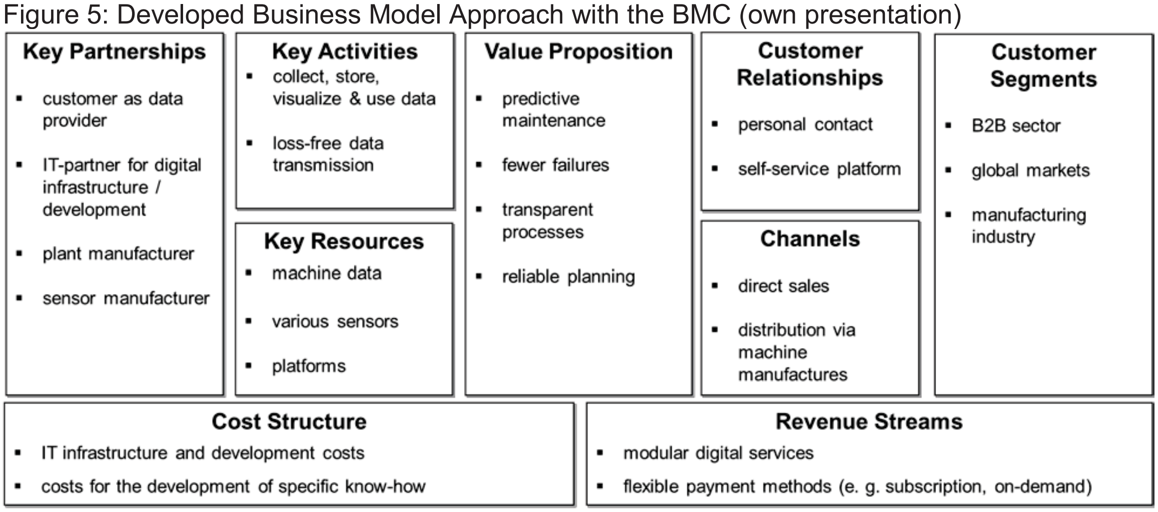
The following showcase is considered when **applying the Lean Procedure Model to business model development**: A fault message from an installed machine initiates the use case. The basis for the evaluation of the current state of the plant is a comprehensive sensor system. Based on the analysis of raw sensor data, a malfunction report is sent to the maintenance technician via a mobile terminal device, which at the same time supports fault identification by allowing fast and precise on-site analysis based on available information, e.g. from the social network. The repair and subsequent recommissioning of the affected machine is also supported and documented by mobile terminals, if necessary with the aid of augmented reality glasses.

On the **software side**, the analysis of sensor raw data with regard to potentially imminent machine faults is carried out on the basis of pattern recognition procedures. In terms of machine learning, algorithms must be trained in such a way that a deviation from the state of health is immediately detected. In addition to the software-based determination of fault messages, the augmented reality-based instructions for the maintenance engineer are based on an assistance system to be provided accordingly. As a **service**, the software used is designed, the data collected and processed and the necessary interfaces created. Smart PSS can thus make relevant information available in a context-sensitive manner, make a considerable contribution to the work design of sustainable maintenance and provide the hardware and software basis for **innovative business models**.

In order to develop the **business model for this maintenance showcase**, the smart PSS will be developed from the perspective of a **machine manufacturer** who wants to expand its target group, in particular by creating new value propositions, and differentiate itself on the market with the help of service-oriented approaches. To design an initial business model approach, the **BMC** is first used in the **development phase of the process model** (see Figure 4). The application of the BMC takes place within the framework of the innovation laboratory in the form of business model workshops. The aim of the workshops is to generate ideas and design the first business model, and at the same time to identify **deviations between value proposition and customer needs**. Identified gaps can be closed in the further product and service development process. In this phase, no implementable business model is generated. Rather, in the sense of **lean startup**, the current product development status is determined from the customer's perspective (*Does the product create benefits for the customer? Are product features still missing?*) and validated regarding internally required resources (*Am I in the position to produce and market it? Which resources and partners do I need for the next development steps?*).

The **value proposition** identified when applying the BMC is therefore based on the reduction of unplanned downtimes through knowledge of impending machine failures (predictive maintenance). This results in cost savings as well as the ability to plan necessary maintenance measures. The **key activities** essential for implementation are primarily based on data acquisition and analysis as well as subsequent reporting of potential anomalies. The basis for this is software and platform development on the one hand (**key resources**) and the establishment and maintenance of contacts to partners on the other, including the development of the necessary IT infrastructure and interfaces (**key partners**). A **predictive maintenance approach** can only be realized by ensuring an extensive and valid database. Figure 5 shows all elements of the business model approach developed in the workshop using the BMC.

Figure 5: Developed Business Model Approach with the BMC (own presentation)



In addition to the BMC, the **55 patterns of the St. Gallen Business Model Navigator are used as reference patterns**. Thus, following an initial design of the nine partial models of the BMC, intersections with existing business model patterns can be uncovered and suggestions for further design possibilities of the own business model can be generated. To this end, the 55 patterns were initially reduced by the workshop participants to five most appropriately rated patterns. The core of the business model, the value proposition and essential resources and skills were always taken into account in the selection process. The **patterns selected** during the workshops are (cf. Gass-mann et al. 2013, p. 72ff.): **Leverage customer data** with a focus on the acquisition and evaluation of machine data, **pay-per-use** as a billing model for each service used on the machine, **performance-based contracting**, which provides for a performance-based billing model (lower production and maintenance costs through predictive maintenance lead to improved machine performance), sales of additional services (predictive maintenance) as **add-on** to the machine, and the pattern of **self-service**, which is characterized by the transfer of maintenance work to the customer through the use of augmented reality-based instructions.

At its core, the business model of the use case thus represents a new combination of existing **business model patterns**. In particular, a multi-page use of customer data (**leverage customer data**) can be highlighted here. **Data as an economic good** can thus be of monetary value and can be prepared and used specifically for value-adding activities.

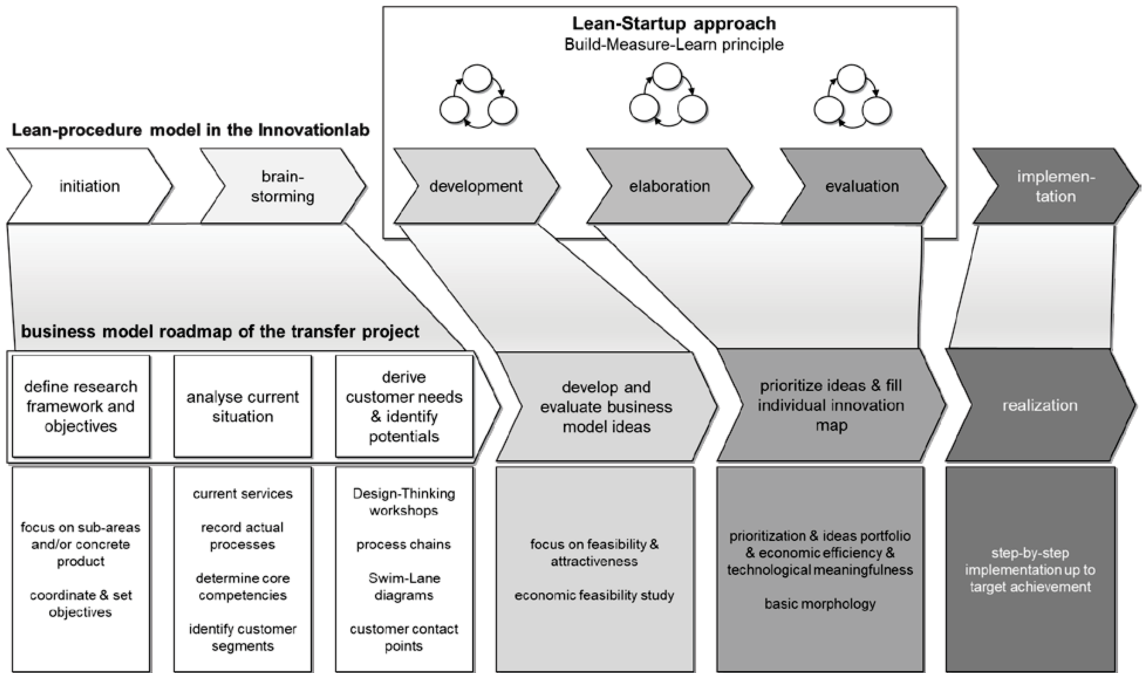
In contrast to other procedural models, the lean procedure model provides for a **continuous questioning and further development of the value proposition integratively to product development**. This is achieved on the one hand by a structured formulation of the business idea, here using the BMC, and on the other hand by comparison with existing business model patterns along the development, elaboration and evaluation phases. The lean startup approach ensures continuous development, measurement and learning through **build measure learn loops**, so that the focus is on market- and customer-oriented product development and business orientation.

Validation of the Procedure Model on the Basis of a Transfer Project

For the supplementary **validation of the lean procedure model**, this contribution uses the results of a transfer project, which was also carried out within the framework of the innovation laboratory in cooperation with a medium-sized plant manufacturer.

The goal of the six-month transfer project was the development of technologically validated, practical business model concepts for industrial services. The results led to a **roadmap** that supports the development of business models in **six phases** and proposes options for action and methods (cf. Michalik et al. 2018b, p. 325 ff.). The steps of this roadmap are comparable with the procedure model used in the innovation laboratory, based on Gassmann et al., in that both basically depict a six-phase process, which ranges from initiation to brainstorming and implementation (cf. Figure 6).

Figure 6: Lean approach compared to the roadmap of the transfer project



The roadmap was developed in cooperation with a medium-sized company from the machine and plant engineering sector. Different **methods** are used in the individual phases, for example creativity methods such as design thinking or business model sample cards, and workshop formats such as process workshops for the creation of swim lane diagrams. The **business model concepts** developed in the workshops were then evaluated and prioritized with regard to customer added value, available resources and competencies. The implementation as the last phase of the roadmap is still pending after the previous concrete elaboration and prototypical implementation. The business model roadmap developed in the transfer project thus offers a guideline for action, which in particular provides an approach for the first **design of a business model strategy for smart PSS**.

An essential **core result of the transfer project** is the realization that the current business model or the current development status of the business model must always be questioned in every phase of the roadmap. The options for maintaining, adapting or distorting the current business model (cf. section 3.4) are examined. The roadmap takes these into account and addresses them in design thinking workshops. However, there is a lack of an integrative perspective that integrates such a questioning into the product development process in a structured way. Without an integrative approach, there is a danger that information from product development will not flow into business model development or that business model development will only be initiated very late in the product development process or even after the development of an initial prototype. It can therefore be seen that iterative loops in the sense of the build measure learn principle are unavoidable for customer-oriented business model innovations.

The lean approach presented here focuses in particular on the creation of MVPs within the framework of the development and elaboration phases and contributes the iterative evaluation. A comparison of both approaches shows that Lean-Startup as well as the roadmap validated in practice use comparable phases with corresponding methods. A **fundamental applicability of the Lean Procedure Model** can thus be regarded as given. In particular, the use of the similarity principle, such as the 55 business model patterns of the St. Gallen Business Model Navigator, makes an important contribution in the early development phase.

Beyond the business model roadmap, the lean startup approach represents a meaningful further development at this point by focusing on **prototype implementation** and its constant comparison with customer needs. As soon as a business model concept has undergone the first development and evaluation steps, the actual market potential and feasibility are checked in the context of a prototype implementation in cooperation with a customer. In accordance with a lean open innovation principle, this enables companies to further develop business models in a structured and resource-efficient manner and to develop a sound basis for evaluation for potential market launches.

SUMMARY AND OUTLOOK

As part of the "Innovation Laboratory Hybrid Services in Logistics", various showcases are being developed to advance the design of smart PSS. The aim is to combine state-of-the-art hardware and software with innovative business models and thus generate **smart PSS for the B2B sector**. The process model presented in this article, which integrates the lean startup approach into business model development, makes a valuable contribution to this.

The results from the conducted business model workshops show the potential, which is opened by the meaningful combination of different methods to a lean procedure model. In this article, the approach is exemplified using a **showcase from maintenance** and validated by a comparison with a transfer project with a medium-sized company from the mechanical and plant engineering sector. When applying the approach in the form of workshops, the focus is on a continuous questioning of the value proposition through new value promises, a (new) combination of business model patterns as well as iterative build measure learn loops. Companies that use this approach not only to question their existing service portfolio, but also to identify and develop completely new service offerings and business models, are enabled to expand their target group and differentiate themselves on the market.

The "Innovation Laboratory Hybrid Services in Logistics" was extended for a further one and a half years until 2020 with the approval and support of the Federal Ministry of Education and Research. Within this framework, **further practical testing** of the lean procedure model is being promoted. Future research efforts will focus on the efficient **identification and evaluation** of (corporate) resources that are required for the market-ready development of business model ideas. On the other hand, **empirical studies** are aimed at that validate the effectiveness of the approach presented here and at the same time identify further research needs.

This publication was produced as part of the research project "Innovation Laboratory Hybrid Services in Logistics" in Dortmund, which is funded by the Federal Ministry of Education and Research (BMBF).

REFERENCES

Abramovici, M.: Engineering smarter Produkte und Services- Plattform Industrie 4.0 STUDIE, URL: https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/hm-2018-fb-smart.pdf?__blob=publicationFile&v=3, published on 01.01.2018, accessed on 25.04.2019

Abramovici, M., Gebus, P., Herzog, O.: Engineering im Umfeld von Industrie 4.0- acatech Studie, Herbert Utz-Verlag, München 2016

Becker, W., Ulrich, P., Ebner, R., Zimmermann, L.: Erfolgsfaktoren der Geschäftsmodelle junger Unternehmen, Bamberger Betriebswirtschaftliche Beiträge 183, 2012

Bilgeri, D., Brandt, V., Lang, M., Tesch, J.; Weinberger, M.: The IoT Business Model Builder; URL: http://www.iot-lab.ch/wp-content/uploads/2015/10/Whitepaper_IoT-Business-Model-Builder.pdf, published on 01.10.2015, accessed on 25.04.2019

Boßlau, M., Gesing, J., Meier, H., Wieseke, J.: Geschäftsmodelle für Industrielle Produkt-Service Systeme, in: Meier, H.; Uhlmann, E. (Hrsg.), Industrielle Produkt-Service Systeme. Entwicklung, Betrieb und Management, Springer Vieweg, Wiesbaden 2017, pp. 299-324

Braun, E.: Vom klassischen Produzenten zum Lösungsanbieter - Ein Ansatz zur Umsetzung von Kundenlösungen im High-End Konsumgüter Bereich, URL: https://www1.unisg.ch/www/edis.nsf/SysLkpByIdentifier/4069/$FILE/dis4069.pdf, published on 29.10.2013, accessed on 25.04.2019

Bruhn M., Hadwich K.: Servicetransformation – Eine Einführung in die theoretischen und praktischen Problemstellungen. In: Bruhn M., Hadwich K. (Hrsg.) Servicetransformation. Springer Gabler, Wiesbaden 2016, pp. 3-22

Bucherer, E.: Business model innovation: guidelines for a structured approach, Shaker Verlag, Aachen 2010

Bühler, P., Maas, P.: Transformation von Geschäftsmodellen in einer digitalisierten Welt, in Bruhn, M.; Hartwich, K. (Hrsg.): Dienstleistungen 4.0 - Geschäftsmodelle – Wertschöpfung – Transformation, Band 2: Forum Dienstleistungsmanagement, Springer Gabler, Wiesbaden 2017, pp. 43-70.

Chesbrough, H.: Business model innovation: it´s not just about technology anymore, in Journal of strategy and leadership, Vol. 35, Nr. 6, 2007, pp. 12-17.

Chowdhury, S., Haftor, D., Pashkevich, N.: Smart Product-Service Systems (Smart PSS) in Industrial Firms: A Literature Review, in: Procedia CIRP 73, 2018, pp. 26-31

Drossel, W.-G., Ihlenfeldt, S., Langer, T., Dumitrescu, R.: Cyber Phsysische Systeme – Forschen für die digitale Fabrik. In: Neugebauer, R. (Hrsg.), Digitalisierung – Schlüsseltechnologien für Wirtschaft & Gesellschaft, Springer Vieweg. Berlin Heidelberg 2018, pp. 197-222

Fujifilm: Innovationsgeschichte, URL: https://www.fujifilm.eu/de/ueber-uns/geschichte/innovationsgeschichte#content7c, published on n.A., accessed on: 31.01.2019.

Gartner: Pressemitteilung Februar 2017, URL: https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016, published on 07.02.2017, accessed on 25.04.2019

Gassmann, O.; Frankenberg, K.; Csik, M.: Geschäftsmodelle entwickeln - 55 innovative Konzepte mit dem St. Galler Business Model Navigator, Hanser, Regensburg 2013

Gausemeier, J.; Klocke, F.: Industrie 4.0 – Internationaler Benchmark, Zukunftsoptionen und Handlungsempfehlungen für die Produktionsforschung. Heinz-Nixdorf-Institut (Universität Paderborn), WZL (RWTH Aachen), acatech, Paderborn Aachen München 2016

Gausemeier, J.; Wieseke, J. (2017): Gemini 4.0: Mit Industrie 4.0 zum Unternehmenserfolg. Integrative Planung von Geschäftsmodellen und Wertschöpfungssystemen, URL: https://www.hni.uni-paderborn.de/fileadmin/Fachgruppen/Seniorprofessur_Gausemeier/gemini/GEMINI_Broschuere_web.pdf, URL: https://www.hni.uni-paderborn.de/fileadmin/Fachgruppen/Seniorprofessur_Gausemeier/gemini/GEMINI_Broschuere_web.pdf, published on 01.04.2017, accessed on 25.04.2019

Geisberger, E.; Broy, M.: agenda CPS - Integrierte Forschungsagenda, Cyber-Physical Systems (acatech Studie), Springer, Berlin Heidelberg 2012

Ghezzi, A., Cavallo, A.: Agile Business Model Innovation in Digital Entrepreneurship: Lean Startup Approaches,in: Journal of Business Research, 2018, pp. 3-50

Giesen, E., Berman, S., Bell, R. und Blitz, A.: Three ways to successfully innovate your business model,in: Strategy and Leadership, Vol. 35 Nr. 6 ,2007, pp. 27-33

Gorldt, C., Wiesner, S., Westphal, I., Thoben, K.-D.: Product-Service Systems im Zeitalter von Industrie 4.0 in Produktion und Logistik – Auf dem Weg zu Cyber-Physischen Product-Service Systemen, in Bruhn, M.; Hadwich, K. (Hrsg.): Dienstleistungen 4.0. Geschäftsmodelle – Wertschöpfung – Transformation, Band 2, Forum Dienstleistungsmanagement, Springer Gabler, Wiesbaden 2017, pp. 363-380

Grandjean, L.; Riss, E.; Steven, M.: Geschäftsmodelltypologie für hybride Leistungsbündel - Einsatz des Business Model Canvas, in: Wirtschaftswissenschaftliches Studium 46, Heft 2, 2017, pp. 7-11

Henke, M.; Hegmanns, T.: Geschäftsmodelle für die Logistik 4.0: Herausforderungen und Handlunsgfelder einer Grundlegenden Transformation, in: Vogel-Heuser, B.; Bauernhansl, T.; ten Hompel, M. (Hrsg.): Handbuch Industrie 4.0, Bnd. 3, Springer Vieweg, Wiesbaden 2017, pp. 335-345

Kagerman, H., Wahlster, W., Helbig, J.: Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0, URL: https://www.bmbf.de/files/Umsetzungsempfehlungen_Industrie4_0.pdf. published on 01.04.2013, accessed on 25.04.2019

Labbé, M., Mazet, T.: Die Geschäftsmodellinnovations-Matrix: Geschäftsmodellinnovationen analysieren und bewerten, in: Der Betrieb 17, 2005, pp. 897–902.

Lindgardt, Z., Reeves, M., Stalk, G., Deimler, M. (2009): Business Model Innovation: when the game gets tough change the game, URL:

https://www.bcg.com/documents/file36456.pdf, published on 01.12.2009, accessed on 25.04.2019

McKinsey: Industry 4.0 – How to navigate digitization of the manufacturing sector; URL: https://www.mckinsey.com/business-functions/operations/our-insights/industry-four-point-o-how-to-navigae-the-digitization-of-the-manufacturing-sector

Meier, H.; Uhlmann, E.: Hybride Leistungsbündel – ein neues Produktverständnis, in: Meier, H.; Uhlmann (Hrsg.), Integrierte industrielle Sach- und Dienstleistungen, Springer, Berlin Heidelberg 2012, pp. 1–21

Meier, H.; Uhlmann, E.; Kortmann, D: Hybride Leistungsbündel – Nutzenorientiertes Produktverständnis durch interferierende Sach- und Dienstleistungen, in: wt Werkstattstechnik online 95, 2005, pp. 528–532

Michalik, A., Möller, F., Henke, M., Otto, B.: Towards utilizing Customer Data for Business Model Innovation: The Case of a German Manufacturer, in: Procedia CIRP 73, 2018a, pp. 310-316

Michalik, A., Förster, F., Besenfelder, C.: Datenbasierte Geschäftsmodelle - Weshalb es dabei nicht um noch mehr Daten geht, In: 35. Deutscher Logistik-Kongress 2018b, Kongressband, Bundesvereinigung Logistik (BVL), pp. 325-344

Mitchell, D., Coles, C.: Business model innovation breakthrough moves, in: Journal of Business Strategy, Vol. 25, No. 1, 2004, pp. 16-26

Mitchell, D.; Coles, C.: The ultimate competitive advantage of continuing business model innovation, in: Journal of Business Strategy, Vol. 24, No. 5, 2003, pp. 15-12

Orngreen R, Levinsen KT. Workshops as a Research Methodology. Electronic Journal of E-Learning 15; 2017. 70–81.

Osterwalder, A., Pigneur, Y.: Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer, campus, Frankfurt New York 2011

Osterwalder, A.; Pigneur, Y.; Tucci, C. L.: Clarifying Business Models. Origins, Present and Future of the Concept, in: Communications of the Association for Information Systems, Vol. 16, No. 1, 2005, pp. 1–25.

Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24; 2007. 45–77.

Porter, M.E.; Heppelmann, J.E.: How smart, connected products are transforming completion, in: Harvard Business Review, Vol. 92, No. 11, 2017, pp. 64-88

Ries, E.: The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses, Crown Business, 2017

Schallmo, D.: Geschäftsmodell-Innovationen. Grundlagen, bestehende Ansätze, methodisches Vorgehen und B2B-Geschäftsmodelle, Springer, Wiesbaden 2013

Schallmo, D.: Theoretische Grundlagen der Geschäftsmodell-Innovation – Definitionen, Ansätze, Beschreibungsraster und Leitfragen, in: Schallmo, D. (Hrsg.): Kompendium Geschäftsmodell-Innovation – Grundlagen, aktuelle Ansätze und Fallbeispiele zur erfolgreichen Geschäftsmodell-Innovation, Springer, Wiesbaden, 2014

Schultz, S.: Kodak-Pleite Geisel verblasster Erfolge; URL: http://www.spiegel.de/wirtschaft/unternehmen/kodak-pleite-geisel-verblasster-erfolge-a-810016.html, published on: 19.01.2012, accessed on: 31.01.2019.

Stähler, P.: Geschäftsmodelle in der digitalen Ökonomie – Merkmale, Strategien und Auswirkungen, St. Gallen 2001

Steven, M.: Produktionscontrolling, Kohlhammer, Stuttgart 2016

Steven, M.: Industrie 4.0 Grundlagen-Teilbereiche-Perspektiven, Kohlhammer, Stuttgart 2019

Teece, D.: Business Models, Business Strategy and Innovation, in: Long Range Planning, Vol. 43, No. 2-3, 2010, pp. 172-194.

ten Hompel, M.; Putz, M.; Nettsträter, A.: Whitepaper »Social Networked Industry« – Für ein positives Zukunftsbild von Industrie 4.0, URL: https://www.e3-produktion.de/content/dam/iwu/e3-produktion/de/documents/E3_Whitepaper_SocialNetworkedIndustry_Web.pdf, veröffentlicht: 2017a, accessed on 25.04.2019

ten Hompel, M; Tüllmann, C., Nettsträter, A, Prasse, C.: Whitepaper Social Networked Industry ganzheitlich gestalten, URL: https://www.iml.fraunhofer.de/content/dam/iml/de/documents/101/06_Whitepaper_SNI-ganzheitlich-entwickeln_WEB.pdf, published on 05.09.2017b, accessed on 25.04.2019

Valencia, A, Mugge, R, Schoormans,J: P. L.,Schifferstein, H. N. J: The Design of Smart Product-Service Systems (PSSs): An Exploration of Design Characteristics, in: International Journal of Design., Vol. 9, No. 1, 2015, pp. 13–28

Wiechers, R.: Maschinenbau macht weiter Tempo im Export, URL: https://www.vdma.org/v2viewer/-/v2article/render/27115762, published on: 19.11.2018, accessed on: 13.05.2019.

Wiesner, S., Gorldt, C., Soeken, M., Thoben, K.-D., Drechsler, R.: Requirements Engineering for Cyber-Physical Systems, in: Grabot, B., Vallespir, B., Gomes, S., Bouras, A., Kiritsis D. (Hrsg.): Advances in Production Management Systems - Innovative and Knowledge-Based Production Management in a Global-Local World, Springer, Berlin Heidelberg 2014, pp. 281-288

Wirtz, B.: Business Model Management – Design – Instrumente – Erfolgsfaktoren von Geschäftsmodellen, Gabler Verlag, Wiesbaden .

Managerial Decisions in Your Firm

DECISION SCIENCES INSTITUTE
A Time Series Analysis Approach to Forecast Price Fluctuation Points

Ching-Chin Chern
National Taiwan University
Email: cchern@ntu.edu.tw

Wei-Yun Chen
National Taiwan University
Email: r03725047@ntu.edu.tw

Bo Hsiao
Chang Jung Christian University
Email: bhsiao@mail.cjcu.edu.tw

ABSTRACT

This study proposes a novel price fluctuation forecast model, Price Fluctuation Point Forecast Approach (PFPFA), to forecast the price fluctuation points and the degrees of price changes. PFPFA proposes four models predict the price fluctuation points and three models to predict the degrees of price changes. In consequence, for a single product, there would be twelve different forecast outcomes. PFPFA is applied on a real world case and compared with the Exponential Smoothing (ES). In the end, PFPFA predict acceptable result on the price fluctuation points and perform better on the degrees of price changes than ES.

KEYWORDS: Price Fluctuation Pattern, Non-uniform Time Series Data, Time Series Analysis, Short Product Life Cycle Product, Price Forecast

INTRODUCTION

For a manufacturer, the material procurement and product sales directly influence the cash flows. Buying material at a lower price and selling products at a higher price are the straightest way to obtain higher revenue. The life cycles of consumer products such as apparel, cosmetic and consumer electronics are getting shorter and shorter because of consumer preference shifts. The demands of these products usually resemble a diffusion process: rapid growth at the beginning and rapid decline at the end of a life cycle while stable in between (Li and Huh, 2012). If there is a way to predict the price fluctuation of material or products accurately, a company can maximize its profit by taking a right action at a right time. This study aims at modifying the existing price estimation models to develop new sourcing policies that can be used under the dynamic environment.

As the technology of producing a product getting mature, the cost of the product decreases, and the demand of the product may drop because a newer product with better functions may appear on the market. In order to eliminate the cost of obsolescence and attract more consumers, the prices of short life cycle products usually are highest when introduced to the market but are cascading down through the entire life cycle (Kurawarwala and Matsuo, 1996). Mostly, the time plot of prices falls in a series of cascades down towards the end of life cycle.

A company can benefit from forecast of price fluctuation either as roles of sellers or as roles of buyers. As a seller, it is crucial to decide when to cut the price and how large a discount to offer. The sensitivities of customers toward prices and demands are varying at different stages of a product's life cycle. Hence, it is crucial to charge different prices at different stages of a product's life cycle to maximize revenue (Li and Huh, 2012). In addition, providing a right discount at right time will increase customer satisfaction, avoid bargaining, and strengthen partnerships. Therefore, analyzing the price strategy from historical data can help managers to making right decisions about price cutting and discount offering for sellers.

On the other hand, as a buyer, it is very important to know the time epochs and the degrees of price fluctuations of a product in order to determine when is the best time to procure material and to review the performance of a supplier. When accumulated purchasing quantity of a material approaches a certain level, suppliers may offer price discount to boost up purchase while buyers can negotiate for discounts to reduce their purchase costs. Furthermore, when the number of products and materials increases and the patterns of price changes become complicated, it is essential to adopt data mining methods to automatically validate that material are purchased at reasonable prices and at the right times.

By analyzing the historical sales or purchase data, we can find the relationship between price change and sales/purchase quantity and build a forecast model to inspect the timing and degree to offer discount. In addition, if a supplier with relatively weak bargaining power fails to offer discounts when it should, buyers can and should initiate the process of negotiation with the supplier for a better price. Procurement at a lower price reduces the total cost, which leads to a higher profit, a better profit margin and a larger return-on-investment (ROI).

Previous researches have proposed many price forecast methods for electricity (Contreras et al., 2003), oil (Morana, 2001), gas (Buchanan et al., 2001) and stock (Baba and Kozaki, 1992). However, consumer products are very different from these commodity markets and stocks in terms of product life cycles and price fluctuation behaviors. Some products may change their prices only once or twice over their life cycles, while other change more times. Thus, time series analysis, a technology to discover the pattern from temporal data and to forecast the future, is a suitable tool to predict the price fluctuation points and the degrees of price changes (Esling and Agon, 2012). There are four components of a time series: the trend, the cyclical fluctuation, the seasonal fluctuation, and the irregular fluctuation (Bowerman and O'Connell, 1979). Based on the theory of product life cycle (Li and Huh, 2012), the demand of a product changes over time through its life cycle. Moreover, according to the law of demand, the price of a product should fluctuate to correspond with the demand for better profit. Therefore, in addition to those four components of a time series, demand variation must be considered when building the forecast model of price fluctuation.

To use demand quantity for price forecast, we need to understand the pattern of demand change through product life cycle. Although actual sales depend on both demand and supply, life cycles are widely used for modeling demand change (Polli and Cook, 1969). Product life cycle (PLC) theory is introduced to explain the different stages of a product from design to obsolescence. Researches showed that demand for most products follows a standard life cycle with five phases: introduction, growth, maturity, decline, and withdrawal (Polli and Cook, 1969).

Diffusion theory is an appropriate tool to model the PLC process (Qin and Nembhard, 2012) and are widely used for demand forecasting (Morrison, 1996). Diffusion Model assumes the demand quantity follows a non-linear logistics curve defined by three parameters: Long Run Saturation

Level, Inflection Point, and Delay Factor (Qin and Nembhard, 2012). If these three parameters are determined, Diffusion Model can be used to forecast the future demand. However, diffusion model represents only the demand of a product at growth and maturity stages and thus is inapplicable at introduction, decline and withdrawal stages. Every product has its own product life cycle curve and a company may have thousands of products. This study first classifies products into different groups whereas in group, the products have similar demand fluctuation behaviors. This study proposes to build a forecast model for each group rather than for each product.

To sum up, price fluctuation forecast is an urgent problem in many industries. Price fluctuation depends on the accumulation of sales/purchase volumes, which represents different stages of a product's life cycle. This study proposes a price fluctuation forecast model by adopting the time series analysis method based on the historic sales/purchase data and establish a classification model to automatically forecast the time and degree of price fluctuation when procuring material as buyers or offering price discount as sellers. The rest of this article is organized as follows: Section 2 describes the problem. In Section 3, we present our PFPFA to resolve the price fluctuation forecast problem efficiently. Section 4 compares the results obtained with our heuristic algorithm against other forecast methods to evaluate the PFPFA's efficiency and accuracy. Finally, in Section 5, we offer conclusions and suggestions for further research.

PROBLEM DESCRIPTION

In order to forecast the time and degree of price fluctuation when procuring material as a buyer or offering price discount as a seller, we first have to know how many times a product will change its prices during the life cycle. Normally, if a new material or product is expected to have similar behavior as the existing one, we can infer the result of a forecast model by using historical data of the existing one to this new material or product. Nevertheless, if there is no existing material or product for reference, managers must make a guess of the price trend based on experience and characteristic of the new material or product, which then lead to the most similar existing product as a role model for analysis and forecast.

After deciding the number of times for price changes, the next question is to determine the time of price fluctuation. Product price and demand quantity are both time series data. As mentioned above, price fluctuation depends on the accumulation of demand volumes at different stages of life cycle. Hence, we want to use the accumulating demand volume as a function of time, and apply time series analysis approaches to find the time of price fluctuation based on the corresponding accumulating demand quantity. For a seller, demand quantity is sales volume whereas for a buyer, demand quantity is purchase volume. Because we use demand as a function of time, the quantity will be accumulated through time.

For a large-scale manufacturer, buyers and sellers play important roles because their jobs influence the profitability of their company. A price fluctuation forecast model can advise sellers and buyers the right time to take appropriate action about product price change. Because the demand and price data are non-uniform time series, they need to be preprocessed before further analysis. Moreover, since the price and demand quantity are both time series data, we need to address and solve the issues related to the time series data. Finally, since this forecast model is novel, we need to propose an evaluation method to assess the forecast performance.

Most of the time series analyses focus on uniform sampled data that are regular recorded at uniform interval of time, e.g., weather records or sensory records. In real life, business transactions do not happen at uniform fixed interval of time, which poses a problem when applying time series analysis techniques. When data is non-uniform sampled, a time bucket is determined and used to convert the data into uniform time series data. Next, many time series analysis techniques, i.e., moving average or exponential smoothing, can be applied to these uniform time series data. These classical techniques focus on the using the historical data to forecast the future price or quantity. However, in real practice, transactions do not occur regularly and we cannot be sure the time when a transaction comes next. Furthermore, the price of a product usually drops consistently at the end of life cycle and time series analysis techniques, e.g., moving average or exponential smoothing, have the problem of delaying price change.

What managers really concern is the time and degree of price fluctuation and thus, using traditional time series analysis techniques cannot achieve this goal. As mentioned above, demand quantity differs at different stages of a product's life cycle. Therefore, using the demand quantity to analyze the price fluctuation points is more accurate rather than time. However, if the price or quantity of all products or material are used to forecast future demand, it will be very difficult to find the pattern of price fluctuation for each individual product or material. Consequently, we need to transform the price and quantity data into percentage to eliminate the effect of scale or unit, which helps us to group the items with similar behavior of price fluctuation.

After data pre-processing, we obtain two time series data. Define Q_{ij} as the demand quantity for transaction record j of product i and P_{ij} as the unit price for transaction record j of product i. Define OP_i as the original unite price for product i. For instance, assume that there are 50 transactions happened in the entire product life cycle of product A. The two time series data for product A, QA_j and PA_j, are listed as follows:

$$\text{Quantity of Product A} = \{Q_{A1}, Q_{A2}, \ldots, Q_{A50}\}$$

$$\text{Price of Product A} = \{P_{A1}, P_{A2}, \ldots, P_{A50}\}$$

This study uses quantity as a function of time to predict price change and thus a new forecast model is developed to predict the time of price fluctuation first followed by the degree of price fluctuation. To concentrate on the time of price fluctuates, we will remove those data points that the price does not change, and keep the data points with price fluctuations. Define the FQ_{ij} as the rate of accumulated demand quantity at price fluctuated data point j of product i and FP_{ij} as the price rate at data point j of product i. Define TQ_j as the total accumulated quantity at the end of price fluctuation of product i. Therefore, if a product changes prices five times in its entire product life cycle, the number of the time series data points to represent demand quantity and price will become 5.

$$\text{Rate of Accumulated Quantity of Product A} = \{FQ_{A1}, FQ_{A2}, \ldots, FQ_{A5}\}$$

$$\text{Rate of Fluctuation Price of Product A} = \{FP_{A1}, FP_{A2}, \ldots, FP_{A5}\}$$

To forecast price fluctuation points, we need know: (1) the number of price fluctuation points; (2) price fluctuation time; and (3) fluctuation price. Before forecasting price fluctuation points, we need to know the number of price fluctuation points from the historical data. As mentioned above, we forecast the demand quantity as the price fluctuation times. Define Num_F_i as the number of the price fluctuation points for product i. For example, we assume that product A has five price fluctuation points, which implies that $Num\_F_A = 5$. Traditional forecasting methods predict the times of these points (periods or dates). This study predicts the accumulated demand quantity level as the times of these points. Define $FARQ_{ik}$ as the fluctuation accumulated rate of quantity for product i and fluctuation record k. For instance, the first price fluctuation point of product A will occur when the demand quantity reached 15% of total demand quantity and thus, $FARQ_{A,1}$ = 15%.

After determining the time of price fluctuation, the degree of price fluctuation needs to be predicted. Define FPR_{ik} as the fluctuation price rate for product i and fluctuation record k. Thus, the price of product A drops to 95% of the original price at the first fluctuation point is noted as $FPR_{A,1}$ = 95%. That is, the price of product A will drop to 95% when the demand quantity reached 15% of total demand quantity.

To find the policy of ($FARQ_{ik}$, FPR_{ik}), we propose several forecasting methods for different kind of products to be chosen from. However, users need to find the best model for their products and employ the best model to forecast the new products. Hence this study constructs a classifier to divide products into different groups in which the products have similar demand fluctuation behaviors and build a forecasting model for each group.

The common error measurements are Mean Absolute Deviation (MAD), Mean Squared Error (MSE) and Mean absolute percentage error (MAPE). Mean absolute deviation (MAD) is the average of total errors from all time units. Mean Squared Error (MSE) is the average of total squared errors from all time units. Mean absolute percentage error (MAPE) is the average of total percentage errors from all time units. However, all these error measurements are for uniform time series data. To overcome this obstacle, we predict both the time and degree of price fluctuation and choose the same error measurement for both forecast results to evaluate the errors.

THE PFPFA

Negotiating lower price at the right time is very important for a manufacturer as well as reducing price at the right stages of a product life cycle is critical for a retailer. This study proposes a method called the Price Fluctuation Point Forecast Approach (PFPFA) to solve these problems effectively. We first introduce the process of PFPFA and transform the data to the format that is ready for forecasting. We then present the models for price fluctuation forecast. There are four models for time forecast and three models for price forecast. In consequence, for a single product, there would be twelve different forecast outcomes. Finally, to evaluate the forecast outcomes, we have to determine an assessment method. Hence, we could find the best combination of these forecasting models for each group.

The main process of PFPFA has four phases: (1) transforming data based on the number of fluctuation points; (2) calculating times with different forecast models; (3) calculating prices based on the results of phase 2 with different forecast models; and (4) evaluating and selecting the best forecast model combination for each group of products.

In phase 1, we transform the data into the format that can be used in phase 2 and phase 3. Because both the time and the degree that the price fluctuates are the main focus of this study, we propose to use demand quantity as the time to forecast the price fluctuation point. Thus, in phase 2, the accumulated rates of quantity are treated as time epochs of price fluctuation points and can be forecasted by four methods. PFPFA also proposes three methods to forecast the fluctuation price rate based on the accumulated rate of quantity derived before in phase 3. Therefore, for each product, there are totally 12 models to perform forecast. Finally, in phase 4, PFPFA chooses the best price forecast method with the smallest MAD for each group.

P1: Transform Data

We want to use the accumulated sales/purchase quantity to represent the time of price fluctuation. Thus, the original data of procurement, including the demand quantity and unit price of each order on different dates, need to be transformed to the accumulated quantities and price change percentages. The first step is to calculate the accumulated rate of quantity at data point j of product i (FQ_{ij}) which is computed as the accumulated quantity up to data point j of product i (AQ_{ij}) divided by the total quantity of product i (TQ_i).

Next, we need to calculate the price rate at data point j of product i, FP_{ij}, which is equal to the unit price (P_{ij}) divided by the original unit price (OP_i). By definition, OP_i is the unit price at first transaction. We also need to calculate the price change rate at data point j of product i (PC_{ij}) by subtracting P_{ij-1} from P_{ij}.

Since we only concern about the time that the price fluctuated, the final step is to filter out those transactions with price change equal to 0 or $PC_{ij} = 0$ and only keep the fluctuation with $PC_{ij} \neq 0$. Furthermore, to avoid confusing, we renamed the FQ_{ij} to $FARQ_{ik}$ and FP_{ij} to FPR_{ik}.

P2: Time Forecast

In this section, we introduce four time forecast models in PFPFA: (1) the linear fluctuation cycle model (LFC); (2) the exponential model (EXQ); (3) the logarithm model (LOGQ); and (4) the diffusion model (DFQ).

First, the linear fluctuation cycle model (LFC) assumes that the price of product i fluctuates every time when the accumulated quantity increased to a certain level. Therefore, the formula, $FARQ_{ik} = FF_i + FC_i \times FID_{ik}$, can be used to predict the fluctuation point k where FF_i is the first time of the price fluctuated and FC is the frequency of price fluctuation of product i. To find the best fit prediction model, we can use the simple regression technique. Assume that the relationship between $FARQ_{ik}$ and FID_{ik} is linear.

Second, the exponential model (EXQ) assumes that the prices of some products fluctuate more frequently at the beginning of their life cycles and getting less frequently toward the end of their life cycles. For this kind of products, we can use an exponential model to forecast the price fluctuation times, $FARQ_{ik} = b \times a^{FID_{ik}}$. To get the best fit prediction for these two parameters, b and a, we can use the simple regression technique. We can take logarithm of the equation and transform the equation to linear.

Third, opposite to the previous exponential model, the logarithm model (LOGQ) assumes that the prices of some products seldom change at the beginning of their life cycles and increase frequently at the end of their life cycles. Hence, we can use a logarithmic model to predict their fluctuation points as $FARQ_{ik} = b + a^* \ln FID_{ik}$. To get the best fit prediction for these two parameters, b and a, we can use the simple regression technique.

Finally, the diffusion model (DFQ) assumes that the prices of some products change frequently at the beginning and the end of their life cycles but keep steady in the middle period of their life cycles. In this case, we can use diffusion model to do the forecast.

$$FARQ_{ik} = \frac{1}{1 + e^{Di\,(Ii \times Num\_F_i - FID_{ik})}} - M_i$$

This model need three parameters, Inflection Point (I_i), Delay factor (D_i), and Middle Point (M_i). Ii is the inflection period when the price fluctuation frequency changes, Di controls the time when the frequency of price fluctuation starts to slow down, and M_i is the middle point or the average value of the accumulated quantity. To derive Ii, we calculate the slopes of $FARQ_{ik}$ every three points, find the most oblique period, and use the middle point of this oblique period divided by Num_F_i as I_i. We then use the deviation of 50% and the average of $FARQ_{ik}$, to calculate the value of M_i as $M_i = 50\% - (\Sigma \text{all } k FARQ_{ik}/Num\_F_i)$. To derive D_i, we set a range of 0.005 to 3, and use computer to automatically choose the value by selecting the one with the minimal sum of errors in absolute value.

P3: Calculating Prices based on Different Forecast Models

The models used to predict prices in this subsection are similar as the models for price fluctuation time proposed in the previous subsection. However, our forecast target is to achieve the minimum error of price forecast. Therefore, one product can use DFQ model for time forecast but use EXP model for price forecast as long as the final error of price forecast is minimized.

First, the fluctuation velocity model for price (FV) assumes that the prices of some products drop at a constant velocity. Therefore, we can predict the FPR_{ik} by the following formula: $FPR_{ik} = MPi + Vi \times FARQ_{ik}$. There are two parameters, Velocity (V_i) and Maximum Price (MP_i), needed in the FV model where MP_i is the price rate at first fluctuated point for product i and V_i is the velocity of price dropping for product i. To get the best fit prediction, we can use the simple linear regression technique by assuming that the relationship between FPR_{ik} and $FARQ_{ik}$ is linear where V is the slope and MP is the intercept.

Second, the exponential model for price (EXP) fits for those products whose prices drop fast at the beginning of their life cycles and become steady at the end of their life cycles. Assume that the relationship between FPR_{ik} and $FARQ_{ik}$ is exponential. We can predict the FPR_{ik} by the following formula: $FPR_{ik} = b \times a^{FARQ_{ik}}$. We then take logarithm of the equation and transform the equation to a linear equation.

Third, the logarithmic model for price (LOGP) is designed for those products whose prices drop slowly at the beginning of their life cycles and drop fast at the end of their life cycles. Assume the relationship between FPR_{ik} and $FARQ_{ik}$ is exponential. Therefore, we can predict the FPR_{ik} by the following logarithmic formula: $FPR_{ik} = b + a \times \ln(1.0001 - FARQ_{ik})$. To get the best fit prediction of these two parameters, b and a, we can use the simple regression technique.

P4: Evaluating and Selecting the Best Forecast Model

Our proposed forecast model involves four models for FARQ and three models for FPR for each product. In other words, when we want to perform a price fluctuation forecast for a single product, we have twelve different results. To evaluate these results, we have to choose an error assessment method. The most common error measurement methods are MAD, MSE and MAPE. Since the value of FPR is already in percentage, it would be inappropriate to use MAPE. In addition, using MSE will make the error become smaller if the data are in percentage. Consequently, we choose MAD as our evaluation method.

Because for each product i, the proposed forecast model not only predicts the future price rate by using model n but also predicts the time that the price fluctuates by using model m, two evaluation measurements are needed as follows:

$$Q\_MAD_{im} = \frac{\Sigma \text{all } k \, |\text{actual}(FARQ_{ik}) - \text{forecast}(FARQ_{ik})|}{Num\_F_i}$$

$$P\_MAD_{imn} = \frac{\Sigma \text{all } k \, |\text{actual}(FPR_{ik}) - \text{forecast}(FPR_{ik})|}{Num\_F_i}$$

For each product, we will run the 12 model combinations and choose the model combination which has the smallest P_MAD_{imn}, then use the parameters to forecast the price of new products. However, not all products have enough data points to perform 12 model combinations. Therefore, to train the system to select the best model for forecasting, all products are classified into several groups with similar price fluctuation patterns. Next, we randomly select 45% of the products in every group for training and the remaining 55% for testing. For each group of products, first we will run the 12 model combinations for each product. Second, we compute the average parameters needed for each model and the run these 12 model combinations again on the testing products in each group with these average parameters. Finally, choose the models with smallest P_MAD_{imn} and Q_MAD_{im} to forecast new products by using these average parameters.

COMPUTATIONAL ANALYSIS

In this section, we implement the PFPFA on a material purchase dataset from Company G, an international motherboard manufacturer in Taiwan, to demonstrate its feasibility. First, we introduce the dataset and divide it into a training set and a testing set. Second, we perform PFPFA on an example in the training dataset to show the result of each step. Third, we find the best forecast model for every group in the training dataset using PFPFA and then use these models to forecast the materials in the same group of the testing dataset. Finally, we interpret and analyze the testing results by comparing them with the models that Company G currently uses.

This dataset from Company G contains all the purchase transaction records within the five years from 2011 to 2015. There are 292,562 records for 17,511 different materials in the dataset. These materials are classified into 2,031 groups based on their price fluctuation patterns. Some materials have more than one supplier and therefore we mark the material numbers by adding their different supplier at the end. For example, MA5243 has two different suppliers, CVX and ASD, and hence is redefined by adding the supplier behind its material no as MA5243_CVX and MA5243_ASD. After we redefine the material numbers, there are 21,342 different materials in the dataset.

Because Company G has a weaker bargaining power than some suppliers do, the prices of 15,826 materials never change and the prices of 5,072 materials change less than 4 times, implying that there is no need to analyze these materials. Therefore, we filter out those materials and the remaining 444 materials in 55 groups are our target data for assessing PFPFA. To testing our PFPFA, we separate the remaining 444 materials into a training dataset and a testing dataset. We randomly select 45% of the materials in every group for training and the remaining 55% for testing.

For every group, PFPFA selects the model combination that best fits the training materials and uses the average of the parameters for the materials in the group in testing dataset. Next, PFPFA performs the forecast for the 55 groups in the testing dataset by the model selected for each group. This study also uses the Exponential Smoothing (ES), commonly used time series analysis and currently used by Company G, as the benchmark for comparisons. However, ES can only predict the degree of price change but not the time when price fluctuates. Therefore, the results of ES can be compared with the price of fluctuation points forecasted by PFPFA but for the time of fluctuation points, no comparison will be available. The smoothing index (α) of ES will be set as 0.9 after several trial and error testing. We use ES with $\alpha = 0.9$ to forecast for the materials in training set and then use the average of the forecast prices of every fluctuation points as the prediction to the materials in the testing set. We then calculate the average P_MAD of PFPFA and ES as 7.65% and 9.75%, respectively. Not only the price change, PFPFA can also predict the times of price change and the resulting average Q_MAD of all 55 groups is 13.45%. PFPFA works better than ES to predict the price changes in all 55 groups in terms of P_MAD.

To prove that PFPFA predict price change more accurate than ES, we first perform Chi-square normality tests on P_MAD of ES and P_MAD of PFPFA. Because p_values are 0.1227 and 0.0652, the null hypotheses are not rejected, both P_MAD of ES and P_MAD of PFPFA are normally distributed. We then perform the pair-t test on the difference between P_MAD of ES and P_MAD of PFPFA. Because t statistics = 6.5317 or p_value of the pair-t test is close to 0, the null hypothesis is rejected and the difference between P_MAD of ES and P_MAD is larger than 0. Based on the above results on the testing dataset, PFPFA performs better than ES. As to the time forecast of price change, since we have no other method to compare with, the result of PFPFA with an error of 13.45% is acceptable for the users to employ PFPFA in practice.

CONCLUSIONS

This study develops the Price Fluctuation Points Forecast Approach (PFPFA) by using the accumulated demand or sale quantity as the time of price fluctuation. PFPFA forecasts the quantity (time) by four quantity (time) forecast models, including a linear model, an exponential model, a logarithm model, and a diffusion model. These four models can represent different kind of relations between price fluctuation point and the accumulated demand or sale quantity. Based on the forecast result of quantity (time) forecast, PFPFA uses them to predict the product price at those fluctuation points by three price forecast models, including a linear model, an exponential model, and a logarithm model. These three models can represent different kind of relations between fluctuation price and the quantity. Therefore, PFPFA offers totally twelve model combinations for different types of products or material to fit. Once the model is selected and trained for a group of products or materials, PFPFA can predict the times and degrees of price fluctuation points for the existing or new products by using the average parameters of the selected best-fit model in the same group.

This study assumes that the number of price fluctuation points is known and can be used directly in training the model. However, in the real world dataset, the number of price fluctuation points is computed from the historical data. Hence, future works should be focused on the analysis of the number of price fluctuation points to enhance the forecast accuracy. Before PFPFA, there is no efficient way to forecast the price fluctuation points for Company G and some materials in the dataset from Company G may not reflect the real possible price. If Company G employs PFPFA for one or two years, they will push for better deals and thus, more price fluctuation points. Future works can observe the change and rebuild the models after applying PFPFA in practice for Company G.

ACKNOWLEDGEMENTS

This research was sponsored by the Ministry of Science and Technology in Taiwan, under project number MOST 107-2410-H-002-102.

REFERENCES

Baba, N. and Kozaki, M., “An Intelligent Forecasting System of Stock Price Using Neural Networks,” IEEE/ INNS International Joint Conference on Neural Networks, 1(1): 371-377, 1992.

Bowerman, B. L. and O’Connell, R. T., “Time Series and Forecasting: an Applied Approach,” N. Scituate, Mass: Duxbury Press, 1979.

Buchanan, W. K., P. Hodges, and Theis, J., “Which Way the Natural Gas Price: An Attempt to Predict the Direction of Natural Gas Spot Price Movements Using Trader Positions,” Energy Economics, 23(3): 279-293, 2001.

Contreras, J., Espínola, R., Nogales, F. J. and Conejo, A. J., “ARIMA Models to Predict Next-Day Electricity Prices,” IEEE Transactions on Power Systems, 18(3): 1014-1020, 2003.

Esling, P. and Agon, C., “Time-Series Data Mining,” ACM Computing Surveys, 45(1): 1-34, Nov. 2012.

Kurawarwala, A. and Matsuo, H., “Forecasting and inventory management of short life-cycle products,” Operations Research, 44(1): 131-150, 1996.

Li, H. and Huh, W. T., “Optimal Pricing for a Short Life-Cycle Product When Customer Price-Sensitivity Varies Over Time,” Naval Research Logistics, 59(7): 552-576, 2012.

Morana, C., “A Semiparametric Approach to Short-Term Oil Price Forecasting,” Energy Economics, 23(3): 325-338, 2001.

Morrison, J., “How to Use Diffusion Models in New Product Forecasting,” Journal of Business Forecasting Methods & Systems, 15(2): 6, 1996.

Polli, R. and Cook, V., “Validity of the Product Life Cycle,” The Journal of Business, 42(4): 385-400 1969.

Qin, R. and Nembhard, D. A., “Demand Modeling of Stochastic Product Diffusion Over the Life Cycle,” International Journal of Production Economics, 137(2):201-210, 2012.

DECISION SCIENCES INSTITUTE
Machine Learning Methods for Financial Prediction

David E. Booth
Kent State University
Email: dvdbooth8@gmail.com
Pervaiz Alam
Kent State University
Email: palam@kent.edu

Ceyhun Ozgur
Valparaiso University
Email: ceyhun.ozgur@valpo.edu

ABSTRACT

Prediction models are an important part of financial research. In this paper, we review statistical prediction models, their application to finance, introduce Machine Learning methods for financial prediction, and investigate their use in banking research. By using stable statistical prediction, investors and regulators can hope to predict future bank failures and other variables of financial interest. In this paper, we report two successful machine-learning methods to predict the size of a bank one fiscal year ahead of the current date. We show that these models are successful.

KEYWORDS: Prediction, Bank failures, Machine learning, Financial research, Statistical predication, R programming language.

INTRODUCTION

Explanatory models (examples include regression models such as arise in experimental designs) assume a cause and effect relationship between the inputs to the system and its output. This allows one to examine such a potential cause and effect relationship. Predictive models on the other hand attempt to find independent predictor variables that in some sense “best predict” the dependent variable of interest (Shuemeli (2010)). As Shuemli(2010) points out variable selection models are most often the models of choice for identifying useful predictive models). In this paper, we build predictive models for bank size and show they are successful.

LITERATURE REVIEW

The data set that we use to illustrate the methods of machine learning based prediction in finance was first reported by Jin et al (2011). Our data set was collected and formed exactly as reported by Jin et al except that we did not use the audit quality variables, which were not available to us.. The variables used as predictors were those shown in Jin et al’s (2011) Table 2 except that the audit quality variables were not used. Further, we did not want to use the scheme proposed by Jin et al (2011) to attempt to predict bank failures because it requires knowledge of troubled banks, which is usually unavailable to most researchers. Since we did not have that information, we chose instead to predict CS or current status of the bank sample. CS is the values of the size variable for a bank one fiscal year ahead of the period where we are building our prediction model. The potential predictors we use are the ones Jin et al (2011) used in their Table 2. The definitions of these variables are given in our appendix. We do note that regulators and to some degree bank auditors have knowledge of some troubled banks which make the failed bank scheme reported by Jin et al (2011) more workable. However most researchers do not. Thus we follow Shumeli (2010) and use predictive models. We will show that two machine learning variable selection models are successful in predicting bank capital size one fiscal year ahead of the sample data. We note that all analysis programs (written in R) and data sets are available from the authors (DEB).

METHODS

An important tenet of modern statistics is that there are two types of statistical models: explanatory and predictive (Shamueli, 2010). Shamueli (2010) means that the objective of statistical modeling can be to explain or to predict. Explanatory models are the stuff of classical statistics. Their basic use is to test hypotheses and make estimates of model parameters. Predictive models on the other hand are models that are built with the primary goal of making predictions. Here we are most interested in, say, predicting values of a dependent variable from, say a regression model. In this case, variable selection is quite important. In the explanatory case choice of variables is largely driven by theory and the requirements of testing theories. Prediction on the other hand is different. Here our main goal is to get the most accurate possible estimates of the dependent variable in, say, a regression model. Prediction is quite different. In the predictive situation our goal is to get the most accurate possible prediction of the dependent variable or variables. This is what drives our variable selection efforts. It may come as a surprise to some that while theory may assist use in choosing independent variables, it is often not the only thing that is important. As Shumeli remarks (2010, p. 297) the choice of variables can be quite different in the two cases. Since the goals can be remarkably different in the two cases, we were lead to two important findings. First, a good set of explanatory variables is not necessarily a good set of predictor variables. In fact Shumeli, in an appendix (Shumeli, 2010) gives a mathematical proof that such is the case. Further she summarizes by quoting Hagerty and Sirinivasan (1991), “We note the practice in applied research of concluding that a model with a higher predictive validity is ‘truer’ is not a valid inference.” Thus we may conclude that the same set of independent variables may not be a good choice for models that are trying to achieve different goals. This brings us to the topic of variable selection.

In this paper we chose two modern machine learning variable selection methods (Booth et al, 2019), gradient boosting (R package GBM)) and adaptive lasso (Booth and Ozgur, 2019) using the least squares approximation(Boos, 2014). Full details can be found in: Booth et al (2019), Booth and Ozgur (2019), Zou (2006), Elith et al (2008). The R implementation was used in all cases.

RESULTS

Table 1: Prediction for Quarters 13 to 17 – Gradient Boosting

Variable	Relative Influence
Loanmix	36.52
NPL	33.77
CAP	28.534
LLP	0.783
GRestate	0.327
Size	0.0567
PSloans	0.00
GComm	0.00
GLoans	0.00
Cross Validation	
Prediciton Error	
Merr	0.2401
Verr	$3.987\text{x}10^{-3}$
Sderr	0.0063

Table 2: Prediction for Quarters 13 to 17 – Adaptive Lasso LSA*

Adaptive Lasso		
OLS	BIC	AIC
0.0820	0.06	0.067
0.666	0.575	-0.007
-0.013	0.00	0.00
-0.211	-0.21	0.00
-0.047	0.00	0.00
0.0073	0.06	0.00
-0.0119	0.00	0.0096
-0.072	0.00	0.00
Cross Validation		
Prediciton Error		
Merr		0.249
Verr		$3.97\text{x}10^{-3}$
Sderr		0.0063

In gradient boosting large values of relative influence and in adaptive lasso large values of BIC or AIC are the criteria by which variables were selected for predictive model inclusion.

Table 3: Prediction for Quarters 17 to 21 – Gradient Boosting

Gradient Boosting	
Variable	Relative Influence
NPL	64.59
CAP	32.53
LLP	2.87
PSloans	0.00
GComm	0.00
GRestate	0.00
GLoans	0.00
Loanmix	0.00
Size	0.00
Prediciton Error	
Merr	0.237
Verr	8.24×10^{-5}
Sderr	0.009

Table 4: Prediction for Quarters 17 t0 21 – Adaptive Lasso LSA

Adaptive Lasso			
Variable	OLS	BIC	AIC
CAP	-0.2960	0.0198	-0.027
PSloans	-0.01	0.00	-0.007
NPL	-0.291	0.00	0.18
Loanmix	0.053	0.024	0.04
Size	0.005	0.00	0.0055
GComm	0.0585	0.00	0.00
GRestate	0.017	0.00	0.00

Table 5: Prediction for Quarters 21 to 25 – Gradient Boosting

Gradient Boosting	
Variable	Relative Influence
NPL	65.28
CAP	32.01
LLP	2.69
PSloans	0.00
Restate	0.00
Assets	0.00
Comm	0.00
Total loan	0.00
Loanmix	0.00
Size	0.00
GComm	0.00
GRestate	0.00
GLoans	0.00
Prediciton Error	
Merr	0.238
Verr	$7.54x10^{-5}$
Sderr	0.0086

Table 6: Prediction for Quarters 21 to 25 – Adaptive Lasso LSA

Adaptive Lasso			
Variable	OLS	BIC	AIC
CAP	-0.296	-0.019	-0.027
PSloans	-0.01	0.00	-0.007
NPL	1.18	1.23	1.20
LLP	0.291	0.00	0.118
Loanmix	0.053	0.024	0.044
Size	0.0056	0.0049	0.0055
GComm	-0.058	0.00	0.00
GRestate	0.017	0.00	0.00

In the Tables, the variable to be predicted was CS, which is, capsize in the predicted quarters.

In order to make the DVs and IVs as corresponding as possible, we tested the quarters:

Start		Stop
17	to	21
21	to	24
21	to	25

By this, we mean we use the data in quarter 17 to predict CS in quarter 21 in our forecasts as shown in Table 1-6. Stop is the quarter that was predicted and Start is the quarter that provided the predictors. The DV was always CS, capsize. The starting set of predictors were taken from Table 2 of Jin et al (2011). Under the gradient boosting Tables 1, 3, and 5, we list the IVs and the relative influence. The higher the relative influence the more important the predictor is. Under Adaptive Lasso, we list the predictors. Under OLS we list the OLS regression coefficients, under BIC we list the Bayesian Information Criterion values and similarly for the Akaike Information Criteria (AIC). The BIC was chosen as explained in Booth and Ozgur (2019), though some would use AIC. We finally cross validate the prediction error where PE=CS at the ending quarter minus CS from the starting quarter and then at the bottom of each table we list the mean, variance and standard deviation of the prediction errors from the gradient boosting estimates. The IVs chosen for the final prediction equations are chosen from the list of predictors in each Tables 1, 3, and 5 with BIC = 0 (Booth and Ozgur, 2019).

DISCUSSION AND CONCLUSIONS

Looking at Tables 1,3 and 5, the first thing we note is that the summary statistics of the cross validated predictions are small indicating that with prediction equations made up of the predictors chosen with this model, we get quite good predictions over the year-time span. Second, we note that all of our initial lasso predictors were the “best” predictors as used in Table 1 of Jin et al (2011). We note two things here:

1. Our equations contain fewer predictors and hence our gradient boosting prediction models are more parsimonious and
2. Jin et al claimed to have “good” prediction equations.

Two other things:

1. The authors do not give any cross validation information to show that their predictions are “good” and
2. They give no objective criteria for why the IVs they chose were in the model.

We do note that they used a longer time span so that their prediction errors are probably a bit larger than ours are (assuming the same DVs) because of this. We could not obtain the audit variables; therefore, we could not exactly check the prediction equations and the cross validation statistics.

We note the following facts from our estimates given in Tables 1-6:

1. Our gradient boosting and adaptive lasso models contain fewer predictors and hence our gradient boosting prediction models are more parsimonious with small prediction errors. These are the strengths of our models.
2. Jin et al’s models are not claimed to be in any sense “best” predictor models. This is a drawback for those models, as they do nothing at all to argue that they have produced models that are in any sense optimal.
3. We go on to consider our adaptive lasso models in comparison with the models produced by Jin et al. First, recall that we started with Jin et al’s ending selected predictor variables. If we can remove any of these predictors from our models than our models are more parsimonious using the BIC criterion of BIC=0 and looking at the adaptive lasso results in Tables 1-6 we see that is the case. Secondly, we know that (as shown by Zou, 2006) the adaptive lasso estimates satisfy an oracle property. This says that under our conditions that an adaptive lasso model is at least as good as any other model. Since we have a more parsimonious model with the same starting variables as Jin et al, we have models that are superior to what Jin et al would have gotten if the DVs were the same. Here, of course, we are comparing model building steps not regression coefficients. This is true

because of the way Jin et al built their prediction without considering the difference between prediction and explanation (Shamueli, 2010). This is of course a bit of an unfair comparison since Jin et al were trying to predict failure while we were predicting current status and failure is likely more difficult. We ,further, note two things:

1. Our models are more parsimonious and
2. Have either demonstrably small prediction errors (gradient boosting) or the models proven optimal (adaptive lasso).

Thus, we believe our conclusion is fair.

Table 7: The world's largest financial companies and their locations

Name	Service	Country
Berkshire Hathaway	Conglomerate	United States
AXA	Insurance	France
Allianz	Insurance	Germany
ICBC	Banking	China
Fannie Mae	Investment Services	United States
BNP Paribas	Banking	France
Generali Group	Insurance	Italy
China Construction Bank	Banking	China
Banco Santander	Banking	Spain
JP Morgan Chase	Banking	United States
Société Générale	Banking	France
HSBC	Banking	United Kingdom
Agricultural Bank of China	Banking	China
Bank of America	Banking	United States
Bank of China	Banking	China
Wells Fargo	Banking	United States
Citigroup	Banking	United States
Prudential	Insurance	United Kingdom
Munich Re	Insurance	Germany
Prudential Financial	Insurance	United States

*Received from: (Dar-YehHwangCheng F.LeeK.Thomas Liaw)

We remark that we have shown several things in this paper:

1. We make successful predictions of capital size over a one fiscal year period.
2. The two machine learning algorithms were able to successfully predict capital size one fiscal year either ahead with small errors (gradient boosting) or with an optimal model (adaptive lasso).
3. Thus, machine-learning algorithms are valuable additions to the quantitative methods used in finance.
4. As Turak (2018) notes being able to predict capital size or even failure may be an advantage in dealing with or preparing for coming financial crises.

Considering the work that Jin et al (2011) did on failing banks, one might wonder why we chose to build models for the prediction of bank size. If we consider Table 7 in the paper we can see that there are banks that were quite large in the data set.. There is value in predicting things other than failure. We note that our models were able to predict the change in size in bank assets over the period of one fiscal year. Given the size of some of these banks, a major change in asset size could have a large impact on the US Economy even if the bank did not fail but was perhaps bailed out. We now know that we have models that can help us deal with effects such as these that may have large impact.

REFERENCES

D. Boos, Adaptive LASSO in R

Zou, Zhang, & Lu. (n.d.). Adaptive LASSO in R. Retrieved from https://www4.stat.ncsu.edu/~boos/var.select/lasso.adaptive.html.D. Booth, V. Gopalakrisha-Remani, M. Coop, F. Green, and M. Rayman. "Boosting and Lassoing Cancer Risk Factors," *Journal of Data Science*. Under review, March 2019.

Hwang, D.-Y., Lee, C. F., & Liaw, K. T. (2002, April 1). Forecasting bank failures and deposit insurance premium. Retrieved from https://www.sciencedirect.com/science/article/pii/S1059056097900411.

D. Booth and C. Ozgur. "The Use of Predictive Modeling in the Evaluation of Technical Acquisition Performance," *Journal of Data Science*, Accepted January 2019.

J. Elith, J. R. Leathwick and T. Hastie. "A working Guide to Boosted Regression Trees*," Journal of Animal Ecology,* 77(2008)802-813.

J. Jin, K. Kanagaretnam and G. Lobo. "Ability of Accounting and Audit Quality Variables to Predict Bank Failure During the Financial Crisis," *Journal of Banking and Finance* 35(2011)2811-2819.

G. Shamueli. "To Explain or to Predict," Statistical *Science* 25 (2010)289-310.

H. Zhang and W. Lu. "Adaptive Lasso for Cox's Proportional Hazards Model," *Biometrika* (2007)1-13. Advance Access.

H. Zou. "The Adaptive Lasso and Its Oracle Properties," Journal *of the American statistical* association (2006)01(476)1418-1429.

Turak, N. (2018, April 19). 75% of the ultra-rich forecast a US recession in the next two years, survey finds. Retrieved from https://www.cnbc.com/2018/04/19/75-percent-of-the-ultra-rich-forecast-a-us-recession-in-the-next-two-years-survey-finds.html

DECISION SCIENCES INSTITUTE

Exploring Forecasting Methodology With Switching Rule For Periods Of Slow And Fast Demand

Matthew Lindsey
Stephen F. Austin State University
lindseymd@sfasu.edu

Robert Pavur
The University of North Texas
pavur@unt.edu

ABSTRACT

Selecting the optimal forecasting methodology when demand patterns change from slow-moving demand to fast may benefit the inventory manager by reducing cost and increasing customer satisfaction. This research assesses an approach that uses the probability of demand and the size of the demand to determine the appropriate forecasting methodology. The proposed procedure outperforms either the Croston method or SES in this simulation.

KEYWORDS: Croston's Method, Forecasting, Intermittent Demand, Single Exponential Smoothing

INTRODUCTION

A large proportion of time periods with no demand characterizes intermittent demand. Instances of this research are plentiful, but many explore applications of Croston's method (1972) or Single Exponential Smoothing (SES). Croston (1972) exposed that SES was not optimum for cases of intermittent demand and introduced a new method. His method has become the foremost methodology for forecasting slow-moving demand. Willemain, Smart, Shockor, and DeSautels (1994), and Johnston and Boylan (1996) have provided evidence for the advantages of Croston's method when forecasting intermittent demand. Variations have been proposed, particularly by Syntetos and Boylan (2005) who provided a bias correction.

Many retail items or spare parts can be classified as having regular or high demand levels versus having intermittent or slow levels of demand. Seasonality, promotions, special events, trends and correlated demands can influence items with intermittent demand and research has begun to manage these effects (Altay, Litteral & Rudisill, 2012; Lindsey & Pavur, 2008). A situation that has only been addressed in a limited manner is how to handle shifts in mean demand. This is a common case in retail inventory management resulting in the item being categorized as slow moving and using forecasting methodology for intermittent demand to forecast demand during the slow period and then change forecasting methodology.
Leven and Segerstedt proposed a general method when demand is slow or regular, but introduces more bias than Croston's method (Teunter & Sani, 2009). Leven and Segerstedt' s (2004) provided a technique for all stocked items and can be utilized for all demand types but

not for an item with distinct demand rates during two different periods. Despite shortcomings, Croston's method (1972) has found widespread use and has been the benchmark for inventory models and forecasts (Syntetos & Boylan, 2005, 2006; Altay, Rudisill & Litteral, 2008; Teunter, Syntetos & Babai, 2010; Xu, Wang, & Shi, 2012; Ramaekers & Janssens, 2014).

The need for categorizing demand patterns to provide guidance in forecasting stock control is stated in the literature (Syntetos, Boylan, & Croston, 2005; Boylan, et al. 2006). By classifying demand, the optimal methodology for the forecast may be selected (Syntetos, et al 2005). An optimal methodology exists for a given demand level, however, no agreement exists on what the measure should be or what the cut off points should be within a measure for different demand levels. Williams (1984) used the coefficient of variation to describe the level of "lumpiness" in relation to the lead-time to define "intermittence" to select the best method. Shortcomings identified by Syntetos, et al (2005) were that it did not distinguish demand patterns sufficiently. Using Croston's method with the Syntetos and Boylan correction factor, Syntetos, et al (2005) proposed categorizing demand as erratic, lumpy, smooth and intermittent using the average number of periods between demands and the coefficient of variation.

This paper will examine a simulated series that shifts from "fast" demand to "slow" demand. It will use a ratio of Bernouli probabilities (likelihood functions) to determine whether the demand rate is at a historically established "fast" rate or a "slow" rate. Forecasts will be computed using SES, Croston's (1972) method and the Bias Corrected Croston's method by Syntetos and Boylan (2001). Additional forecasts will be computed using two switching rules and a combination of the switching rules. A rule of "no demand for two periods" and a "control chart" for the average demand mean will be used to indicate the need to switch methodologies. Finally, a combination of the two rules will be used. The Average Root Mean Square Error (RMSE) will be reported to compare forecasts.

LITERATURE REVIEW

Croston's method (1972) is a popular technique to forecast when demand is intermittent and is the standard for inventory models (Teunter, Syntetos & Babai, 2010; Xu, Wang, & Shi, 2012; Ramaekers & Janssens, 2014). Croston's method has documented benefits by Willemain, Smart, Shockor, & DeSautels (1994), Johnston and Boylan (1996) and others. Syntetos and Boylan (2001) corrected a bias in Croston's method. Many studies have compared the bias corrected method with the original without identifying a superior method in the evaluations (Eaves & Kingsman, 2004; Syntetos & Boylan, 2005; Teunter & Sani, 2009; & Teunter & Duncan, 2009). A generalized method when demand is slow or regular has been suggested, but it potentially introduces more bias than Croston (Teunter & Sani, 2009).

Categorization

Categorizing demand distributions to provide guidance in forecasting and stock control is suggested in the literature. (Syntetos, Boylan, & Croston, 2005; Boylan, et al. 2006). By correctly categorizing demand, the practitioner can select an optimal forecasting method (Syntetos, et al 2005). The literature suggests a proper course of action exists for a given demand level, but not what the measure should be or the cutoff point within a measure.

Aggregation Approaches

Aggregation provides an alternate approach to forecasting slow demand. It may eliminate slow periods altogether. An aggregate-disaggregate intermittent demand Approach (ADIDA) is proposed by Nikolopoulos, Syntetos, Boylan, Petropolulos and Assimakopoulos (2011). The ADIDA method eliminates intermittent data by condensing it into larger periods that can be forecasted and then redistributed. While the original ADIDA methodology was identified to have specific weaknesses (Petropoulos & Kourentzes, 2015; Rostami-Tabar, Babai, Syntetos, & Ducq, 2013; Spithourakis, Petropoulos, Nikolopoulos, & Assimakopoulos, 2012) several variations have been suggested for specific demand patterns and to address the problems with ADIDA (Petropoulos, Kourentzes, & Nikolopoulos, 2016; Kourentzes, Rostami-Tabar, & Barrow, 2017; Athanasopoulos, Hyndman, Kourentzes, & Petropoulos, 2017; and Li & Lim, 2018.) ADIDA could be applicable to shifts in demand.

Croston's Method

Willemain, et al. (1994) should be consulted for the methodology for implementing Croston's method used in this paper. Croston's technique is comparable to SES and assumes a constant demand mean of size of μ taking place every p periods, so the average demand is not μ/p but

$$y^* = (\mu/p)\ (p\alpha)/(1-(1-\alpha)^p) \tag{1}$$

Syntetos and Boylan's Bias Corrected Modification for Croston's Procedure

Syntetos and Boylan (2001) identified a bias in Croston's (1972) method. The paper should be consulted for the methodology. The bias increases as the value of the smoothing constant increases. It can be minimized by using a small alpha level. Boylan and Syntetos (2005) correct for the bias using Croston's method by multiplying the demand per period by $1-\alpha/2$. Teunter and Sani (2009) advocate that in some cases when only a limited periods have no demand, Croston's method excels. However, in the case when most periods have no demand, the use of the Syntetos and Boylan's correction is better. Syntetos and Boylan (2005) also remind us that Croston's assumptions apply to the new estimator.

Single Exponential Smoothing

SES is a weighted-moving average technique conventionally used to forecast demand for inventoried items. Any general statistics book will explain the method in detail. Trigg and Pitts (1962) provided guidelines for alpha selection. Later Ekern (1983) warns against "adaptive" methods that automatically adjust the constant in response to the forecast performance and identifies the case of the nonstationary demand rate as being especially difficult. Saygin (2005) considers the ability to maximize service level and minimize inventory cost in alpha selection.

Hybrid Approach: Adapting Croston's Method for Two Demand Levels

The conventional use of Croston's method is that it is based on exponential smoothing and will adapt to changing data patterns. Using the same notation as in Croston's method, the following approach is considered a modification of Croston's method. A superscript of S and F are used to distinguish between two modes of sales, slow and fast. For this study, the rule that if the last two time periods did not have a sale, then the current mode of sales is classified as “slow.”

If $X_{t-1} = 0$ and $X_{t-2} = 0$,

Then If $X_t = 0$, $Z^S_t = Z^S_{t-1}$
$P^S_t = P^S_{t-1}$
$q = q + 1$ (2)
Else $X_t = 1$ $Z^S_t = Z^S_{t-1} + \alpha(y_t - Z^S_{t-1})$
$P^S_t = P^S_{t-1} + \alpha (q - P^S_{t-1})$
$q = 1.$
Count_Slow = Count_Slow + 1 (3)

If $X_{t-1} = 1$ or $X_{t-2} = 1$,

Then If $X_t = 0$, $Z^F_t = Z^F_{t-1}$
$P^F_t = P^F_{t-1}$
$q = q + 1$ (4)
Else $X_t = 1$ $Z^F_t = Z^F_{t-1} + \alpha (y_t - Z^F_{t-1})$
$P^F_t = P^F_{t-1} + \alpha (q - P^F_{t-1})$
$q = 1.$
Count_Fast = Count_Fast + 1 (5)

The mean demand per period is then
(Count_Slow * Z^S_t / P^S_t + Count_Fast * Z^F_t / P^S_t)/(Count_Slow + Count_Fast) (6)

EXPERIMENTAL SIMULATION STUDY

To determine whether the data can be classified as fast moving or slow moving, a ratio of Bernoulli likelihood functions are used. If this ratio is greater than one than the data are slow moving, otherwise, the data are classified as fast moving.

Two rules and a combination of the two will be assessed in this simulation. One rule will be whether 2 previous periods have no demand. If true, it will signal to switch from SES. A second rule, “control chart”, based on the average demand per period will be used to signal to switch from SES when the average demand deviates from the mean. A combination of the two switching rules will also be used. Previous research by the authors compared forecasts using a switching rule of 1 period with no demand up to 6 periods with no demand for a variety of demand rates. To investigate the performance of different switching rules on a modification to Croston’s method in the presence of two distinct demand rates, a simulation was performed in which 1000 time periods are computed. Forecasts by SES, traditional Croston, and the bias corrected Croston approach are computed. The Average Root Mean Squared Error will be used to measure the accuracy of these methods.

A smoothing constant of 0.1 was used in this simulation. The length of slow moving periods was set at 300 and the length of the fast moving periods was set at 200 and then again 300 slow and 200 fast for a total number of time periods of 1000. The mean product demand for slow-moving periods was set at 200 and the mean demand for fast moving periods was fixed at 230. The standard deviation for the demand was fixed at 10. Each test situation consisted of 500 simulation repetitions to measure the root mean square error for each method forecasting the average demand over 1000 periods of time. Simulation parameters are provided in Table 1.

Table 1. Simulation Parameters

Simulation Periods :	1000
Number of Simulations:	500
Slow Periods	300
Fast Periods	200
Slow	**Fast**
Probability of Demand	
40% or 60%	80%
Mean Average Demand	
200	230
Demand Standard Deviation	
10	10

Forecasts were computed using SES, Croston method and the bias corrected Croston's method. To compare the methods, the reduction in forecast error in percent was computed for Croston's method and SES and between the modified Croston method and SES.

RESULTS

The forecast error using SES, Croston and the Bias Corrected Croston Method are given in Table 2. The average RMSE is then provided with a switching rule. The Hybrid Croston method percent improvement improved over forecasts using SES and Croston's method as the slow demand rate decreased and the number of periods of no demand required to switch methods decreased. The switching rule was one of three rules: 1.) Switch to Slow when no demand occurs in the last two periods. 2.) Switch to slow when the difference between the sample mean and the true value of the Mean Demand under the Fast Scenario is greater than 1.6 times the standard deviation of demand and dividing by the square root of 5. 3.) Switch to slow when both of the conditions in 1 and 2 occur.

DISCUSSION AND CONCLUSIONS

Knowing when to use the proper forecasting methodology improves the quality of the forecast. As shown in the table 2 above, one size does not fit all. Forecast error is decreased when the methodology is adapted for slower demand periods. These results suggest two things. First, utilizing more information in the forecasting procedure is beneficial and knowing when to reclassify a product based on demand is important. Second for slow moving inventory the sooner the slow-moving case can be identified, the better, for forecasting accuracy. It also appears that when the probability of demands are closer between the slow moving and fast demand rates, that is the rates are similar, the advantage from switching at the right time is less.

Table 2. Performance of SES and Croston's Method with and without Switching Rules

No Switching Rule - Probability for Slow is 0.4 & Probability for Fast is 0.8			
Forecasting Method	**Average Root Mean Square Error**		
SES (alpha= .1)	53.7		
Croston (1972)	49.9		
Croston Bias Corrected	50.2		
Croston Bias Correct (known P)	48.8		
Switching Rule			
Switching Rule Croston/SES	**Average Root Mean Square Error**	**% RMSE reduction over SES**	**Percent RMSE reduction over Croston**
No Demand for 2 periods	26.1	51.4%	47.7%
Control Chart for Mean Demand	31.8	40.8%	36.3%
Control Chart and no Demand	22.4	58.3%	55.1%
No Switching Rule - Probability for Slow is 0 .6 & Probability for Fast is 0.8			
Forecasting Method	**Average Root Mean Square Error**		
SES (alpha= .1)	37.3		
Croston (1972)	34.2		
Croston Bias Corrected	34.6		
Croston Bias Correct (known P)	34.0		
Switching Rule			
Switching Rule Croston/SES	**Average Root Mean Square Error**	**% RMSE reduction over SES**	**Percent RMSE reduction over Croston**
No Demand for 2 periods	24.8	33.5%	27.4%
Control Chart for Mean Demand	23.7	36.5%	30.7%
Control Chart and no Demand	21.2	43.2%	38.0%

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Application of Auxiliary Variable Parameters for Enhanced Estimation of Population Variance for Business Decisions

Dinesh K. Sharma
University of Maryland Eastern Shore
Email: dksharma@umes.edu

S. K. Yadav[1]
Babasaheb Bhimrao Ambedkar University
Email: drskystats@gmail.com

Hari Sharma
Virginia State University
Email: hsharma@vsu.edu

ABSTRACT

In the paper, we utilized the given parameters of the auxiliary variable for improved estimation of population variance of the study variable. We proposed a Searls type estimator of population variance using known population tri-mean and third quartile of auxiliary variable. The natural properties of the proposed estimator are studied up to the approximation of degree one. The precision of the estimation of variance is tested using mean squared error (MSE) with the competing estimators of the population variance. The efficiency constraints are obtained for the proposed estimator and are justified through the numerical example. Finally, the improvement over the competing estimators is demonstrated by lower mean squared error (MSE).

KEYWORDS: Population Variance, Searls Type Estimator, Auxiliary Variable, Bias, MSE.

INTRODUCTION

One of the important measures of dispersion is the Population Variance. Variation is natural and it can be seen in our day-to-day life. However, precise estimation is crucial for planning, policy making and other business decisions. The most suitable estimator for population variance is the sample variance, but its sampling distribution is widely dispersed around the true population variance. Thus, we search even for a biased estimator to bridge the gap between sampling variance and true population variance. This is achieved through the use of auxiliary information. Many authors suggested various improved estimators of population variance using auxiliary variable. They used the known parameters of auxiliary variable like coefficient of variation, coefficient of skewness and kurtosis along with many conventional and non-conventional parameters of auxiliary variable. This study is motivated by Searls (1964) and Yadav et al. (2019). In this study, we propose a Searls type estimator of the population variance and apply the known population tri-mean and third quartile of the auxiliary variable for improving the estimation.

[1] Corresponding Author

NOTATIONS USED

N - Size of Population

n - Size of Sample

$f = n/N$ - Fraction of Sampling

Y - Main Variable

X - Secondary variable

$\bar{Y}$ - Population Mean of Study Variable

$\bar{X}$ - Population Mean of Auxiliary Variable

S_y - Population Standard Deviation of Study Variable

S_x - Population Standard Deviation of Auxiliary Variable

ρ - Population Correlation Coefficient between Study and Auxiliary Variable

S_{yx} - Population Covariance between Study and Auxiliary Variable

C_y - Population Coefficient of Variation of Study Variable

C_x - Population Coefficient of Variation of Auxiliary Variable

Q_i - Population i^{th} Quartile of Auxiliary Variable

M_d - Population Median of Auxiliary Variable

β_1 - Population Coefficient of Skewness of Auxiliary Variable

β_2 - Population Coefficient of Kurtosis of Auxiliary Variable

$Q_r = Q_3 - Q_1$ - Population Quartile Range of Auxiliary Variable

$Q_d = \frac{Q_3 - Q_1}{2}$ - Population Quartile Deviation of Auxiliary Variable

$Q_a = \frac{Q_3 + Q_1}{2}$ - Population Quartile Average of Auxiliary Variable

$TM = \frac{Q_1 + 2Q_2 + Q_3}{4}$ - Population Tri-mean of Auxiliary Variable

LITERATURE REVIEW

Various modified estimators of population variance have been given by various authors till date. Different estimators of population variance by different authors are presented in Table 1.

Table 1: Various estimators of population variance in the literature

Year	Authors	Journal
1983	Isaki	Journal of American Statistical Association
1999	Upadhyaya and Singh	Biometrical Journal
2006	Kadilar and Cingi	Hacettepe Journal of mathematics and Statistics
2012a	Subramani & Kumarpandiyan	International Journal of Probability and Statistics
2012b	Subramani & Kumarpandiyan	International Journal of Statistics and Applications
2013	Subramani & Kumarpandiyan	Pakistan Journal of Statistics and Operation Research
2013	Khan & Shabbir	Journal of Statistics Applications & Probability
2017	Maqbool and Javaid	American Journal of Biological and Environmental Statistics
2018	Khalil *et al.*	Oriental Journal of Physical Sciences
2019	Yadav *et al.*	Int. J. Business and Data Analytics

PROPOSED ESTIMATOR

Motivated from the literature, we suggest the following estimator as,

$$t_p = \kappa s_y^2 \left[\frac{S_x^2 + (TM + Q_3)}{s_x^2 + (TM + Q_3)} \right] \tag{1}$$

Where, κ is a characterizing scalar which is so obtained that the MSE of t_p is a minimum.

The bias and MSE of suggested estimator up to first order of approximation respectively are,

$$B(t_p) = \gamma \kappa S_y^2 [R_{17}^2 (\lambda_{04} - 1) - R_{17} (\lambda_{22} - 1)] + S_y^2 (\kappa - 1) \tag{2}$$

$$MSE(t_p) = S_y^4 [\kappa^2 \gamma (\lambda_{40} - 1) + (3\kappa^2 - 2\kappa) R_{17}^2 \gamma (\lambda_{04} - 1) - 2(2\kappa^2 - \kappa) R_{17} \gamma (\lambda_{22} - 1) + (\kappa - 1)^2] \tag{3}$$

The Minimum MSE for optimum κ is,
The MSE of the proposed estimator is obtained for the optimum value of κ as,

$$\kappa = \frac{A}{B} \tag{4}$$

Where,

$A = 1 + R_{17}^2 \gamma (\lambda_{04} - 1) - R_{17} \gamma (\lambda_{22} - 1)$ and

$B = 1 + \gamma (\lambda_{40} - 1) + 3R_{17}^2 \gamma (\lambda_{04} - 1) - 4R_{17} \gamma (\lambda_{22} - 1)$

$$MSE_{\min}(t_p) = S_y^4 \left[1 - \frac{A^2}{B} \right] \tag{5}$$

EFFICIENCY COMPARISON

Under this section, various competing estimators are compared with the suggested estimators and are presented in Table 2. The conditions under which proposed estimators performs better than the competing one are also given in Table 2.

Table 2: Theoretical comparison of suggested estimator with various competing estimators of population variance		
S. No.	Estimator	Efficiency condition
1.	Sample variance	$V(t_0) - MSE_{\min}(t_p) > 0$
2.	Isaki (1983) estimator	$MSE(t_R) - MSE_{\min}(t_p) > 0$
3.	Estimators in Table 1	$MSE(t_i) - MSE_{\min}(t_p) > 0$

NUMERICAL STUDY

For theoretical verification, we have used the population given in Yadav *et al.* (2019). The parameters of this population are presented in Table 3.

Table 3: Parameters of the population in Yadav *et al.* (2019)
$N = 80$, $n = 20$, $\bar{Y} = 51.8264$, $\bar{X} = 11.2646$, $\rho = 0.9413$, $S_y = 18.3549$, $C_y = 0.3542$, $S_x = 8.4563$, $C_x = 0.7507$, $\lambda_{04} = 2.8664$, $\lambda_{40} = 2.2667$, $\lambda_{22} = 2.2209$, $Q_1 = 5.1500$, $Q_3 = 16.975$, $Q_r = 11.825$, $Q_d = 5.9125$, $Q_a = 11.0625$, $T_m = 9.318$, $M_d = 7.575$

The biases and the mean squared errors and percentage relative efficiency (PRE) of some estimators along with the proposed estimators are presented in Table 4.

Table 4: Bias, MSE and PRE of suggested estimator over the competing estimators			
Estimator	Bias	MSE	PRE
Sample variance	0	5,393.89	271.57
Isaki (1983) estimator	10.87	3,925.16	197.62
Upadhyaya and Singh (1999) estimator	9.29	3,658.41	184.19
Kadilar and Cingi (2006) estimator	10.44	3,850.16	193.84
Subramani & Kumarpandiyan (2013) estimator	-0.94	2,467.88	124.25
Khan & Shabbir (2013) estimator	3.62	2,878.56	144.93
Maqbool and Javaid (2017) estimator	3.03	2,820.06	141.98
Khalil *et al.* (2018) estimator	-0.56	2,580.75	129.93
Yadav *et al.* (2019) estimator	1.33	2,040.12	102.71
Sharma and Yadav (Proposed) estimator	-5.90	1,986.22	100.00

CONCLUSION

In the present manuscript, we have suggested an improved estimator of population variance using known auxiliary variable. The biases and MSEs of suggested estimators are obtained up to the first order of approximation. The efficiency comparisons have been made with the competing estimators. The conditions have been defined for the suggested estimator under which it performs better than competing estimators of population variance under simple random sampling scheme. The study for the selected data set shows that the suggested estimator has least MSE among the competing estimators. We found that the proposed estimator is the best among the competing estimators of population variance under simple random sampling scheme.

The suggested estimator will be benefial in planning and making better business decisions for professionals.

REFERENCES

Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of American Statistical Association,* 78, 117–123.

Kadilar, C., & Cingi, H. (2006). Improvement in variance estimation using auxiliary information Hacettepe, *Journal of mathematics and Statistics,* 35, 111–115.

Khalil, M., Ali, M., Shahzad, U., Hanif, M., & Jamal, N. (2018). Improved Estimator of Population Variance using Measure of Dispersion of Auxiliary Variable. *Oriental Journal of Physical Sciences*, 3, 1, 33-39.

Khan, M., & Shabbir, J. (2013). A ratio type estimator for the estimation of population variance using quartiles of an auxiliary variable. *Journal of Statistics Applications & Probability*, 2, 3, 319–325.

Maqbool, S., & Javaid, S. (2017). Variance estimation using linear combination of tri-mean and quartile average, *American Journal of Biological and Environmental Statistics*, 3, 1, 5–9.

Searls D.T. (1964). The Utilization of Known Coefficient of Variation in the estimation procedure. *Journal of American Statistical Association*. 59, 1125-1126.

Subramani, J., & Kumarapandiyan, G. (2013). A new modified ratio estimator of population mean when median of the auxiliary variable is known, *Pakistan Journal of Statistics and Operation Research*, 9, 2, 137–145.

Upadhyaya, L.N., & Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population mean, *Biometrical Journal*, 41, 5, 627–636.

Yadav, S.K., Sharma, D.K., & Mishra, S.S. (2019). Searching efficient estimator of population variance using tri-mean and third quartile of auxiliary variable, *Int. J. Business and Data Analytics,* 1, 1, 30-40.

(A complete list of references is available upon request)

DECISION SCIENCES INSTITUTE

When a Variable Is Not Perfectly Correlated with Itself:
Time to Retire Kendall's Tau_a Rank Correlation

John R. Dickinson
University of Windsor
Email: MExperiences@bell.net

ABSTRACT

Kendall's tau_a rank correlation coefficient is commonly presented in research methods texts and is readily available in statistical software packages. Tau_a is not defined to accommodate ties in the ranked data it analyzes. Kendall's tau_b correlation, though, is defined to accommodate ties in the ranked data. Where there are no ties in the ranked data, the tau_b formula is the same as tau_a, i.e., tau_b completely supplants tau_a. The present research suggests that tau_a not be presented in textbooks or statistical software packages.

KEYWORDS: Kendall's tau_a, Kendall's tau_b, rank correlation

INTRODUCTION

Kendall's (1938) tau_a rank correlation coefficient is commonly presented in research methods texts. Too, it is readily available in statistical software packages. Tau_a is not defined to accommodate ties in the ranked data it analyzes. While this limitation is not unknown, it is not always reported in texts and statistical software guides and is routinely produced by the latter. These conditions set the stage for the misapplication of tau_a. The present research suggests that tau_a be retired from textbooks and statistical software packages. Such retirement would not spell the end of Kendall's rank correlation for data without ties.

Kendall's tau_b correlation is defined to accommodate ties in the ranked data. Where there are no ties in the ranked data, the tau_b formula is the same as tau_a, i.e., tau_b completely supplants tau_a. Tau_a serves no purpose beyond tau_b and, again, its presence in texts and software packages facilitates its misapplication.

TEXTBOOK TREATMENTS

Though not presenting formulae or examples, Malhotra (2010) notes Spearman's rho and Kendall's tau (p. 536) without distinguishing between the two versions–not accounting and accounting for ties–of tau (or the two versions of rho). Stated without explanation, though, is that, "...when the data contain a large number of tied ranks, Kendall's τ seems more appropriate." p. 536) Malhotra's presentation does not include formulae or examples of nonmetric correlation. Stated, though, is that, "Both vary from -1.0 to 1.0.", again not apprising the student that this may not be the case in the presence of tied data.

Iacobucci & Churchill (2010) presents Spearman's rank correlation coefficient (pp. 385-386), but none of Kendall's coefficients. The possibility of tied ranks is not noted and the formula presented for Spearman's rho does not account for ties. Stated, though, is that, "The upper limit for the Spearman rank correlation coefficient is 1.0..." (p. 386) (Kendall is mentioned on page 388 in the context of the coefficient of concordance.)

Burns & Bush (2003) describes monotonic relationships (pp. 518-520, 542) but does not present any rank order correlation statistics: "...a rank order correlation's relationship falls between these two [chi-square and Pearson] with a monotonic relationship." (p. 542)

Zikmund (2003, pp. 571-572) presents the uncorrected formula for tau, noting only that, "Tied observations are treated in the same way as for the Spearman coefficient." (p. 571) That "same way" is to assign the average rank to tied observations and, "...a correction factor [unspecified] can be introduced to offset their effect..." (p. 570)

A textbook that does reflect the suggestion put forth here–that Kendall's tau_a be retired–is Cooper & Schindler (2006). Kendall's tau_b is presented (pp. 561-563), including the formula correcting for ties and notes, "Its compensation for ties.." (p. 561). Spearman's rho, too, is presented, though with the formula not correcting for ties and no mention of the treatment of tied data.

Cited here are instances where textbook treatments of rank correlation are incomplete. The implication of those instances is **not** that textbooks *should* include more complete presentations. The intended implication is that it can be easily understood why researchers may not appreciate the limitation of Kendall's tau_a.

OTHER PRESENCES OF TAU_a

In the context of the R statistical package Yao's (undated) online *R Tutorial eBook* presents "Kendall Rank Coefficient" and states the formula for the uncorrected-for-ties tau_a. (An endnote refers to a significance test for tau_b.) Wikipedia (2019) reports tau_a as well as tau_b and the Stuart-Kendall tau_c.

The "Nonpar Corr" procedure in SPSS Version 25 produces Kendall's tau_b and tau_a is not noted (p. 1301).

THE SUFFICIENCY OF KENDALL'S TAU_b RANK CORRELATION STATISTIC

Supporting the essential point of the present report, i.e., that Kendall's tau_a should be retired, is that his tau_b statistic completely supplants tau_a. That this is so is easily demonstrated. One formula for tau_b, the corrected statistic, is (Siegel 1956, p. 218, Equation 9.10):

$$\tau_b = \frac{S}{\sqrt{(1/2)N(N-1)-T_x}\,\sqrt{(1/2)N(N-1)-T_y}}$$

Excepting T_x and T_y, definitions of the other terms are not pertinent here. The "T" terms represent tied data. Where no ties are present they equal 0, leading to:

$$\tau_b = \frac{S}{\sqrt{(1/2)N(N-1)-0}\ \sqrt{(1/2)N(N-1)-0}}$$

$$\tau_b = \frac{S}{\sqrt{(1/2)N(N-1)}\ \sqrt{(1/2)N(N-1)}} = \frac{S}{(1/2)N(N-1)}$$

Siegel's formula for tau_a (p. 217, Equation 9.9) is:

$$\tau_a = \frac{S}{(1/2)N(N-1)}$$

In the absence of ties, tau_b equals tau_a. With this, it seems that there is no use for Kendall's tau_a that is not inherently present in tau_b.

AN INSTANCE

In research in the social sciences the correlation matrix is well-known. The square matrix comprises correlations between pairs of variables, the same variables comprising the rows and columns of the matrix. Entries on the matrix diagonal are correlations of a variable with itself, tautologically perfect correlation. That is, the entries on the diagonal of the correlation matrix must equal 1.0. Perhaps the most common type of correlation is the Pearson product moment correlation. However, for some applications, rank order correlation may be more appropriate.

For example, the Taxonomy Index (*TaxI*, Dickinson 2013) is a statistic describing the accuracy of published multiple-choice question difficulty taxonomies. *TaxI* commences with rank ordering questions according to their observed difficulty. A variety of measures of item difficulty exists, including the most common one, the percent of students answering the item correctly (Haladyna 2004, p. 207; Murphy & Davidshofer 1988, p. 131). A second measure of item difficulty divides examinees into the 27 percent of examinees receiving the highest score on an exam and the 27 percent receiving the lowest score. (Aiken (1991) The item difficulty index, then, is the percent of those combined subgroups answering the question correctly. Yet other measures are those of Horst (1933) and Guilford (1936) that correct for guessing.

Since the purpose of the specific measure employed is to rank order questions, that purpose is sufficiently met by any difficulty measure among those that are monotonically related. The measures need not be linearly (i.e., Pearson) correlated. Thus, a rank order correlation statistic might be employed to compare/contrast different measures of difficulty to determine which are substitutable for one another. Kendall's (1938) tau_a is one such rank correlation statistic.

Accordingly, using Kendall's tau_a rank correlation, for a random sample of 612 questions from the published bank of multiple-choice questions accompanying Levy, Weitz, and Grewal (2014) a correlation matrix was calculated (Table 1). As may be seen, the correlations on the diagonal do not equal 1.0. The reason that the entries on the diagonal in Table 1 do not equal 1.0 is, of course, that Kendall's tau_a is not defined to handle tied ranks.

Table 1: Difficulty Measure Intercorrelations (Kendall's Tau_a)

Difficulty Measure	Percent Correct	Percent Correct, Upper~Lower	Horst	Guilford
Percent Correct	0.9912	0.7952	0.9129	0.9887
Percent Correct, Upper~Lower 27%	0.7952	0.9725	0.7633	0.7821
Horst	0.9129	0.7633	0.9940	0.9149
Guilford	0.9887	0.7821	0.9149	0..9918

Correlations of four measures of item difficulty for a random sample of 612 questions

ERROR IN CALCULATED TAU_a VALUES AND THE PRESENCE OF TIED DATA MAY NOT BE OBVIOUS, BUT THAT SHOULD BE MOOT

The instance presented in Table 1 is unusual in that the incorrectness of the diagonal elements is readily apparent; it is known that they should equal 1.0. In most applications, though, no known value for the rank correlation exists. The example in Table 1 is of happenstance. Generally, the error in a calculated tau_a value is not known from the calculated value itself.

Textbook examples of rank order correlations typically use very small numbers of observations with the data exhaustively listed in a table: e.g., 15 observations (Iacobucci & Churchill 2010, Table 14A.4, p. 385), 10 observations (Cooper & Schindler 2006, Exhibit 19-24, p. 563), 11 observations (Gibbons 1993, Table 2.4, p. 10), 12 observations (Siegel 1956, Table 9.6, p. 209). With this, instances of tied ranks are easily identified and, of course, the authors explicitly point them out.

Real world applications, though, may involve hundreds or thousands of observations. (The numbers of observations in the multiple-choice question data that gave rise to this research range from 612 to 958 across six published question banks.) Accordingly, identifying tied observations is much more problematic.

With the retiring of Kendall's tau_a, happenstance or inspection of the data would be made moot. Kendall's tau_b formula, as has been shown above, accommodates both tied and not-tied ranks.

DISCUSSION

Per its title, the purpose of the present study is to suggest that Kendall's tau_a be retired; from textbooks, statistical computer packages, and applications. Kendall's tau_b inherently corrects for ties where they exist and subsumes tau_a thereby obviating the latter. That tau_a does not take into account tied data has been long known (Siegel 1956, pp. 217-219; Gibbons 1993, pp. 15-16). That noted, as cited above it is common for research methods textbooks to give but passing attention to this limitation if any attention at all is given. Also as cited above, popular statistical packages, e.g., R, produce Kendall's tau_a as well as Kendall's tau_b when taub is sufficient and producing tau_a sets the stage for its misapplication.

Mitigation

It might be maintained–with a degree of correctness–that the "limitation" of Kendall's tau_a highlighted here is more accurately positioned as a limitation of using that correlation in the first place. That is, the limitation is with the researcher who used an inappropriate statistical analysis. That granted as a normative prescription, several conditions mitigate that prescription.

Many, perhaps most, researchers are knowledgeable of the substance of their research, and not necessarily thoroughly knowledgeable of the statistical techniques they apply. *Should* researchers be expected to be thoroughly knowledgeable of the statistical methods they apply?

The standard deviation is a commonly applied measure of dispersion. Its calculation involves squaring of differences of observations from the mean of those observations. The effect of small deviations becomes enlarged with their being squared. But the effect of large deviations becomes much more enlarged with their being squared. It may be observed that this disproportionately increases the effect of large deviations. The less commonly applied mean absolute deviation (MAD) is a function of the absolute value of deviations, not their squared value; the deviations are taken as is.

Few researchers could explain the rationale for squaring the deviations. Nor should they be expected to be able to. The researchers sensibly apply what "everyone else" applies without belaboring its founding reasoning. A parallel situation exists in the least-squares method of estimation in linear regression. Why are the error terms squared, rather than using their absolute values?

There are, of course, tenable reasons for these statistical approaches. The issue, though, is whether researchers should be expected to pursue to their origins statistical methods already widely accepted.

REFERENCES

Aiken, Lewis R. (1991). *Psychological testing and assessment*, Seventh Edition. Boston: Allyn and Bacon. ISBN: 0-205-12864-5

Burns, Alvin C. & Bush, Ronald F. (2003). *Marketing research: online research applications*, Fourth Edition. Upper Saddle River, NJ: Prentice Hall. ISBN: 0-13-035135-0

Cooper, Donald R. & Schindler, Pamela S. (2006). *Business research methods*, Ninth Edition New York: McGraw-Hill Irwin. ISBN: 0-07-297923-2

Dickinson, John R. (2013). *TaxI*: a statistic describing the accuracy of multiple-choice question difficulty classifications. In Kubacki, Krzystof (Ed.), *Proceedings of the 2013 Academy of Marketing Science Annual Conference*, *XXXVI*, 337.

Gibbons, Jean Dickinson (1993). *Nonparametric measures of association*. Newbury Park, CA: Sage Publications, Inc. ISBN: 0-8039-4664-3

Guilford, J. P. (1936). The determination of item difficulty when chance success is a factor. *Psychometrika*, *1*, December, 259-264.

Haladyna, Thomas M. (2004). *Developing and validating multiple-choice test items*, Third Edition. London: Routledge; Taylor & Francis Group. ISBN: 0-8058-4661-1

Horst, Paul (1933). The difficulty of a multiple choice test item. *Journal of Educational Psychology*, *24*(3), 229-232.

Iacobucci, Dawn & Churchill, Gilbert A., Jr. (2010). *Marketing research: methodological foundations*, Tenth Edition. Mason, OH: South-Western Cengage Learning. ISBN: 13-978-0-324-35995-4; ISBN: 10-0-324-35995-0

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), June, 81-93.

Levy, Michael, Weitz, Barton A., & Grewal, Dhruv (2014). *Retailing management*, Ninth Edition. New York: McGraw-Hill Irwin. ISBN: 978-0-07-802899-1, MHID: 0-07-802899-x

Malhotra, Naresh K (2010). *Marketing research: an applied orientation*, Sixth Edition. Upper Saddle River, NJ: Prentice Hall. ISBN: 978-0-13-608543-0

Murphy, Kevin R. & Davidshofer, Charles O. (1988). *Psychological testing*. Englewood Cliffs, NJ: Prentice Hall. ISBN: 0-13-732587-8

Siegel, Sidney (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill Book Company. ISBN: 07-057348-4

SPSS (undated). Chapter 170. Nonpar Corr. *IBM SPSS statistics 25 command syntax reference*, downloaded May 10, 2019.

Wikipedia (2019, last edited March 4). Kendall rank correlation coefficient.

Yau, Chu (undated). *R tutorial ebook*. www.r-tutor.com/gpu-computing/correlation/kendall-rank-coefficient, downloaded December 28, 2018.

Zikmund, William G. (2003). *Business research methods*, Seventh Edition. Mason, OH: Thomson South-Western. ISBN: 0-03-035084-0

DECISION SCIENCES INSTITUTE

Privacy, Awareness, and the Internet of Things

C. Bryan Foltz
University of Tennessee at Martin
Email: foltz@utm.edu

Laura Foltz
University of Tennessee at Martin
Email: lfoltz@utm.edu

ABSTRACT

This research examines the applicability of the Dinev and Hu (2007) awareness centric model of human behavior to the Internet of Things domain. This model is used to examine user intention to modify security settings on Internet of Things devices.

KEYWORDS: Privacy and Security, Consumer Choice and Judgement, Internet of Things

INTRODUCTION

Mason (1986) listed privacy as one of four primary ethical issues facing the information age. This early concern over privacy has only increased as technology has changed over time with modern technology posing a plethora of new privacy threats (Belanger & Crossler, 2011). Privacy models must, therefore, evolve to incorporate the technological changes.

The Internet of Things (IOT) is one such technological advancement. Despite rapid growth, the Internet of Things poses multiple security risks by collecting a variety of private data through smart devices such as watches, dish washers, and digital assistants. Many Internet of Things devices allow users to alter privacy settings to mitigate these risks. This research seeks to understand the degree to which users are aware of these configuration options and utilize these protective settings.

Zheng et at (2018) suggest that any solution to Internet of Things (IoT) privacy must consider user attitudes and awareness. This research examines user awareness of privacy settings in conjunction with privacy concerns in terms of IoT devices by utilizing the Dinev & Hu (2007) instrument.

LITERATURE REVIEW

Privacy

Although privacy has been a concern within the extant information systems literature for many years, interest in this topic extends across multiple disciplines (Mason, 1986; Belanger & Crossler, 2011; Dinev & Hart, 2004). Interest in privacy has increased as a result of the rapid growth of technology (Dinev & Hart, 2004; Belanger & Crossler, 2011). Belanger and Crossler (2011) note that many definitions of privacy incorporate the same basic concept: control over the use (especially the secondary use) of one’s information. Further, the authors suggest that

"Information privacy refers to the desire of individuals to control or have some influence over data about themselves." (Belanger & Crossler, 2011).

Information privacy concern is often used as a proxy for the idea of privacy within the Information Systems literature (Xu et al, 2008). Information privacy concern may be defined as "concerns about possible loss of privacy as a result of information disclosure to a specific external agent" (Xu et al, 2012, p2).

The Internet of Things

The Internet of Things (IoT) refers to a "dynamic global network infrastructure, where objects with unique identifications, wireless communication and computing capabilities are integrated into an information network" (Boos et al, 2013). The IoT acts as a bridge to connect physical and digital worlds by collecting data, largely without human interaction. IoT devices range from wearable devices to items created to automate cities (Weinberg et al, 2015). IoT devices are rapidly gaining acceptance, with some predictions of 50 billion IoT devices in use by 2020 (Weinberg et al, 2015).

Privacy and the Internet of Things

As the IoT grows, users voluntarily install sensor-equipped devices which often passively and aggressively collect data about daily activities (Lee & Kobsa, 2016). Since these devices are generally connected to the Internet, they also share data between multiple devices and services which leads to security and privacy issues (Sfar et al, 2018; Zheng et al, 2018). These issues include questions about what data is collected, where that data is stored, how the data is used, who owns the data, and who can access the data (Zheng et al, 2018).

Bannon (2016) notes that consumers often replace or upgrade home appliances with "smart" replacements but may be unaware of the data being collected and shared by these devices. Further, Bannon (2016) suggests many users fail to read privacy policies. Finally, each new IoT device installed provides yet another opening for attackers (Poudel, 2006).

METHOLODOGY

Information Privacy Concern Measurement

Within the Information Systems field, multiple instruments have been constructed to measure information privacy concern. Preibusch (2013) summarizes many of these.

Preibusch (2013) considers the Concern for Information Privacy (CFIP) instrument developed by Smith et al (1996) to be the "most influential" information privacy instrument. This instrument focuses on offline marketing and consists of four dimensions: "collection of personal information, unauthorized secondary use of personal information, errors in personal information, and improper access to personal information" (Xu et al, 2012). The CFIP scale contains a second-order factor structure (Stewart & Segars, 2002).

The CFIP instrument was modified by Malhotra et al (2004) to focus upon the information privacy concerns of Internet users. This new instrument, the Internet Users Information Privacy Concern (IUIPC) instrument, utilized three dimensions: "collection of personal information, control over personal information, and awareness of organizational privacy practices" (Xu et al 2012).

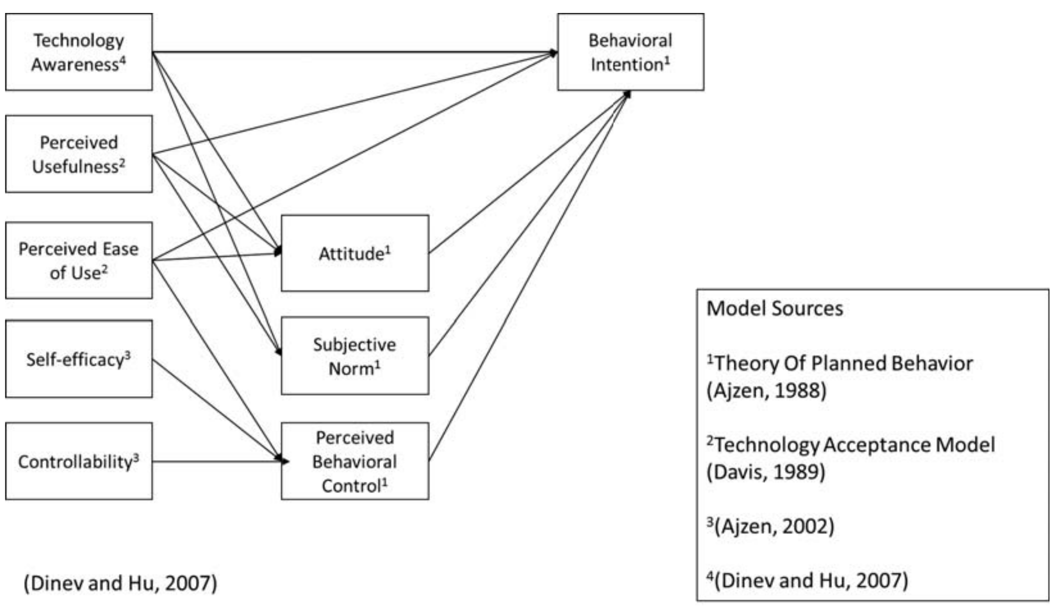
The IUIPC instrument itself was later modified to focus more upon mobile users. Xu et al (2012) based the Mobile Users' Information Privacy Concerns (MUIPC) instrument upon Communication Privacy Management theory. This instrument focuses on three dimensions: perceived surveillance, perceived intrusion, and secondary use of information (Xu et al, 2012). Other research suggests that solutions to IoT privacy concerns must consider both user attitudes and awareness (Zheng et al, 2018). A privacy instrument developed by Dinev and Hu (2007) specifically focuses upon these constructs.

The Dinev and Hu (2007) model focuses upon protective information technologies, which the authors define as "information technologies that protect data and systems from disturbances such as viruses, unauthorized access, disruptions, spyware, and others" (p. 386). The authors distinguish protective technologies from positive technologies, which "benefit their users in terms of productivity, efficiency, competitiveness, or entertainment" (Dinev & Hu, 2007, p. 387). Further, the authors suggest that awareness of threats strongly predicts user intention to use protective technologies.

Within the IoT domain, user awareness of threats appears to be low. Bannon (2016) observes that "the most dangerous part of IoT is that consumers are surrendering their privacy, bit by bit, without realizing it, because they are unaware of what data is being collected and how it is being used." This lack of awareness, coupled with user attitudes, is key to understanding the use of security and privacy settings within the IoT (Zheng et al, 2018).

The Dinev and Hu (2007) model (Figure 1) is based upon Ajzen's Theory of Planned Behavior (1988) but incorporates a number of other factors, including the Technology Acceptance Model (Davis, 1989).

Figure 1: The Dinev and Hu Model (2007)



The Theory of Planned Behavior

The Theory of Planned Behavior (TPB), which is based upon the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980), suggests that human behavior can be predicted by intention and that intention is formed from three factors: attitude toward the behavior, subjective norm, and perceived behavioral control. Attitude refers to individual evaluation about performing the behavior (good or bad), while subjective norm captures perceptions of how others feel about performing the behavior. Perceived behavioral control captures individual perceptions regarding the ease or difficulty of performing the behavior in question. Perceived behavioral control can also directly impact behavior (Ajzen, 1988).

Self-Efficacy and Controllability

Although the TPB has been widely used, perceived behavioral control has caused issues for many researchers. Ajzen (2002) suggests that self-efficacy and controllability are both components of perceived behavioral control, and that these may be measured separately. Self-efficacy refers to an individual's judgement of their ability to perform the behavior (Bandura, 1986), while controllability references the individual's judgement regarding the opportunity to perform the behavior and the availability of required resources (Ajzen, 2002).

Technology Acceptance Model

The Technology Acceptance Model (TAM) is also an extension of the TRA (Dinev & Hu, 2007). TAM suggests that two factors are key in understanding user acceptance of new technology. These determinants are perceived ease of use and perceived usefulness. Perceived ease of use captures user perceptions of the difficulty or ease of using the new technology, while perceived usefulness captures the degree to which the user believes the new technology will prove beneficial (Davis, 1989).

Technology Awareness

Technology awareness refers to "a user's raised consciousness of and interest in knowing about technological issues and strategies to deal them" (Dinev & Hu, 2007, p. 391). The authors also suggest that an individual cannot form attitudes toward protective technologies without being aware of those technologies in the first place (Dinev & Hu, 2007).

The Dinev and Hu Model

The Dinev and Hu Model encompasses the Theory of Planned Behavior, the Technology Acceptance Model and contains the attitude and awareness components that Zheng et al (2018) required for a model understanding IoT privacy. While the original model focused on spyware protection, the IoT technology aspect appears similar and suggests that the Dinev and Hu model might provide a methodology for thinking about IoT privacy. Since the current research setting differs from the initial 2007 article, all relationships and hypotheses will be examined.

Research Question: Does the 2007 Dinev and Hu awareness centric model of human behavior toward protective technologies apply to the Internet of Things?

Hypotheses

H1: Technology awareness is positively correlated with attitude toward the behavior.

H2: Technology awareness is positively correlated with subjective norm.

H3: Technology awareness is positively correlated with behavioral intention.

Users who are unaware of the privacy threats posed by IoT devices, or more specifically unaware of their ability to modify settings to enhance their own privacy will lack attitudes toward the behavior of changing those settings. These same users are also unlikely to be influenced by others regarding changing settings. On the other hand, informed users who are aware of their ability to modify settings are more likely to have a positive attitude toward doing so, are more likely to be influenced to do so by others, and are more likely to actually change those settings.

H4: Attitude is positively correlated with behavioral intention.

H5: Subjective norm is positively correlated with behavioral intention.

H6: Perceived behavioral control is positively correlated with behavioral intention.

H4, H5, and H6 are drawn directly from the TPB (Ajzen, 1988, Dinev & Hu 2007). Within the current study, individuals with positive attitudes toward modifying security settings should experience a stronger intention to do so. Also, pressure from important others should also encourage individuals to modify security settings. Finally, individuals who feel they are able to make these changes should be more inclined to do so.

H7: Self-efficacy is positively correlated with perceived behavioral control.

H8: Controllability is positively correlated with perceived behavioral control.

Ajzen (2002) suggests that both controllability and self-efficacy are components of perceived behavioral control. Individuals who feel they have the ability (self-efficacy) and resources and opportunities (controllability) to modify security settings should experience higher degrees of perceived behavioral control.

H9: Perceived usefulness is positively correlated with attitude.

H10: Perceived ease of use is positively correlated with attitude.

Although the TAM and later versions such as the Unified Theory of Acceptance and Use of Technology Model (Venkatesh et al, 2003) suggest a direct relationship between perceived usefulness and perceived ease of use and actual behavior (Dinev & Hu, 2007), the initial version of TAM (Davis et al, 1989; Dinev & Hu, 2007) incorporated attitude. H9 and H10 are drawn directly from the Dinev and Hu (2007) study. Essentially, individuals who perceive benefits associated with modifying IoT security settings should experience a positive attitude toward doing so. In addition, those who find changing these settings to be easy should also have a positive attitude toward doing so. Dinev and Hu (2007) did not find support for the relationship between perceived ease of use and attitude.

H11: Perceived usefulness is positively correlated with subjective norm.

Dinev and Hu (2007) suggest this relationship based upon work by Venkatesh and Davis (2000). Essentially this suggests that a protective technology that is perceived as being useful may lead referent others to feel the technology is beneficial.

H12: Perceived ease of use is positively correlated with perceived behavioral control.

Although these factors are quite similar, a user could consider changing security settings easy to do while feeling they lacked the ability to do so (Dinev & Hu, 2007). However, Pavlou and Fygenson (2006) noted a relationship between perceived ease of use and perceived behavioral control. Dinev and Hu (2007) confirmed this relationship.

H13: Perceived usefulness is positively correlated with behavioral intention.

H14. Perceived ease of use is positively correlated with behavioral intention.

H13 and H14 are duplicated from the TAM model (Davis, 1989) and suggest that individuals who perceive changing security settings to be useful and easy to do are more likely to change those settings. However, these were not supported in the Dinev and Hu (2007) study.

Survey Instrument

The original Dinev and Hu (2007) survey instrument was modified to focus upon the Internet of Things. Items that referenced spyware protection were reworded to address IoT.

<u>Survey Administration and Descriptive Statistics</u>

1257 students enrolled in undergraduate courses at a Carnegie Master's level university located in the mid-south were invited to complete the survey online. The survey itself was created as an Office 365 form. Students were asked to complete the survey by their normal instructors. Respondents were offered the opportunity to enter their email addresses into a separate survey form (administered by a faculty member not involved in the project) to enter a random drawing for three gift cards. 192 students completed the survey, for a response rate of roughly 15.3%.

The majority of respondents were females (60.5%) between the ages of 18 and 24 (81.1%). As expected, most respondents (68.6%) had completed some college, with almost half (47.8%) reporting business as their major. Further, most of the respondents (73.5%) reported over seven years of experience using the Internet; however, the majority (71.5%) reported less than 7 years of experience utilizing the IoT. The majority of respondents reported only infrequently being a privacy victim (24.3% responded very infrequently, while 59.4% selected the three least frequent categories). While only 13% of respondents reported having no IoT devices in their residence, 59.4% reported 1-4 such devices. Further, 57.4% of respondents utilize 1-2 mobile IoT devices, while 25.1% utilize no such mobile IoT devices. Descriptive statistics are presented in Appendix Tables 1-7.

RESULTS

SmartPLS, a structural equation modeling software, was used to validate the measurement model and test hypotheses. This software utilizes a variance-based approach to estimate parameters and a least-squares procedure to limit demands on measurement scales and sample size (Chin, 1998; Falk et al, 1992; Fornell & Bookstein, 1982). The analysis does not assume normally distributed data and utilizes a nonparametric bootstrap procedure (Davison & Hinkley, 1997; Efron & Tibshirani, 1986).

The measurement model was examined by comparing discriminant validity, convergent validity, and internal consistency reliability (ICR) to established heuristic values.

To establish discriminant validity, the average variance extracted (AVE) should exceed .707 and exceed the correlation between construct pairs (Fornell & Larcker, 1981). Also, the items should load more heavily upon their respective constructs than on others (Gefen & Straub, 2005). Within the current study, several criterion fail to meet the thresholds for discriminant validity. For example, the AVE for technology acceptance is .63, and the AVE for perceived usefulness is .70. In addition, several of the correlations between construct pairs exceeds the specified AVE values. Discriminant validity is not completely supported. Table 8 presents the ICR, AVE, AVE square root values, and correlations.

To establish convergent validity, the factor loadings should load upon the respective construct and should exceed .70 (Carmines & Zeller, 1979). In the current study, all loadings exceed .70 and load as expected. In addition, the Internal Consistency Reliability (ICR) values should exceed .70 (Fornell & Larcker, 1981) to establish convergent validity. In the current study, ICR values range from .87 to .94. Finally, the AVE values should exceed the accepted cutoff value of .50 to establish convergent validity. Within the current study, AVE values ranged from .63 to .87. As shown in Table 9, convergent validity is supported.

Next, all hypotheses were evaluated. Table 10 summaries the hypotheses, path coefficients, and t-statistics. H4, H6, H10, H12, and H13 are not significant.

Table 8: Internal Consistency Reliability, Convergent Validity, and Discriminant Validity

	ICR	AVE	A	SN	PBC	BI	SE	C	TA	PU	PEOU
A	.94	.84	**.92**								
SN	.89	.81	.66	**.90**							
PBC	.93	.87	.56	.56	**.93**						
BI	.93	.81	.56	.74	.51	**.90**					
SE	.93	.82	.53	.60	.76	.59	**.91**				
C	.91	.78	.66	.68	.78	.63	.80	**.88**			
TA	.89	.63	.61	.73	.48	.73	.58	.67	**.79**		
PU	.87	.70	.69	.72	.58	.72	.63	.70	.75	**.84**	
PEOU	.91	.77	.56	.67	.75	.68	.85	.80	.62	.71	**.88**

Internal Consistency Reliabilities (in ICR column), Average Variance Extracted (in AVE column), AVE square roots (on diagonal), and correlations (below diagonal).

A = Attitude toward the Behavior, SN = Subjective Norm, PBC = Perceived Behavioral Control, BI = Behavioral Intention, SE = Self Efficacy, C = Controllability, TA = Technology Awareness, PU = Perceived Usefulness, PEOU = Perceived Ease of Use

Discussion

Dinev and Hu (2007) found that perceived ease of use and self-efficacy were not significant within the model while examining the protective technology of anti-spyware. The authors suggest that ease of use and self-efficacy were not significant due to the differences in protective and positive technologies.

The current study focused upon altering security settings within IoT devices. As discussed earlier, the IoT presents numerous privacy concerns, and the only true method of alleviating many of these concerns (other than not using IoT devices) is altering the settings.

Of course, all results must be carefully considered, since discriminant validity was not completely supported. However, the first three hypotheses, which focused upon technology awareness, were supported. While expected, this suggests that awareness of an issue and the methods to prevent the issue are indeed important within the IoT arena. Since 87% of respondents reported having some form of IoT device within their home, these results are somewhat encouraging…awareness of issues does contribute to attitude, subjective norm, and (most importantly) to behavioral intention. Future research should explore this area more—for example, are IoT users concerned with the possibility of privacy intrusions?

Table 9: Factor Loadings

	A	BI	PBC	PEOU	PU	SN	TA	C	SE
A1	**0.8567**	0.4893	0.4771	0.4549	0.5732	0.5874	0.4778	0.5863	0.4419
A2	**0.9281**	0.4905	0.5435	0.5153	0.6326	0.5972	0.5745	0.5742	0.5046
A3	**0.9548**	0.5537	0.5259	0.5461	0.6851	0.6303	0.6107	0.6402	0.5169
BI1	0.4627	**0.9389**	0.498	0.645	0.6353	0.6658	0.6327	0.5999	0.5601
BI2	0.5778	**0.8276**	0.3476	0.4949	0.6375	0.607	0.6677	0.5037	0.4445
BI3	0.4796	**0.9295**	0.5183	0.6762	0.6787	0.7139	0.671	0.6032	0.5711
C1	0.5424	0.6264	0.6495	0.7352	0.6636	0.6333	0.5693	**0.8503**	0.6515
C2	0.6086	0.5552	0.7356	0.7173	0.5998	0.608	0.6411	**0.9195**	0.7635
C3	0.5875	0.5031	0.6846	0.6597	0.5913	0.5789	0.5566	**0.8777**	0.7052
PBC1	0.5839	0.5149	**0.9434**	0.7398	0.5913	0.5434	0.4992	0.7902	0.738
PBC2	0.4617	0.4296	**0.9239**	0.6656	0.4847	0.5111	0.3953	0.663	0.6856
PEOU1	0.4903	0.5696	0.5698	**0.7863**	0.5533	0.5164	0.5057	0.5414	0.619
PEOU2	0.4639	0.5742	0.6851	**0.9144**	0.6441	0.5981	0.5459	0.7405	0.7926
PEOU3	0.5035	0.6306	0.7205	**0.9205**	0.6513	0.6325	0.569	0.7951	0.8032
PU1	0.6276	0.5738	0.4057	0.4954	**0.8095**	0.5675	0.622	0.5658	0.4304
PU2	0.559	0.6527	0.5225	0.6809	**0.8763**	0.6593	0.6977	0.6151	0.5938
PU3	0.5465	0.5817	0.5239	0.5856	**0.817**	0.5804	0.5662	0.5656	0.561
SE1	0.4344	0.542	0.6504	0.7761	0.6289	0.5103	0.4994	0.669	**0.8419**
SE2	0.511	0.519	0.7313	0.7602	0.534	0.5666	0.5296	0.7637	**0.9357**
SE3	0.5022	0.5312	0.6871	0.7636	0.5631	0.5577	0.5514	0.7394	**0.9318**
SN1	0.5167	0.7412	0.5196	0.6511	0.6687	**0.9138**	0.662	0.6256	0.5322
SN2	0.6871	0.5767	0.4977	0.5438	0.6309	**0.8858**	0.6599	0.6089	0.5563
TA1	0.4561	0.6195	0.4328	0.5869	0.6437	0.6006	**0.8256**	0.5745	0.5073
TA2	0.4319	0.5361	0.3394	0.4014	0.5941	0.5913	**0.7959**	0.4823	0.4206
TA3	0.4102	0.5727	0.2652	0.4236	0.5156	0.4996	**0.7003**	0.4273	0.3476
TA4	0.421	0.5434	0.334	0.4515	0.5859	0.5473	**0.8287**	0.4892	0.4352
TA5	0.657	0.6118	0.5061	0.5609	0.64	0.6518	**0.8085**	0.6437	0.5687

Only one of the three TPB hypotheses (H4, was supported. The relationship between subjective norm and behavioral intention was significant, while the others were not. This is curious indeed, and simply raises more questions. Are users concerned with privacy in the IoT? If not, their attitudes toward protecting themselves by changing settings could reasonably be found insignificant.

Within the TPB, perceived behavioral control (PBC) has often been troublesome. In this study, no support was found for the anticipated relationship between perceived behavioral control and behavioral intention. However, both self-efficacy (H7) and controllability (H8) are indeed positively correlated with PBC. Further, perceived ease of use (H12) was not significantly correlated with perceived behavioral control.

Table 10: Hypotheses, Path Coefficients, and t-statistics

Hypotheses	Path Coeff.	t-statistic
*H1: Technology awareness is positively correlated with attitude toward the behavior.	0.1822	1.9791
*H2: Technology awareness is positively correlated with subjective norm.	0.4377	3.4523
*H3: Technology awareness is positively correlated with behavioral intention.	0.2615	1.9313
H4: Attitude is positively correlated with behavioral intention.	-0.0308	0.3276
*H5: Subjective norm is positively correlated with behavioral intention.	0.2981	2.594
H6: Perceived behavioral control is positively correlated with behavioral intention.	-0.0686	0.6305
*H7: Self-efficacy is positively correlated with perceived behavioral control.	0.2568	1.6717
*H8: Controllability is positively correlated with perceived behavioral control.	0.4059	3.4651
*H9: Perceived usefulness is positively correlated with attitude.	0.4821	3.7151
H10: Perceived ease of use is positively correlated with attitude.	0.102	0.9303
*H11: Perceived usefulness is positively correlated with subjective norm.	0.3924	2.9789
H12: Perceived ease of use is positively correlated with perceived behavioral control.	0.2143	1.3224
H13: Perceived usefulness is positively correlated with behavioral intention.	0.1989	1.172
*H14. Perceived ease of use is positively correlated with behavioral intention.	0.2444	1.8194

*These hypotheses are significant.

These results definitely cast doubt upon the applicability of the TPB within this setting. Of course, this could be a result of the specific survey instrument or the sample. However, the authors suspect this may simply represent the same underlying issue as mentioned above.

The remaining hypotheses (H9, H10, H11, and H12) examine two factors from the TAM model (Davis, 1989). The hypothesized relationship between perceived ease of use and attitude, (H10), PBC (H12) are not supported, although the relationship between perceived ease of use and behavioral intention is significant. The hypothesized relationship between perceived usefulness and attitude (H9), subjective norm (H11) are both supported; however, the relationship between perceived usefulness and behavioral intention is not significant. As Dinev and Hu (2007) suggested, these results may simply reflect the differences between positive and protective technologies. Respondents may simply feel that protective technologies are somewhat necessary even if less than convenient to utilize. This does not explain the unexpected results leading to intention however.

Limitations

As with any research, multiple limitations exist with this study. First, only students enrolled within a specific university were surveyed. This convenience sample may pose issues. Second, the survey instrument failed to achieve discriminant validity in a number of areas. This could be caused by refocusing the instrument on a somewhat different area. The respondents self-selected and there was a low, 15%, response rate. Thus, the study could suffer from self-selection bias and limited representativeness.

CONCLUSION

The results of this study are somewhat disappointing. The Dinev and Hu (2007) model did not perform well as used within this research. However, the results do suggest that technology awareness is critical to IoT privacy. Users who are unaware of (or unconcerned with) privacy may simply not realize that IoT devices jeopardize that privacy. Future research should explore this area.

In addition, other research could examine both the TPB and TAM within IoT to determine applicability of those instruments.

APPENDIX

Table 1: Gender

Gender	Frequency	Percent
Female	112	60.5
Male	68	36.8
Prefer not to say	5	2.7
Total	185	100

Table 2: Age

Age	Frequency	Percent
18-24	150	81.1
25-34	14	7.6
35-44	11	5.9
45-54	7	3.8
55-64	3	1.6
Over 65	0	0
Total	185	100

Table 3: Highest Education Level Obtained

Educational Level	**Frequency**	**Percent**
Some school, no degree	3	1.6
High school graduate	26	14.1
Some college, no degree	127	68.6
Bachelor's degree	28	15.1
Master's degree	0	0
Professional degree	0	0
Doctorate degree	1	.5
Total	185	100

Table 4: Major Area of Study

Major	**Frequency**	**Percent**
Agriculture	13	7.1
Behavioral Sciences	12	6.5
Biological Sciences	1	0.5
Business	88	47.8
Computer Science / Management Information Systems / Information Technology	7	3.8
Education	19	10.3
Engineering	15	8.2
Human Performance	1	0.5
Humanities	14	7.6
Mathematics and Statistics	1	0.5
Nursing / Medical / Human Performance	6	3.3
Interdisciplinary	3	1.6
Criminal Justice	3	1.6
Geoscience	1	0.5
Total	184	100

Table 5 Internet and IoT Experience

	Internet Experience		**IoT Experience**	
Time Frame	**Frequency**	**Percent**	**Frequency**	**Percent**
Less than a year	3	1.6	27	14.5%
1 – less than 2 years	1	0.5	26	14.0%
2 – less than 3 years	2	1.1	21	11.3%
3 – less than 4 years	9	4.9	29	15.6%
4 – less than 5 years	7	3.8	13	7.0%
5 – less than 6 years	12	6.5	11	5.9%
6 – less than 7 years	15	8.1	6	3.2%
More than 7 years	136	73.5	53	28.5%

Table 6: Invasion of Privacy Victim

Privacy Invasion	Frequency	Percent
1 Very Infrequently	45	24.3
2	37	20.0
3	28	15.1
4	35	18.9
5	20	10.8
6	11	5.9
7 Frequently	9	4.9
Total	185	100

Table 7: Number of IoT Devices and Mobile IoT Device Utilization

	Iot Devices in Residence		Mobile IoT Device Utilization	
Device Count	**Frequency**	**Percent**	**Frequency**	**Percent**
0	24	13	46	25.1
1-2	59	31.9	105	57.4
3-4	51	27.6	24	13.1
5-6	23	12.4	5	2.7
7-8	11	5.9	1	.5
9-10	7	3.8	1	.5
>10	10	5.4	1	.5
Total	185	100	183	100

REFERENCES

Ajzen, I. (1988). *Attitudes, Personality, and Behavior*. The Dorsey Press, Chicago IL.

Ajzen, I. (2002). “Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior.” *Journal of Applied Social Psychology*, 32, pp. 665-683.

Ajzen, I. & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall, Englewood Cliffs, NJ.

Bannan, C. (2016). “The IoT Threat to Privacy.” *TechCrunch*. https://techcrunch.com/2016/08/14/the-iot-threat-to-privacy/

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive*. Prentice-Hall, Englewood Cliffs, NJ.

Belanger, F. & Crossler, R. E. (2011). "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems." *MIS Quarterly*, 34(4), pp. 1017-1041.

Boos, D., Guenter, H., Grote, G., & Kinder, K. (2013). “Controllable Accountabilities: The Internet of Things and its Challenges for Organisations.” *Behaviour & Information Technologies*, 32(5), 449-497.

Carmines, E.G. & Zeller, R.A. (1979). *Reliability and Validity Assessment*, Beverly Hills, CA: Sage Publications.

Chin, W. W. (1998). “Issues and Opinion on Structural Equation Modeling.” *MIS Quarterly*, 22(1), pp. vii – xvi.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Vol. 1). Cambridge University Press.

Davis, F. D. (1989). “Perceived Usefulness, Ease of Use, and User Acceptance of Information Technology.” *MIS Quarterly*, 13(3), pp. 319-340.

Davis, F. D., R. P. Bagozzi, and P. R. Warshaw. (1989). “User Acceptance of Computer Technology: A Comparison of Two Theoretical Models.” *Management Science*, 35(8), pp. 982-1003.

Dinev, T. & Hart, P. (2004). "Internet Privacy Concerns and their Antecedents--Measurement Validity and a Regression Model." *Behavior & Information Technology*, 23(6), pp. 413-422.

Dinev, T. & Hu, Q. (2007). “The Centrality of Awareness in the Formation of User Behavioral Intention toward Protective Information Technologies.” *Journal of the Association for Information Systems*, 8(7), pp. 386-408.

Efron, B., & Tibshirani, R. (1986). “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy.” *Statistical science*, pp. 54-75.

Falk, R., Miller, F. & Miller, N.B. (1992). *A Primer for Soft Modeling*, Akron, OH: University of Akron Press.

Fornell, C. & Bookstein, F.L. (1982). “Two structural equation models: LISREL and PLS applied to consumer exit-voice theory.” *Journal of Marketing Research*, 19(4), 440–452.

Fornell, C. & Larcker, D. F. (1981). “Evaluating Structural Equation Models with Unobservable Measurement Error.” *Journal of Marketing Research*, 18(1), pp. 39-50.

Gefen, D. & Straub, D. (2005). “A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example”, *Communications of the AIS*, 16, pp. 91-109.

Lee, Hosub & Kobsa, A. (2016). "Understanding User Privacy in Internet of Things Environments." *IEEE 3rd World Forum on Internet of Things*, Reson VA.

Mason, R. O. (1986). "Four Ethical Issues of the Information Age." *MIS Quarterly*, 10(1), pp. 5-12.

Pavlou, P.A. & Fygenson, M. (2006). “Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior.” *MIS Quarterly*, 30(1), pp. 115-143.

Poudel, S. (2006). “Internet of Things: Underlying Technologies, Interoperability, and Threats to Privacy and Security.” *Berkeley Technology Law Journal*, 998-1021.

Preibusch, S. (2013). "Guide to Measuring Privacy Concern: Review of Survey and Observational Instruments." *International Journal of Human-Computer Studies*, 71(12), pp. 1133-1143.

Sfar, A. R., Natalizio, R, Challal, Y. & Chtourou, Z. (2018). "A Roadmap for Security Challenges in the Internet of Things." *Digital Communications and Networks*. 4(2), pp. 118-137.

Smith, H. J., Milberg S. J., & Burke, S. J. (1996). "Information Privacy: Measuring Individual Concerns About Organizational Practices." *MIS Quarterly*, 20(2), pp. 167-196.

Stewart, K. A. & Segars, A. H. (2002). “An Empirical Examination of the Concern for Information Privacy Instrument.” *Information Systems Research*, 13(1), pp. 36-49.

Venkatesh, V. & Davis, F. D. (2000). “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies.” *Management Science*. 46(2), pp. 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). “User Acceptance of Information Technology: Toward a Unified View.” *MIS Quarterly*, 27(3), pp. 425-478.

Weinberg, B. D., Milne, G. R., Andonova, Y. G., & Hajjat, F. M. (2015). “Internet of Things: Convenience vs. Privacy and Secrecy.” *Business Horizons*, 58(6), pp. 615-624.

Xu, H., Gupta, S., Rosson, M. B. & Carroll, J.M. (2012). "Measuring Mobile Users' Concerns for Information Privacy." *33rd International Conference on Information Systems (ICIS) Proceedings*, Orlando.

Xu, H., Dinev, T., Smith H. J., & Hart, P. (2008) “Examining the Formation of Individual’s Privacy Concerns: Toward an Integrative View.” *International Conference on Information Systems (ICIS) Proceedings*, Paris, pp. 1-16.

Zheng, S., Apthorpe, N, Chetty, M., & Feamster, N. (2018). “User Perceptions of Smart Home IoT Privacy.” *Proceedings of the ACM Human-Computer Interactions*. 2, CSCW, Article 200, https://doi.org/10.1145/3274469.

DECISION SCIENCES INSTITUTE

Contextualizing Fear Appeals for Improved Information Security

Jordan Shropshire
Professor
Information Systems Department
University of South Alabama
jshropshire@southalabama.edu

Philip Menard
Assistant Professor
Information Systems Department
University of Texas at San Antonio
philip.menard@utsa.edu

ABSTRACT

Fear Appeals are a common tool for encouraging individuals to adopt information security practices. A fear appeal describes a potential threat and suggests a course of action for avoiding the threat. Most fear appeals are static, generically worded messages designed for mass audiences. They lack personalization and are not sensitive to the interests, values, and preferences of individual recipients. As one of the primary tools for encouraging adoption of cyber security policies, fear appeals often have a modest effect on end user behavior. This research holds that fear appeals would be more effective change catalysts if their content was contextualized to the individual. Accordingly, this research describes a new framework for producing more relevant fear appeals. It uses machine learning to curate the most appropriate fear appeal for each individual. The efficacy of this framework is examined in an experiment involving mobile device security. Early results confirm the utility of the proposed framework. Implications and future research directions are also described.

KEYWORDS: Fear Appeal, Information Security, Machine Learning, Personalization, Context

INTRODUCTION

Convincing individuals to adopt new information security and privacy recommendations remains a hard problem within the information security field. A common approach for affecting behavioral change is to communicate via the fear appeal. Fear appeals are mechanisms for encouraging behavioral changes in the security practices of end users (Johnston et al. 2010). These messages are carefully designed to evoke fear among their recipients. Fear appeals also contain recommendations for minimizing information security and privacy risks. Many researchers have spent years experimenting with variations in the wording of the fear appeal in order to make them more persuasive (Ruiter et al. 2014). However, it appears that these efforts have had a relatively modest effect on the influence of fear appeals (Barlow et al. 2013). Individuals and organizations continue to eschew good security recommendations and leave themselves vulnerable to the malicious actions of outsiders.

The present study holds that fear appeals may be ineffective because they lack the personalization necessary to be relevant to individual recipients (Peters et al. 2012). People are increasingly deaf and blind to generic communications (Calder et al. 1980; Grass et al. 1969). As big data and mass customization continue to merge, people seek more tailored, thoughtful experiences. Enterprises which successfully curate their product recommendations to the tastes and preferences of individual shoppers are enjoying record sales growth (Peppers et al. 1997). Organizations which tailor each communication to the background and interests of each recipient are rewarded with increased awareness (Wind et al. 2001). Conversely, those who stick with the conventional one-size-fits-all approach are becoming irrelevant. Individuals have come to expect a high degree of personalization throughout their lives (Cöner 2003). Hence, should be of little surprise that a static, unitary fear appeal will have little effect on its recipients.

This research proposes a new framework for developing and delivering highly contextual fear appeals (see Figure 1). At the crux of this framework are processes for developing user insights and sending fear appeals which are personally relevant to each recipient. The proposed framework is conceptualized in terms of a front end which allows for data gathering, behavior observation, and fear appeal delivery, and a back end which performs analytics, machine learning, and message curation. The front end is implemented as a mobile device-based agent for the Android ecosystem. The back end is operationalized as a cloud-based platform which integrates a noSQL database for building user intelligence and Anaconda machine learning libraries for fear appeal curation. It uses textual analysis and machine learning to categorize individuals according to their profiles and then select the most relevant fear appeal. A data pipeline allows communication between the endpoints. The proposed framework is evaluated in a small-scale experiment in which it is pitted against the contemporary approach of the static fear appear.

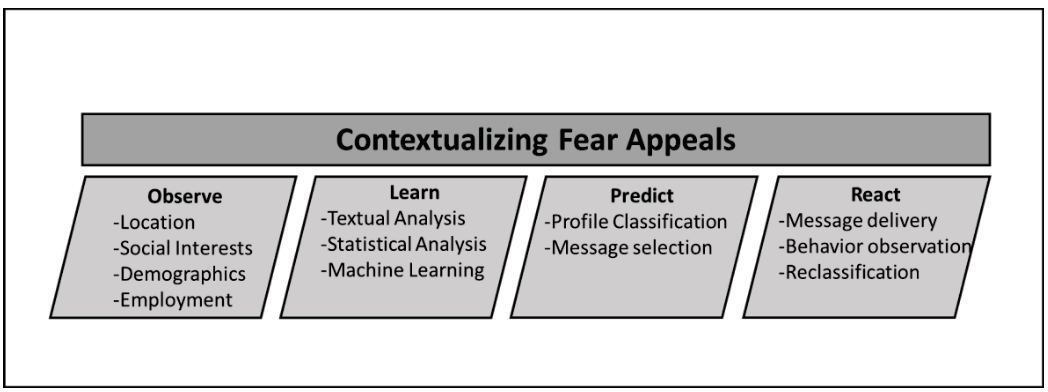


Figure 1. Creating Contextual Fear Appeals

The remainder of this paper is organized as follows. The next section contains the background. It describes the concept of message contextualization. The conceptual development section follows the background. It describes the proposed framework and explains the expected benefit. The process of testing and comparison described in the evaluation section. The results of the tests are described next. Implications for research and practice are then discussed. Finally, conclusions and future research directions are shared.

BACKGROUND

Until recently, it was widely assumed that the best way to reach a mass audience is to develop a single, clearly articulated message. Communicators sought to saturate as many channels as possible with that message in order to produce a desired effect (Belch 1982; Grass et al. 1969). This communication approach is widely seen in advertising campaigns for products and services, during election seasons, and for the promotion of social change. Indeed, change agents tasked with promoting the diffusion of new innovations are often ordered to adhere to a well-defined script for introducing the concept and persuading people to try it. This approach was widely regarded because it maximizes distribution of a consist message in a relatively short period of time. As the breadth of audience increases, the potential scope of impact also increases. Therefore, in a short period of time, it is possible to build awareness and effect change on a national or even international level (Calder et al. 1980). Hence, mass messaging is the preferred communication method for many industries and agencies.

Mass communication is very much a balancing act: the message must be broadly acceptable but it must also be vivid enough to trigger a desired response (Cho et al. 2007; Peters et al. 2012). A message which is too contextual will be misunderstood by vast swaths of the audience. Some percentage of the audience will not draw the correct conclusions. Their interpretation of the message could result in unexpected responses and undesirable behavior. Hence, mass messages should be standardized for broad interpretation (Jensen et al. 2014). However, a message which is too generic will be unmemorable and unpersuasive. People are unmoved by abstract information (Calder et al. 1980). A message must convey an individually meaningful consequence or benefit or else the recipient would have no cause to react. People respond to information which is relevant to their frame of reference. Hence, it is difficult to create a powerful message which is universally interpreted across a wide audience.

The solution lies in mass customization. This is the concept of modifying something for each member in a targeted group (Adomavicius et al. 2005; Wind et al. 2001). With mass customization of communications, it is no longer necessary to send a single message to everyone. A core message can be modified to create dozens or even hundreds of variants. This enables a finer-grained match between individual and message (Allen et al. 2001). Each person can receive the message which is most appropriate for their personal frame of reference (Peracchio et al. 1998). In the era of big data, it is now possible to rapidly develop in-depth profiles for scores of individuals and programmatically select the best message for each person (Peppers et al. 1997; Wind et al. 2001). Contextual factors, such as background, interests, goals, abilities, income, demographics, geographic location, social status, education, and career can be intelligently factored into message selection (Moon 1999). To sum, mass customization allows for massive distribution of information which is made contextually relevant for each recipient.

Business and communications professionals are aware of this capability and are employing it within their own enterprises with great success (Simonson 2005). For instance, E-commerce companies use this concept to tailor their storefront according to the tastes of each visitor (Kim et al. 1997). Some US government agencies modify the contents of their webpages according to the interests of each visitor. Even the advertising copy in postal marketing campaigns is increasingly adjusted according to the data that the sender has on each address. Message contextualization is effective because it increases the personal relevance of the information (Moon 1999). It raises the specter that a person will be persuaded to follow the recommended actions (Koch et al. 2013). To sum, organizations increasingly rely on mass customization of their communications in order to effectively reach their audiences.

CONCEPTUAL DEVELOPMENT

Given the demonstrated effectiveness of message customization in other fields, it makes sense to apply this concept to fear appeals (Ruiter et al. 2014). Fear appeals work on the principle that people will follow a recommended precaution in order to mitigate the risk of a potential threat (Rogers 1975). Fear appeals which are personally relevant are more apt to be persuasive (Roskos‐Ewoldsen et al. 2004). They draw a closer link between the threat and the personal impact of the loss. For instance, telling someone that they are susceptible to a potential risk is not as effective as telling someone what they personally stand to lose if something bad happens (Barlow et al. 2013). The increase in contextual relevance creates a more salient threat and loss. In turn, this amplifies the possibility that the individual will follow the recommendations associated with the fear appeal (Moore et al. 1996). Hence, it makes sense to send each person the most contextually relevant fear appeal as possible.

A framework for selecting and delivering contextualized fear appeals is herein proposed. The framework is comprised of two parts: a front-end host-based agent and a back-end cloud-based service. A data pipeline connects the endpoints. An outline of this arrangement is depicted in Figure 2 (below).

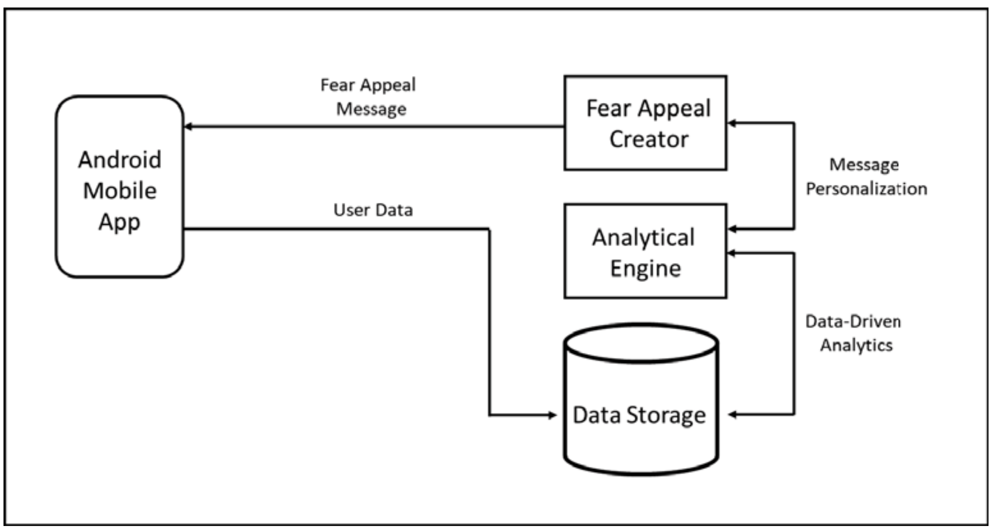


Figure 2. Proposed Framework

In general, the framework works as follows: the front-end collects data and sends it to the backend for analysis. The backend performs a number of analytical processes in order to gain insights into each individual. Individuals are sorted into one of many predetermined profiles using textual analysis, principal components analysis and Naïve Bayes classifiers. The pre-determined profiles are derived from marketing personas. Textual analysis is performed to transform string attributes into features which are amenable to quantitative analysis. Principal components analysis converts dozens of personal attributes into a smaller subset of independent features. The Naïve Bayes classifier assigns individuals to profiles based on the similarities between their principal components and the principal components associated with each profile. The front-end also observes the device's security and privacy posture. This information is also relayed to the backend. When it is time to deploy a new fear appeal the backend determines which users are susceptible to the potential security threat. Each individual in this subset is sent a fear appeal which is unique to the profile associated with their bin. The front-end displays the fear appeal and observes response behavior. The response is delivered to the backend for future analysis.

Based on the outcomes of studies in related fields, it is predicted that contextualizing the contents of fear appeals at the individual level will result in a higher rate of adoption of the associated recommendations (Adomavicius et al. 2005; Cöner 2003; Kim et al. 1997; Peppers et al. 1997). Further, it is expected that fear appeals contextualized for information security and privacy recommendations will be especially effective. It is assumed that individuals will be surprised to find out how much can be learned about them from a few pieces of unguarded information (Jensen et al. 2014). Thus, the following hypothesis is offered:

H_1: Individually contextualized fear appeals will be more effective than generic fear appeals.

EVALUATION

A small-scale experiment was performed in order to evaluate the efficacy of individually contextualized profiles. In general, the experiment was conducted as follows:

One hundred US-based subjects over the age of 18 were recruited via Amazon Mechanical Turk. An inducement of $2 US dollars was offered for fully participating in the experiment. Each subject was asked to install an app on their mobile phones. Those who complied with this requirement were randomly assigned to either a test group or a control group. A round-robin approach to group assignment was undertaken in order to assure equality of group size. Each subject entered a provided identification code into the mobile app. This identifier was relayed to the backend, which in turn linked the participant back to the tracking system within Mechanical Turk.

The proposed two-part framework was fully operationalized. The front-end consists of an Android Oreo-based application which performed the duties described in the previous section. The backend is instantiated as a series of micro-services within the Amazon AWS PaaS ecosystem. The service was built for the Ubuntu Linux operating system. In addition, a MongoDB NoSQL database was developed for data management. The front-end was coded in JAVA with Android libraries while the backend was constructed using python 3.7 and the Anaconda collection of machine learning libraries (e.g., Weka, pykit, scikit, pandas, etc.).

Once the experimental groups were fully formed, a notification was sent to each app to perform initial data collection. This data was pipelined to the back-end apparatus as WIFI permitted. From there a number of analytical processes were performed to assess device security and privacy settings, build user insights, and appropriately classify each individual. A security recommendation which would be applicable for all participants was then identified. The recommendation was to disable a feature in the Android Oreo operating system which defaulted to sharing user location when Bluetooth was also enabled. (Note: Following this experiment the principal investigator reported this bug to the Android team at Google, which subsequently issued a corrective patch.)

Those subjects in the control group received a standard fear appeal while those in the test group received a fear appeal which was contextualized to the profile they were associated with. A total of 25 a priori profiles were created. Accordingly, 25 different fear appeals were also created (one fear appeal was contextualized for each profile). The 46 subjects in the test group were binned into 18 profiles. Thus, 18 different fear appeals were used for the test group. The 45 members of the control group received a fear appeal which was designed to balance persuasiveness with broad applicability. The subjects received their fear appeals approximately 24 hours after they installed the mobile application.

After displaying the appropriate fear appeal, the front-end then polled the device settings and security settings every hour for 24 hours in order to determine if the subject complied with the recommendation. On success, the front-end sent a report to the backend to appraise it of the result. After 24 hours of polling the front-end released an alert message indicating that the experiment was over, and the subject could uninstall the app.

RESULTS

Following the experiment, the demographic data and test results were inputted into a spreadsheet for further analysis. The demographics indicate that the subjects skewed towards a younger age although gender was more evenly distributed. These data are illustrated in Table 1 (below). It should be noted that the previously developed profiles were defined in terms of 96 attributes. However, a principal components analysis was performed to transform the attributes into 7 principal components based on 40 significant features (Table 2). This preprocessing step is essential for ensuring the classifier's strong assumption of independence. The Naïve Bayesian classifier developed parameter estimates using a set of 111 training identities which were previously classified into 25 classes. The results of training indicated an 88.3% overall accuracy and a Kappa Statistic of .64. These results are further summarized in Table 3. The trained algorithm classified experimental subjects according to the probability that they belonged to a given profile. Subjects were sent the corresponding fear appeals and their behavior was noted. The effectiveness of the contextualized fear appeal is compared against the effectiveness of the standard fear appeal using a T-test of significant differences in adoption rate (Table 4). The results of the test support the hypothesis that contextualization provides a meaningful improvement in user response.

Age	18-24	25-29	30-39	40-49	50-59	60+
	26	31	22	11	1	0
Gender	Male	Female	Other			
	53	38	0			
Ethnicity	White	Black	Hispanic	Asian	Am. Indian	Other
	72	4	3	11	0	1

Table 1. Demographics

Feature	PC1	PC2	PC 3	PC4	PC5	PC6	PC7
Demographics							
Age	.87						
Gender	.72						
Ethnicity	.61						
Health	.51						
Relationship							
Current Status		.84					
Relat. Trajectory		.76					
Affinity		.73					
Activity Level		.71					
Fidelity		.67					
Marriage		.66					
Domestic Arrange.		.57					
Divorced		.51					
Geographics							
Present location			.81				
Home			.79				
Proximity to Family			.72				
Business Travel			.69				
Travel frequency			.64				
Vacation frequency			.53				
Preferred Destination			.49				
Career							
Current Position				.71			
Prestige				.68			
Previous				.67			
Work Classification				.66			
Employment Status				.65			
Side Hustle				.58			
Upward mobility				.51			
Open to new				.50			
Social Interests							
Participation Level					.67		
Network size					.65		
Influence					.63		
Interest Category I					.58		
Interest Category II					.56		
Interest Category III					.53		
Education							
Attainment Level						.65	
Major/Concentration						.65	
Major Prestige						.63	
University						.54	
University Prestige						.52	
Routine							
Active Periods							.56
Exercise							.51
Eigenvalue	3.66	3.35	3.08	2.47	2.13	1.54	1.21
Percentage of Variance	18.44	12.72	10.69	10.46	8.43	8.40	4.73
Cumulative Per. of Var.	18.44	31.16	41.85	52.31	60.74	69.14	73.87

Table 2. Results of Principal Components Analysis

Overall Accuracy	88.3%
Kappa Statistic	.64
Mean Squared Error	.198
Root Mean Squared Error	.445
Relative Absolute Error	52.4%
Root Relative Absolute Error	92.8%
Total Number of Instances	111

Table 3. Naïve Bayes Classifier Results

	Levene's Test		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Err. Diff.	95% Confidence Interval	
								Lower	Upper
Equal variance assumed	1.332	.217	2.163	8	.023	6.17	1.49	3.39	10.28
Equal variance not assumed			2.163	6.91	.024	6.17	1.49	3.37	10.32

Table 4. t-Test of Significant Differences

IMPLICATIONS

This initial study was conducted on a small scale. Further, the proposed framework is limiting in that it is only considers one machine learning approach. A large-scale study would provide more definitive results while a variety of classification approaches would clarify the role of machine learning in message selection. These issues can be addressed in future research. Despite these issues, the present study clearly indicates that message contextualization provides a significant benefit over traditional, generalized messages. Fear appeals which are aligned with each individual's frame of reference present a more realistic threat than those which are crafted for a general audience. The next logical step in this research stream is to augment the proposed model by incorporating the role of behavioral response in follow-up messages. Continually refining user classification will allow for even closer matches between individuals and fear appeals.

CONCLUSIONS

Fear appeals are common tool for eliciting behavioral change. Within the information security research field, this topic has been well-explored. However, most of the research has centered on developing message for broad audiences. However, in the age of mass-customization people expect more relevant and tailored experiences. Hence, this study proposed a smarter framework for effecting behavioral change. It suggests a two-part approach to first building individual insights and then selecting and delivering the most relevant fear appeal to each individual. The proposed framework was evaluated using a small experiment. In this experiment the prevailing techniques were compared against the new framework in order to determine the relative effectiveness of contextualization. The results confirm that contextualized fear appeals are significantly more effective than generic fear appeals. This research represents an early first step in improving the effectiveness of fear appeals. Future research should build on this method by tailoring follow-up messages according to the responses of recipients.

REFERENCES

Adomavicius, G., and Tuzhilin, A. 2005. " Personalization Technologies: A Process-Oriented Perspective," *Communications of the ACM* (48:10), pp 83–90.

Allen, C., Kania, D., and Yaeckel, B. 2001. *Internet WorldGuide to One-to-One Web Marketing.*, (Wiley: New York.

Barlow, J., Warkentin, M., Ormond, D., and Dennis, A. 2013. "Don't make excuses! Discouraging neutralization to reduce IT policy violation," *Computers & Security* (39:B), pp 145-159.

Belch, G. 1982. "The effects of television commercial repetition on cognitive response and message acceptance," *Journal of Consumer Research* (9:1), pp 56-65.

Calder, B., and Sternthal, B. 1980. "Television commercial wearout: An information processing view," *Journal of Marketing Research* (17:3), pp 173–186.

Cho, H., and Salmon, C. 2007. "Unintended effects of health communication campaigns," *Journal of Communication* (57:1), pp 293–317.

Cöner, A. 2003. "Personalization and Customization inFinancial Portals.," *Journal of the American Academy of Business* (2:2), pp 498–504.

Grass, R., and Wallace, H. 1969. "Satiation effects of TV commercials," *Journal of Advertising Research* (9:1), pp 3-8.

Jensen, J., Carcioppolo, N., King, A., Scherr, C., Jones, C., and Niederdeppe, J. 2014. "The cancer information overload (CIO) scale: Establishing predictive and discriminant validity," *Patient Education and Counseling* (95:3), pp 90-96.

Johnston, A., and Warkentin, M. 2010. "Fear Appeals and Information Security Behaviors: An Empirical Study," *MIS Quarterly* (34:3), pp 549-566.

Kim, J., and Rubin, A. 1997. "The Variable Influence of Audience Activity on Media Effects," *Communication Research* (24:2), pp 107-135.

Koch, T., and Zerback, T. 2013. "Helpful or harmful? How frequent repetition affects perceived statement credibility," *Journal of Communication* (63:4), pp 993-1010.

Moon, Y. 1999. " Interactive Technologies and RelationshipMarketing Strategies," *Harvard Business School Cases* (1), pp 1-12.

Moore, D., and Harris, W. 1996. " Affect Intensity and the Consumer's Attitude toward High Impact Emotional Advertising Appeals," *Journal of Advertising* (25:2), pp 37-50.

Peppers, D., and Rogers, M. 1997. *Enterprise One-to-One:Tools for Competing in the Interactive Age*, (Currency Doubleday: New York.

Peracchio, L., and Luna, D. 1998. "The Development of an Advertising Campaign to Discourage Smoking Initiation among Children and Youth," *Journal of Advertising* (27:3), pp 49-56.

Peters, G., Ruiter, R., and Kok, G. 2012. "Threatening communication: a critical re-analysis and a revised meta-analytic test of fear appeal theory," *Health Psychology Review* (7:1), pp 8-31.

Rogers, R. 1975. "A Protection Motivation Theory of Fear Appeals and Attitude Change," *The Journal of Psychology* (91:193-114).

Roskos-Ewoldsen, D., Yu, J., and Rhodes, N. 2004. "Fear appeal messages affect accessibility of attitudes toward the threat and adaptive behaviors," *Communication Monographs* (71:1), pp 49-61.

Ruiter, R., Kessels, L., Peters, G., and Kok, G. 2014. "Sixty years of fear appeal research: Current state of the evidence," *International Journal of Psychology* (49:2), pp 63-70.

Simonson, I. 2005. " Determinants of Customers' Responses to Customized Offers: Conceptual Framwork and Research Proposition," *Journal of Marketing* (69:2), pp 32-45.

Wind, J., and Rangaswamy, A. 2001. "Customerization: The next revolution in mass customization," *Journal of Interactive Marketing* (15:1), pp 13-32.

DECISION SCIENCES INSTITUTE

How to Achieve a Seasoned Cybersecurity Workforce

Patricia A. McQuaid
California Polytechnic State University
Email: pmcquaid@calpoly.edu

Stephanie Cervantes
Secure Smart Solutions
Email: stephanie@securesmartsolutions.com

ABSTRACT

Ransomware, phishing schemes, voting improprieties, and data breaches are increasingly challenging our infrastructure and privacy. We need to be positioned to fight off these cyberattacks and improve the quality of our security. But there is a severe shortage of cybersecurity workers in the United States that puts our digital privacy and infrastructure at risk. One of the big problems is in finding qualified people to fill these jobs. This paper discusses these issues and suggests a range of options to both attract new employees and keep those we have.

KEYWORDS: Cybersecurity, Workforce Development, National Initiative for Cybersecurity Education (NICE), Education, Training

INTRODUCTION

"Cybersecurity workers protect our most important and private information, from bank accounts to sensitive military communications" (Cyberseek, 2019). Cyberseek is a project that is supported by the National Initiative for Cybersecurity Education (NICE), a program of the National Institute of Standards and Technology (NIST) in the U.S. Department of Commerce. Among other important services, they have created an "interactive heat map" that provides a granular snapshot of demand and supply data for cybersecurity jobs at state and metro area levels, used to grasp the challenges and opportunities facing regional cybersecurity workforces (www.cyberseek.org/heatmap.html).

CYBERSECURITY SKILLS SHORTAGE

As of June 2019, in California Alone, there are more than 36,000 cybersecurity job openings, with more than 313,000 openings nationwide (Cyberseek, 2019). The data reflects the number of online job listings for cybersecurity-related positions from September 2017 through August 2018. We need to address the cybersecurity gaps identified by the current work shortage of cyber professionals, while preparing the public sector to defend itself against critical infrastructure attacks as identified by the Department of Homeland Security (DHS).

Mr. Jonas Prising, Chairman & CEO of ManpowerGroup emphasized the strategy of developing your own talent in an organization, in addition to hiring. He stated that, "With record talent shortages around the world, employers should shift their focus from just in time hiring strategies to becoming builders of talent for today and tomorrow. Developing the right blend of people, skills, processes and technology is the only way to execute your business strategy, create value and improve people's lives." (ManpowerGroup, 2018)

Reasons for the Shortage

So why is there such a shortage? There are several reasons for this problem, including that cybersecurity is a relatively new field. As a result, there is no current standardization on training or curriculum, although the NICE framework has done great work to address this issue (see below). Not only are there not enough programs to train people, there are not enough candidates applying. Part of the reason for that is that education is not started early enough.

Cyber awareness should begin in elementary school, not waiting for high school or college or trade school. There are a number of colleges that have cybersecurity programs, and many that are working with industry and government programs to give students practical experience before they graduate. There are also a number of cyber-related challenges and competitions for high-school students. One initiative that one of the authors is affiliated with is described in the reference, (McQuaid, et al, 2019)

<u>The Traditional Pathway Is Not Working</u>

Currently, women only make up 14% of the US cybersecurity workforce, so there needs to be an increased emphasis of training and hiring women. In the recent 2017 Global Information Security Workforce survey conducted by the Center for Cyber Safety and Education (Center, 2017), two thirds of the 20,000 respondents said they lack the number of cybersecurity personnel needed to combat today's threat. They also found that 30% of cybersecurity professionals have a background outside of IT, and that 33% of cybersecurity executives came from non-technical careers. So, we need to increase the number of cyber professionals and make people aware that they do not need to have a degree in computer science to be effective in the cybersecurity field.

<u>Other Reasons Why</u>

There is a shortage of qualified, seasoned professionals to teach or mentor. NICE stands for National Initiative for Cybersecurity Education, and the NICE Framework is "is a national focused resource that categorizes and describes cybersecurity work. The NICE Framework establishes a taxonomy and common lexicon that describes cybersecurity work and workers irrespective of where or for whom the work is performed. The NICE Framework is intended to be applied in the public, private, and academic sectors." Also, traditional computer science programs at postsecondary institutions align with only the Operate and Maintain category in the NICE Framework (NICE, 2019), while there are 6 other categories, as listed in Figure 1.

Figure 1: NICE Framework

Also, many traditional programs don't bring the soft skills (e.g., analysis, business knowledge, compliance & regulatory knowledge, ethics, some basic legal knowledge, troubleshooting, program management, general management, etc.) into their programs.

REASONS CYBERSECURITY WORKERS LEAVE THEIR JOBS

Using the David Letterman "Top 10" framework, below are the "Top 10" Reasons Cybersecurity Workers Leave Their Jobs (Figure 2). This list is based on an article written by Steve Morgan (Morgan, 2016), and also based on knowledge from the authors.

Figure 2: "Top 10"

#1 Reason for Leaving - They are Bored.

While there are numerous reasons for being bored, it is possible that they are mis-assigned for their Knowledge, Skills, and Abilities (KSAs) or they haven't received training, and/or they cannot stand doing their current job.

#2 Reason for Leaving - They feel Under Appreciated

We all know this one!

#3 Reason for Leaving - They feel Underpaid.

Oftentimes people think the grass is greener someplace else, so want to try something different. They may also be overestimating their own knowledge, skills, and abilities (KSAs). It could be that they really are underpaid, or think they are underpaid because they don't count their perks and retirement plans as their actual monthly compensation.

#4 Reason for Leaving - They Want to Switch Industries.

As Steve Morgan wrote, you might want more of a challenge in a different industry. He wrote, “If you're an experienced cyber pro working in a field that isn't a top hacker target, then maybe you should be.”

#5 Reason for Leaving - They Want Upward Mobility

Perhaps there is not room to progress upward in your organization, after all, there is only one CISO at a company. But you could create more of a career path for your employees.

#6 Reason for Leaving - Their Company has Lousy Security

Perhaps your current employer has their defenses down and they are practicing poor cybersecurity -- not allocating enough budget, not training employees on security awareness, not engaging senior management, and basically a sitting duck.

#7 Reason for Leaving - They Want to Master New Skills

Perhaps you are reading about IoT and mobile security, cyber forensics, automotive security, AI and other cutting-edge technologies that excite you -- but there's no clear path to get hands-on experience in those areas with your current employer.

#8 Reason for Leaving - They have a long commute

Perhaps your employees would like to work remotely, but policy prohibits it, or their work schedule is not flexible – making them commute at peak times.

#9 Reason for Leaving - They don't like or get along with their boss, co-workers, etc.

Enough said, there.

#10 Reason for Leaving - They won't say........

In conclusion, there are always a variety of factors that cause employees to want to change jobs, so if you want to retain your talent, be proactive and be sure to communicate with them and help them achieve their career goals.

So, How Do We Move Forward? Desire vs. Reality

Recognize that you already have people in your organization that want to move into cybersecurity and who have a talent for it. Be realistic about the talent in the marketplace and what your organization can acquire. Realize that you do not need the “perfect” cybersecurity person, and you might not find them nor be able to afford them. Instead, consider growing cybersecurity professionals and retain them.

“While cybersecurity education is maturing and improving at all levels, there is still work to do, including attracting young students to cybersecurity careers”, says David Shearer, CEO of

(ISC)[2] Inc., a global, not-for-profit that educates and certifies information security professionals throughout their careers (Collett-2, 2017).

Skills Needed Knowledge of Technology; Knowledge of Business

Cybersecurity is a hybrid of IT and business knowledge and skills. Someone can be good at Information Technology (IT) functions, but fail at cybersecurity. In terms of the knowledge of IT, it is best to have overall knowledge of all functions security touches, which is most areas of your business. It is important to learn how to work cooperatively within IT and which area to bring into security solutions or policy development sessions.

Having knowledge of how businesses operate is also a strength in a cyber-employee. For some job functions, it is good to have specific technical skill sets, but we also need people that have the breadth of knowledge. Having the ability to work well across the organization – not just within IT, is an asset. Knowing when to include parts of the organization and how to communicate across all realms in their own language can help improve overall security. Having basic business knowledge will give you an edge when talking to the program/business side of the organization. Let us not forget risk management, knowing how to identify and propose ways to mitigate risks.

Additional Factors

Using the full NICE Framework to align and standardize curriculum will help ensure students build the knowledge, skills and abilities needed.

It is very helpful to have a love for security. Do not select it just for the paycheck or because it pays more..., make it your passion!

Do you like puzzles and solving them? Good cybersecurity professionals have an “And then What?” inquisitive nature, and do not take things at face value. They have the ability to dig and piece the puzzle together – calling on things they may have seen weeks ago. You need ethics, along with some basic legal knowledge -to draw the line between White Hat vs Black Hat and what is Grey Hat.

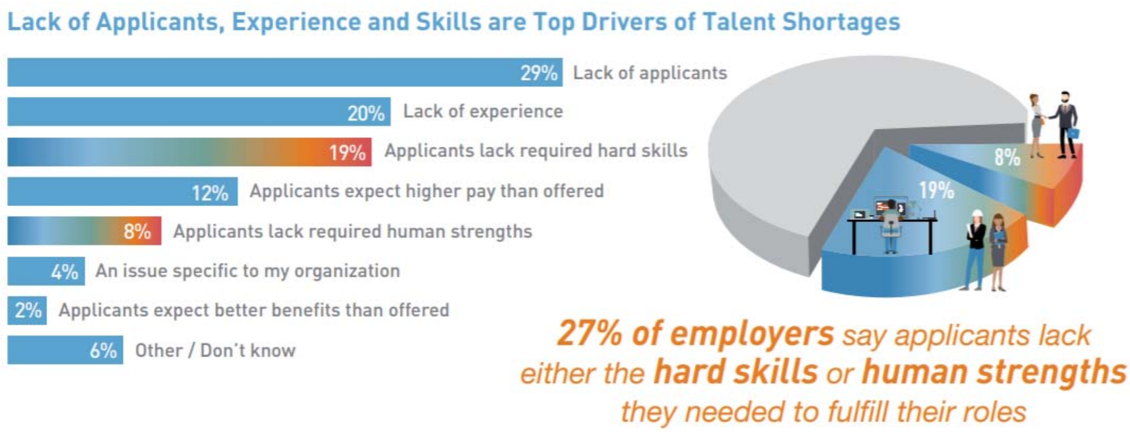
2018 MANPOWER REPORT, “SOLVING THE TALENT SHORTAGE”

In 2018, the Manpower Group issued a report entitled: “Manpower Report: Solving the Talent Shortage” (ManowerGroup, 2018). Below is are some of the key findings and a discussion of how we can apply the lessons learned to the cybersecurity industry. They found that the talent shortages have reached a 12-Year high and that the talent shortages are growing around the world.

Why Employers Can’t Find the Talent They Need

Nearly one-third of employers say the main reason they can't fill vacant positions is a lack of applicants. Another 20% say candidates lack the necessary experience. More reasons are detailed in Table 1 below.

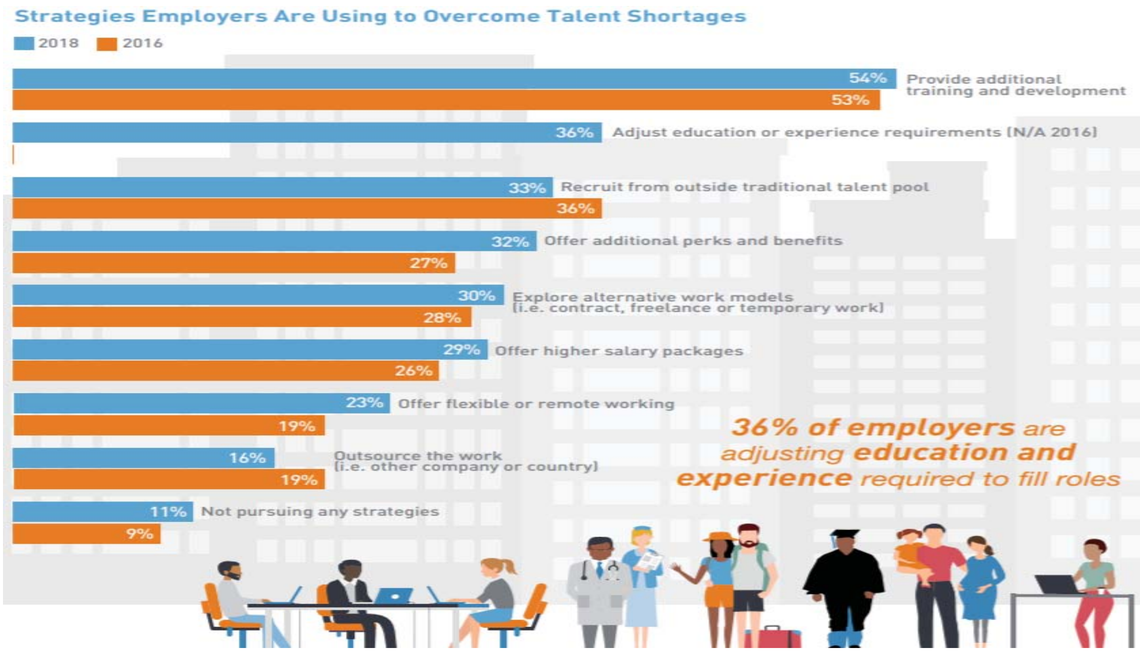
Table 1: Why Employers Can't Find the Talent They Need



The Talent Shortage Conundrum: What Employers Are Doing Today

Companies are beginning to tackle talent shortages by upskilling their own workforce: over half are investing in learning platforms and development tools to build their talent pipeline, up from just 20% in 2014. The data is listed in Table 2. This approach leads into the next section of this paper.

Table 2: Strategies Employers are Using to Overcome the Talent Shortage



Build, Buy, Borrow and Bridge: Overcoming Talent Shortages in the Future

The ManpowerGroup study emphasizes that employers are more active than ever in applying a mix of workforce strategies to address talent shortages. They have come up with a very interesting approach that can apply to many fields, suggesting that to win in the digital age an effective talent strategy should have four parts: build, buy, borrow and bridge. “Build your talent pipeline, buy skills where necessary, borrow from external talent sources and bridge people with adjacent skills from one role to another to complement existing skills.”

Build: Invest in learning and development to grow your talent pipeline.

There are several ways to build a cybersecurity team and deter them from leaving.
Put the Right Candidate in the Right Job. One way is to put the right candidate in the right job. Someone who has advanced skills will not be happy doing Account Provisioning, spending all day on Log Reviews or sitting in meetings – they will get bored and leave. Putting the wrong person in a job over their KSAs will make them frustrated, cause them to make errors, possibly create havoc or take down a network and put them on the Network Team’s and/or Management’s Most Hated or Useless List. And, instead of asking for help, they will find another job and leave.

Invest in them with training and create individual employee development plans. Do cross-training within the team with both hard and soft skilled members and mentor them. Include a budget for outside training and conferences, and allow time for webcasts and other video training during the workweek. Encourage staff to join professional organizations to find outside mentors and colleagues to learn from, and to always keep learning. Things change daily in cybersecurity; the day they quit learning is the day they should retire or move on.

Remember that you are building a workforce and not just a worker. Have senior team members or seasoned security pros (could be contractors) mentor others on the team and do knowledge transfers.

Michael Eisenberg, vice president in the Office of the CISO at Optiv Technologies, cautions that “You don’t want to bring somebody over [to a security role] if they don’t have an interest in security. It’s a different animal from IT”, he says (Collette-1, 2017).

<u>How to Find Cybersecurity Talent in Your Own Organization</u>

Ms. Stacy Collette has some suggestions for locating talent in your own organization (Collette-1, 2017). First, lower your expectations-look outside of traditional IT experience to fill cybersecurity roles. Also, be realistic of who your organization can actually attract; if you find the super qualified cybersecurity professional, you may not be able to afford them. Second, consider mid-and-later career-change employees. "There are many mid-and later-career IT pros who are not considered for redevelopment into cyber roles simply because they’ve been around too long”.

Mark Coleman, research director at Gartner indicates that organizations have become overly ambitious in their job descriptions that profile the ideal candidate, Coleman says. Companies must “free up your demands, talk to your HR department and unplug some of the requirements, such as [requiring] a degree in computer science or x number of years of information security experience,” Coleman says. Alternatively, “take a look at people in the process of achieving qualifications,” he says.

Another benefit -- “Those folks who want to make the jump can understand how their day-to-day responsibilities work right now, and how they can apply them to the security side”, says John Masserini, CISO at MIAX Options Exchange.

"There are many mid- and late-career IT pros who are not considered for redevelopment into cyber roles simply because they’ve been around too long,” Coleman says. In reality, their depth and breadth of experience could translate well into a security role.

You can also do mentoring or volunteering to grow your workforce. Yu can encourage and support coding clubs, youth hackathons and cyber challenges, and foster cybersecurity camps. A different building block could be to reach out to ex-military employment groups.

Buy: Acquire Talent.

Go to the external market to find the best talent that cannot be built in-house in the timeframe required. Consider hiring contractors to do the work or augment your staff. Be sure you require them to do knowledge transfer to your staff. You might also pay a little more to get seasoned talent. Additionally, you could invest in in-house and outside training. Be sure to choose appropriately to make sure it’s the right fit to keep an employee in place, and/or to grow them.

Borrow: Connect with Professional Groups.

Cultivate communities of talent outside of the organization, including part-time, freelance, contract and temporary workers to complement existing skills. There are numerous professional groups, such as (ISC)², ISACA, ISSA, WiCys, InfraGard, WITI, ASIS, and others...

Doing so allows your staff to connect with people who have other skill sets that can be called upon in a pinch or for mentorship. Professional groups are also a good way to remain current and to obtain Continuing Professional Education (CPEs) for their certifications.

Bridge: Create Professional Development Plans.

Help people move on and help them move up to new roles inside or outside of the organization. You can help to bridge the gap by creating professional development plans with your employee's input. You can set goals and training to grow them in their current jobs, and also for those roles they aspire to. You can utilize the NICE Framework to assist in making the plan step by step, and not an unrealistic leapfrog. You can encourage help them study and earn professional certifications.

CONCLUSIONS

It is imperative that we make some significant changes to our approach of training and hiring the cybersecurity workforce. We need to start cyber-awareness education much earlier in the lives of the upcoming workforce and cultivate an interest in cybersecurity among more than computer science students. Utilizing the NICE Framework to train, develop and hire cybersecurity professionals will most certainly help. And, we should all recognize that we may have more cyber-talent and resources in our organizations, than we might think!

REFERENCES

Center for Cyber Safety and Education. (2017). 2017 Global Information Security Workforce Study, https://iamcybersafe.org/wp-content/uploads/2017/06/Europe-GISWS-Report.pdf

Collett, S. (2017). 4 Places to Find Cybersecurity Talent in Your Own Organization, April 5.
https://www.csoonline.com/article/3186470/it-careers/hiding-in-plain-sight-4-places-to-find-cybersecurity-talent-in-your-own-organization.html#tk.cso_fsb

Collett, S. (2017). Five Ways the U.S. is Educating Cybersecurity Talent – and What's Still Missing, April 26.
https://www.csoonline.com/article/3191649/it-careers/five-ways-the-u-s-is-educating-cybersecurity-talent-and-what-s-still-missing.html#tk.csoendnote

CyberSeek. (2019). Cyberseek, An Interactive Resource for Cybersecurity Career Information.
www.cyberseek.org

The California Community Colleges Centers of Excellence for Labor Market Research. (2018). Cybersecurity: Labor Market Analysis and Statewide Survey Results.
http://business.ca.gov/Portals/0/Files/CASCADE/cybersecurity-labor-market-analysis.pdf

ManpowerGroup. (2018). 2018 Talent Survey: Solving the Talent Shortage - Build, Buy, Borrow and Bridge.
https://go.manpowergroup.com/hubfs/TalentShortage%202018%20(Global)%20Assets/PDFs/MG_TalentShortage2018_lo%206_25_18_FINAL.pdf?t=1538944138917

McQuaid, P., Britton, B., Minnich, M., Borrelli, D., Baker, J., Burton, B. (2019). University and Government Unite to Address Homeland Cybersecurity Issues. Software Quality Professional Journal, June 2019, Volume 20, Issue 3, pp. 26-32. http://asq.org/pub/sqp/past/vol21_issue3/index.html

Morgan, S. (2016). Top 5 Reasons to Quit Your Cybersecurity Job, CSO, November 2. https://www.csoonline.com/article/3137722/it-careers/top-5-reasons-to-quit-your-cybersecurity-job.html

National Initiative for Cybersecurity Education (NICE), a program of the National Institute of Standards and Technology in the U.S. Department of Commerce. (2019) NICE Cybersecurity Workforce Framework. https://www.nist.gov/itl/applied-cybersecurity/nice/resources/nice-cybersecurity-workforce-framework

Schwantes, M. (2017). The Most Effective Way to Keep Star Employees from Leaving, March 15. https://www.inc.com/marcel-schwantes/first-90-days-how-to-keep-your-best-employees-from-quitting.html

DECISION SCIENCES INSTITUTE

Using automation to grade student assignments submitted in Microsoft Excel

Cody Baldwin
Brigham Young University-Hawaii
Email: cody.baldwin@byuh.edu

James Watkins
Brigham Young University-Hawaii
Email: james.watkins@byuh.edu

Brent White
Brigham Young University-Hawaii
Email: brentwhite@byuh.edu

ABSTRACT

Across business, Microsoft Excel has become an essential skill. For this reason, it is used in a variety of university courses. However, grading assignments in Excel can be tedious and time-consuming. Automation is a solution to this problem. This paper outlines a tool that automates the grading of these assignments.

KEYWORDS: Automation, Student Assessment, Microsoft Excel, Visual Basic for Applications (VBA), Macros

INTRODUCTION

According to a recent report by a labor market analysis firm, skills in Microsoft Excel are critical to accessing higher-paying jobs in the United States (Burning Glass Technologies, 2015). Employers are demanding these skills. In response, universities, and business schools in particular, are incorporating Excel into their courses and assignments. This is a positive change, but it requires unique, and sometimes more time-consuming grading methods. Automation can be a solution this problem.

Across business (not only manufacturing), automation has added value by removing humans from repetitive, mundane tasks and shifting them to more important, value-added activities. Automation can be used in academia to fulfill the same objectives. When automated grading is implemented, faculty members and teaching assistants can spend more time on improving the curriculum, coaching students, and conducting research.

With all of that being said, the main question that guides this work is as follows: *Can business departments at universities save time and resources by automating student grading of Microsoft Excel assignments while maintaining (or possibly improving) student satisfaction?* In order to make that determination, we developed an automated grading tool called EAGR, and data was collected from courses, where this tool was implemented, that includes (1) the time required to perform manual and automated grading, (2) the positive and negative effects of shifting to the automated method of grading, and (3) student satisfaction scores that were collected before and after the tool was implemented.

LITERATURE REVIEW

This experiment was not the first attempt to use automated tools to grade student assignments. Several studies have been documented on the topic. It is becoming increasingly popular as student-to-teacher ratios climb at universities across the world (Kovacic, 2012). Also, as online education becomes more popular, universities are looking to automation in order to scale their online programs (Geigle, 2016). Most of the tools and methodologies that have been developed and published were for classes in computer science or business management.

The objectives of previous studies varied. As an example, some sought to improve the quality and quantity of feedback that was provided to students (Matthews, 2012). Others sought to decrease the time required to grade assignments (Anglin, 2008). However, across those studies, several benefits were recognized: automation saves time (Anglin, 2008, Jones, 2001, Kovacic, 2012); automation allows grading to be consistent (Jones, 2001); automation enables fast feedback and more detailed feedback for the student (Kovacic, 2012, Matthews, 2012); and automation did not appear to negatively impact student satisfaction scores (Anglin, 2012). In general, it was noted that there is an upfront investment in migrating to automated grading methods, but that investment is paid off over several semesters (Jones, 2001). If class sizes are larger, the payback period is much shorter.

In our opinion, many of the solutions that have been developed are too technical and challenging to implement for an average user of Microsoft Excel. As an example, one solution requires the use several database tables in Microsoft Access in addition to worksheets in Microsoft Excel (Hill, 2003). The tool presented in this paper keeps everything in Microsoft Excel and does not require the transfer of data between software packages. Additionally, some of the solutions only automate a small number of grading steps or only reduce grading time by 50% (Anglin, 2008). Our solution is almost entirely automated, and when tested, it reduced grading time by more than 99%. Finally, with the solution documented in this paper, the upfront investment is minimal and can be paid back within the same semester in which the solution is implemented.

OVERVIEW OF THE MANUAL GRADING PROCESS

Before automation is discussed, we need a clear understanding of the steps required to manually grade assignments in Microsoft Excel. This will serve as a way to compare manual and automated methods. In a typical class, the process for manually grading an assignment that is submitted in Microsoft Excel might follow these steps:

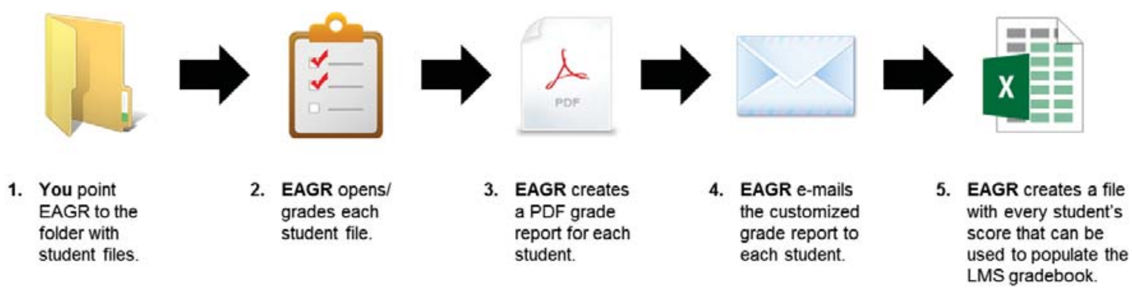
1. Download an individual student file.
2. Open the student file.
3. Evaluate each item in the rubric individually.
4. Compare student answers to the official answer key.
5. Capture feedback for the student about what was correct and incorrect.
6. Send that feedback to the student or upload it to the learning management system.
7. Add the student's grade to the learning management system.
8. Repeat steps 1-7 for each student file.

This is a time-consuming process with several mundane steps that are low-hanging fruit for automation. The solution discussed in this paper automates steps 2 through 6 above and eliminates step 8. Also, within the paper, in the “Results and Evidence” section, the time spent to manually grade assignments is compared to that of the automated method.

OVERVIEW OF THE AUTOMATED GRADING PROCESS

At Brigham Young University-Hawaii (BYUH), we programmed an automated grading tool in Microsoft Excel using the Visual Basic for Applications (VBA) programming language – called the *Excel Automated Grading Robot (EAGR)*. It automatically grades a batch of student assignments that are submitted in Microsoft Excel and e-mails students a customized grade report in a PDF format. EAGR also creates a spreadsheet for the professor that summarizes all student grades, which can be used to upload those grades to a learning management system. (At BYUH, we use Canvas.) Figure 1 below provides an overview of the solution. It has been used in the "Business Spreadsheets and Modeling Class" (BUSM 230), which is an online course required for all business students at Brigham Young University-Hawaii. Typically, there are about 175 students registered for the course in a given semester, and there are 8 assignments and 2 tests – all of which are completed by the student in Microsoft Excel. With EAGR, each of those assignments and tests can be graded in a matter of minutes. The data to demonstrate this is provided later in the paper. Even though EAGR was originally built for the BUSM 230 class, it is being tailored for other courses.

Figure 1: Overview of Automated Grading with EAGR



DETAIL BEHIND THE AUTOMATED SOLUTION

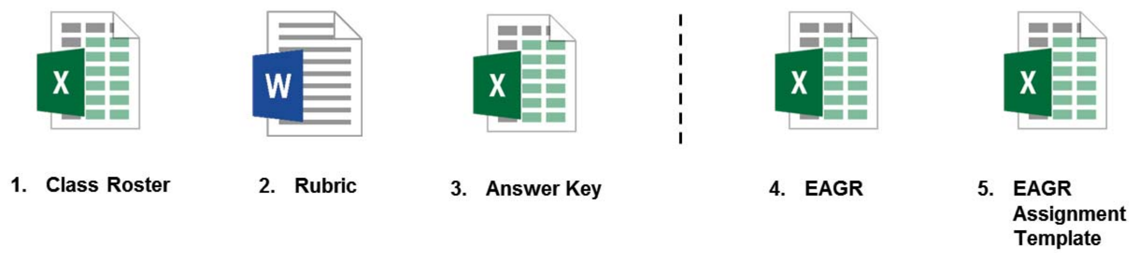
Technical Requirements

In order for EAGR to be used, the faculty member and/or teaching assistant must have a Windows PC, Microsoft Excel, Microsoft Outlook, and Adobe Acrobat. It may be possible to use other technical configurations, but that is outside of the scope of this paper.

Inputs

Additionally, several inputs are required before automated grading can be performed: (1) A class roster with student IDs, names, and e-mail addresses. This information should be gathered each semester from your learning management system. Then, (2) a rubric and (3) an answer key for the assignment. Finally, (4) the EAGR grading tool, which is a macro-enabled Microsoft Excel file and (5) the EAGR starter assignment template.

Figure 2: Inputs Required for Automated Grading with EAGR



Initial Setup

Prior to beginning the use of EAGR, the faculty member must complete the following steps: (1) Copy their rubric and paste it into the “GRADE” worksheet within the assignment template. (2) Create formulas to check whether each requirement from the rubric has been met. This can be done by comparing a student’s answer to the official answer key, which would be a hidden worksheet in the assignment template. (3) Enable restrictions within the learning management system, such as accepted file types, and the assignment template. The details of how to do so are in the “Implementation Tips” section later in this paper. These restrictions help to prevent grading errors. (4) Add student IDs, names, and e-mails to the “Student E-mails” worksheet.

Grading Process

Once the initial setup has been complete, you can begin using EAGR to grade assignments. To do so, the faculty member or teaching assistant performs the following steps: (1) Download all student files from the LMS; extract them, if they are compressed; and put them in an empty folder. (2) Open EAGR and click the “Grade Assignments” button. (3) When the first pop-up appears, select the folder where student files have been stored. On the second pop-up, enter the password used to protect the assignment file. At that point, EAGR will automatically grade the assignments and e-mail a customized grade report to each student. (4) After grading has completed, go to the folder where student files are stored and open the “Grade Spreadsheet” folder and select the “student_scores” file, which summarizes each student’s score. Use this data to update the LMS gradebook. (5) In the “student_scores” spreadsheet, address any errors that were identified in the “Errors” worksheet. (This usually occurs when students submit a different file type, such a Word document or PDF.)

Outputs

During the grading process, EAGR generates a few outputs: (1) a detailed grade report for each student in a PDF format (see Figure 2), (2) an e-mail to each student with their detailed grade report attached (see Figure 3), (3) a spreadsheet that summarizes all student grades and documents errors, if any occurred, and (4) folders that organize the student grade reports (PDFs) and the summary spreadsheet.

Figure 3: Student Grade Report (PDF)

GRADE REPORT

BUSM 230 - Module 2 - Intermediate Excel Skills

Problem Requirement	Score	
1. Correctly Sorted the "Sales" Column	0	out of 1.0 point
2. Correctly Calculated "Profit Margin"	0	out of 1.0 point
3. Added a Drop-Down Menu for "Order Method"	0	out of 1.0 point
4. Correctly Populated the "Region Manager" Column	0	out of 1.0 point
5. Formatted Appropriate Columns into Currency	0	out of 1.0 point
6. Correctly Created the "Orders Summary" Pivot	0	out of 1.5 points
7. Correctly Populated the "Returned?" Column	0	out of 1.5 points
Total	**0**	**out of 8.0 points**

Figure 4: Automated E-mail to Student

PDF [redacted].pdf
184 KB

Aloha, Excel Ninja!

Attached is your grade report for a recent assignment. If you have questions about your grade, you can (1) review course lessons to identify where a mistake may have been made, (2) attend the lab and speak with the teaching assistant, or (3) visit me during office hours.

Assignment: 1 - Basic Excel Skills
Score: 5 out of 6.0 points

RESULTS AND EVIDENCE

Time Savings

The primary objective of automating grading was to save time and university resources. To assess the solution's ability to meet this objective, time tests were performed. For three assignments in the BUSM 230 class, we evaluated the time required to grade assignments using normal manual methods and the automated method (EAGR). The grading activities for which time was accounted include steps 2-6 in the "Overview of the Manual Grading Process" earlier in this paper.

Table 1: Grading Time – Manual vs Automated Methods

	# of Files	Manual Grading Time (MM:SS)	EAGR Grading Time (MM:SS)
Assignment 1 *("loans")*	30	04:01.0 per file	00:01.2 per file
Assignment 2 *("stocks")*	30	02:37.9 per file	00:01.5 per file

Assignment 3 *("goal seek")*	30	01:59.1 per file	00:01.1 per file
Average	--	02:52.3 per file	00:01.2 per file

These results show that EAGR can save considerable time and money, even with small class sizes. Those savings grow exponentially as class sizes increase. As an example, in our BUSM 230 class that had 175 students, given the average times above, manually grading an assignment would take 8 hours, 22 minutes, and 32 seconds. Using the automated grader (EAGR), it would take 3 minutes and 30 seconds

At this point, an important note should be made about the upfront investment to move an assignment to the automated grading method. It does take time and planning. However, manual methods have initial development and training time, too. For example, as teaching assistants leave the university, new resources need to be trained on grading methods and techniques. In our tests, the upfront investment with EAGR repaid itself many times over.

Student Outcomes

As part of our original research question, we wanted to determine if we could save time and resources by automating the grading of Microsoft Excel assignments, while still maintaining (or possibly improving) student satisfaction. After confirming time was saved, we evaluated student satisfaction scores before and after the semester in which EAGR was implemented in the BUSM 230 class. That data is available in Table 2 below. As context, during the Fall 2017 semester, manual grading was performed by the faculty member and teaching assistant. After that semester, EAGR was used to grade assignments. We included data for Winter 2018 and Spring 2018 as we assumed that some normalization was needed. Additionally, when students provide feedback on a course, they answer several questions using a scale from 0 to 7. Some of those questions deal with whether or not the students felt they were graded fairly and whether or not the feedback they received was timely and useful.

Table 2: Student Feedback Scores (BUSM 230)

Student Feedback Score	Fall 2017	Winter 2018	Spring 2018	Change
Overall Course Score	6.35	6.49	6.42	+0.07
"Fair in Grading"	6.34	6.39	6.61	+0.27
"Timely Feedback"	6.20	6.25	6.61	+0.41
"Useful Feedback"	6.22	6.08	6.56	+0.34

Based on these results, it is clear; student satisfaction was not negatively impacted by the shift to automated grading. Between Fall 2017 and Spring 2018, before and after the implementation of EAGR, each metric shown above improved, which was a promising sign.

ADDITIONAL BENEFITS

Beyond considerable savings of time and resources, EAGR provided additional benefits, which are documented below. In many cases, this confirmed outcomes from previous studies that were performed and are referenced in the literature review.

(1) More consistent grading. When manual grading is performed, especially by a teaching assistant, mistakes are inevitable. Two assignments that are exactly the same could be

assessed different scores. When automation is used, the grading operations are performed the same way each time. Everyone is treated fairly.

(2) More detailed grading. Sometimes, as faculty, we limit the items (or the scope of items) on our grading rubrics because we may not have time to conduct the necessary checks. This can be harmful to students. They may not have as much actionable information about how to address weaknesses in their skills. Based on our experience, students prefer detailed grade reports. They want to know exactly what they did wrong. Additionally, with more detailed information in their grade report, they tend to ask fewer clarifying questions about their score, which can be another source of time savings and increased student satisfaction.

(3) Shorter feedback cycles. When grading is automated, feedback can be given to students much quicker. We found that students appreciate this. Feedback on student performance is more valuable when the assignment is still fresh on their mind. If students have to wait days or weeks, they may not remember as much about the specifics of the assignment. Additionally, Matthews (2012) noted that feedback given to students – its quality, quantity, and speed – is an important key to unlocking student learning.

(4) Easier cheating detection. When using VBA in Microsoft Excel, several features about the file itself can be collected, such as the user and last update time. This metadata allows more cases of possible cheating to be identified, and when cheating is punished, students share this information with their classmates, which reduces instances of cheating.

RESULTS IN OTHER COURSES

In addition to being implemented in BUSM 230, EAGR was also tested in ACCT 312 (“Managerial Accounting”) at BYU-Hawaii. During this process, 30 student files were graded – using both manual and automated methods. Using manual grading, it took an average of 5 minutes and 22 seconds to grade each file. Using EAGR, it took an average of 1.2 seconds per file. As in other time trials in this experiment, those times included steps 2-6 in the “Overview of the Manual Grading Process” earlier in this paper.

As the tool was explored in other courses, it was found that the time and difficulty of transitioning to EAGR depends on the following: (1) the quality of your existing assignments, (2) the presence and quality of your rubrics, and (3) the types of grading checks you are performing. If you are only assessing students on whether or not they reached the correct answer – as opposed to validating all of their calculations – the transition will be faster and smoother.

IMPLEMENTATION TIPS

Based on experiences within the BUSM 230 class and other courses, the following implementation recommendations have been developed. Faculty members should consider these, if they seek to deploy the EAGR tool.

(1) Restrict file types. If possible, you should adjust the assignment settings within the learning management system to only allow Microsoft Excel files (i.e. the .xls or .xlsx file extensions). If not, over the course of time, you are likely to have Microsoft Word documents and PDF files

submitted -- whether intentional or not. Additionally, if possible, only allow one file to be submitted.

(2) Add data validation to the assignment template. To ensure proper grading and to avoid errors, when you are looking for a numeric answer, add data validation. For example, if "20" is the correct answer, you do not want the student to enter "20 widgets."

(3) Protect the workbook and each worksheet in the assignment template. Grading may fail, if a student alters the template, which inevitably occurs when students are given freedom to do so. For example, if you have a formula that looks for an answer in a certain cell and a student created a new worksheet to enter their answers, even if those answers are correct, they will not be given any points.

(4) Explain to students why it is important to use the assignment template and why there are protections in that file. At times, students get frustrated with limitations in the file. However, if you explain that, by doing so, you are able to give them more timely, more detailed, and more consistent feedback, most students understand and comply.

CONCLUSION & NEXT STEPS

This automated grading solution has already produced substantive results, but it is clear that it might not suit every course. Additionally, more work need to be done to improve the effectiveness of the solution. As such, next steps for this work include (1) improving the grading code and adding more capabilities to detect cheating and (2) incorporating feedback from other universities and faculty members to make their transition to the solution easier.

REFERENCES

Anglin, L., Anglin, K., Schumann, P. L., & Kaliski, J. A. (2008). Improving the Efficiency and Effectiveness of Grading Through the Use of Computer-Assisted Grading Rubrics. *Decision Sciences Journal of Innovative Education, 6*(1), 51-73.

Burning Glass Technologies. (2015). Crunched by the Numbers: The Digital Skills Gap in the Workforce, Retrieved from https://www.burning-glass.com/wp-content/uploads/2015/06/Digital_Skills_Gap.pdf. Last accessed on May 17, 2019.

Bertheussen, B. A. (2014). Power to business professors: Automatic grading of problem-solving tasks in a spreadsheet. *Journal of Accounting education*, 32(1), 76-87.

Geigle, C., Zhai, C., & Ferguson, D. C. (2016, April). An exploration of automated grading of complex assignments. *In Proceedings of the Third (2016) ACM Conference on Learning@ Scale* (pp. 351-360). ACM.

Hill, T. G. (2003). MEAGER: Microsoft Excel automated grader. Journal of Computing Sciences in Colleges, 18(6), 151-164.

Jones, E. L. (2001). Grading student programs-a software testing approach. *Journal of Computing Sciences in Colleges, 16(*2), 185-192.

Koorsse, M., Calitz, A. P., & Zietsman, J. (2016, July). Criteria for Evaluating Automated Grading Systems to Assess Microsoft Office Skills. In Annual Conference of the Southern African Computer Lecturers' Association (pp. 51-58). Springer, Cham.

Kovačić, Z. J., & Green, J. S. (2012). Automatic grading of spreadsheet and database skills. *Journal of Information Technology Education, 11*(1), 53-70.

Matthews, K., Janicki, T., He, L., & Patterson, L. (2012). Implementation of an automated grading system with an adaptive learning component to affect student feedback and response time. Journal of Information Systems Education, 23(1), 71.

McNeil, C. R. (2015). Excel Modeling Assignments and Automatic Grading. *Journal of Financial Education*, 86-97.

DECISION SCIENCES INSTITUTE

Business Process Redesign and the Learning Organization: Exploring Single and Double Loop Learning

Mary Helen Fagan
The University of Texas at Tyler
Email: mfagan@uttyler.edu

ABSTRACT

Business process redesign requires the reinvention of business processes to achieve transformational improvements. Individuals must re-examine their assumptions about how to best achieve the desired results in order to envision and implement innovative organizational processes. As a result, process design must include organizational learning in order to be effective. This study illustrates how single and double loop learning can occur by analyzing an example of business process redesign undertaken in a large retail organization. This organization's experience can provide useful insights for other organizations that are trying to do the same and for change agents who are facilitating such efforts.

KEYWORDS: Organizational learning, Learning organization, Business process redesign, Double loop learning

INTRODUCTION

There are many different ways that an organization can adapt and change. One of the most important approaches to bringing about organizational change involves viewing the organization from a process perspective, as opposed to taking a functional, divisional or product view (Davenport and Beer, 1995; Hammer, 2015). Both quality management and business process redesign take a process-oriented approach to organizational change. However, business process redesign differs from business process improvement in that it seeks to engage in a more radical rethinking of business processes in order to achieve transformational, not incremental, improvements (Davenport, 1993; Huang, Lee, Chiu, & Yen, 2015). Business process redesign involves individuals in a change program where they re-examine their assumptions about how to best achieve the desired results in order to envision and implement radically innovative organizational processes.

This article posits that business process design, in order to be effective, must include organizational learning. The objective of this article is to illustrate how single and double loop learning can occur by analyzing an example of business process redesign at FoodCo (a pseudonym). Lessons that were learned at FoodCo as they implemented and modified their version of business process redesign can provide useful insights for other organizations that are trying to do the same and for change agents who are facilitating such efforts.

This paper is organized as follows. First, conceptual background information is provided on business process redesign and organizational learning. Then the study's methodology is described. Next, the FoodCo case is presented, and its experience with business process design is analyzed in terms of the learning experiences it underwent. Finally, conclusions are drawn.

CONCEPTUAL BACKGROUND

Systems research documents the use of technology for the automation of existing functions. The traditional approach to systems development emphasizes a logical, structured problem-solving process and, in general, automates routine operations and does not seek to introduce radical changes to an organization:

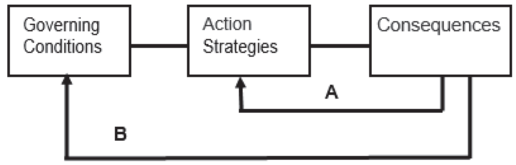
> "The organization is assumed to have a set of well-defined information processing problems which the MIS designer expertly analyzes. A rational decision-making process which examines alternative designs in terms of the goals of the system is followed to select the specific design. Implementation of this expert solution is then assumed to transform the organization immediately into a more effective state." (Bostrom and Heinen, 1977, p. 28).

The traditional systems development approach typically results in incremental improvements by speeding up processes through automation. Michael Hammer, one of the first proponents of business process reengineering, suggested that “heavy investments in information technology have delivered disappointing results - largely because companies tend to use technology to mechanize old ways of doing business. They leave the existing processes intact and use computers simply to speed them up" (Hammer, 1990, p. 104). In his work on business process redesign, Davenport (1993) contrasted the process improvement obtained by a traditional systems approach, which involves "performing work with slightly increased efficiency or effectiveness", with the objectives of process redesign, which involves "performing a work activity in a radically new way" (p. 10).

Some proponents of business process redesign suggest that organizations must radically change their business processes in order to survive in today's competitive environment (Vanwersch, Shahzad, Vanderfeesten, Vanhaecht, & Grefen, 2016). Setting aside the existing principles on which businesses processes are organized and endeavoring to “start from scratch” is one way organizations can work to achieve radical process redesign. When organizations seek to innovate in radically new ways, process redesign can be contrasted with traditional systems development approaches which have typically automated existing processes and resulted in incremental improvement. One author has called business process redesign "an organizational learning methodology that emphasizes a process view with its unique questioning: 'What do we do?' 'If we were doing it from scratch, how would we do it?'" (McGill and Slocum, 1994, p. 251). Organizational learning occurs, according to Argyris and Schon (1996), "when individuals within an organization experience a problematic situation and inquire into it on the organization's behalf. They experience a surprising mismatch between expected and actual results of action and respond to that mismatch through a process of thought and further action ... " (p. 16).

Argyris and Schon (1996) describe the two types of learning that can occur in organizations as single and double loop learning. Single loop learning is "instrumental learning that changes the strategies of action or assumptions underlying strategies in ways that leave the values of a theory of action unchanged" (Argyris and Schon, 1996, p. 20). In Figure 1, the loop labeled 'A' represents single loop learning. A double-loop learning process results in a "change in the values of the theory-in-use, as well as in its strategies and assumptions" (Argyris and Schon, 1996, p. 21). Together, the loops labeled 'A' and 'B' in Figure 1 represent the changes initiated by double loop learning.

Figure 1: Single and double loop learning



METHODS

This study takes a qualitative field study approach to studying a number of different business process management projects in one organization. The researcher was first involved as an observer and took a case study approach to studying the activities at FoodCo (Yin, 1994). Subsequently the researcher became a consultant to the business process redesign projects at FoodCo and eventually played the role of an active participant on a business process redesign team. The fact that the author participated in making decisions that would influence the success of the project and help to shape its direction meant any study coming from the author regarding these redesign activities could not consist of a typical case study where the author takes the role of a disinterested observer. Thus, an action research approach (Lau, 1999) the appropriate for the portion of the study involving process redesign. The data gathering process involved a combination of interviews, review of internal documentation and publicly available materials, and, when the researcher was involved as an active participant, materials gathered from internal project meetings and interactions with various vendors.

CASE DESCRIPTION

FoodCo is a vertically integrated food producer involved in the entire process of food production from growing the product to trucking the product to stores. FoodCo is a privately held firm with approximately 13,000 employees and $1.5 billion in revenue. The company has a culture that is highly focused on the key goals of efficiency and quality. FoodCo pursues their goal of increased efficiency through extensive measurement and management of details in all aspects of their operations. The slim profit margins in their segment of the industry mean that no wasted effort can be sustained in any area of the organization. This concern for costs is reflected in the fact that many individuals think of operational costs in terms of what percentage the costs constitute in terms of the cost per pound of the final product. For example, a manager might describe how his individual salary constitutes .xxx¾ of the cost per pound for the product. Their focus on measurement and efficiency is closely tied to the stress placed on quality. The commitment to quality starts with the chief executive and is mirrored throughout the organization. FoodCo has actively incorporated quality management into all of its operations through an intensive program of education, the articulation of the firm's Quality Improvement Process (QIP) and Error Cause Removal Process (ECR), and the cultivation of Corrective Action Teams (CAT's). The development of an organization that is focused on quality is a conscious, ongoing process at FoodCo.

The results of the twin focus on quality and efficiency can be readily seen in the changes that have taken place in the FoodCo Management Information Systems (MIS) department. The MIS department has shifted from a focus on application maintenance to a focus on determining the source of problems and redesigning processes to eliminate the root causes of ineffectiveness or inefficiency. A new relationship was developed with each functional area of the organization such that each group is treated as a customer of the MIS group, setting their own priorities and managing the development hours which they have been allocated and for which they are charged monthly. A set of critical success factors (CSFs) was defined for the MIS department which guide the activities of MIS associates and their organizational customers.

Business Process Redesign Activities

The MIS department took on the critical responsibility of introducing the concepts of business process redesign into the FoodCo organization. Characteristically, the introduction of process redesign at FoodCo was not the result of a top-down push from management. Instead, the initial introduction of process redesign concepts was in response to a dilemma that arose in the course of a typical MIS system development effort. This initial application of the process redesign approach was therefore appropriate to the problem, and, over time, the project was able to successfully demonstrate the utility of redesigning business processes in order to gain significant improvement at FoodCo.

The FoodCo Logistics organization was facing a changing business environment. To achieve their stated mission of providing quality transportation services that meet or exceed customer's requirements, they defined specific strategies that included reducing costs and providing a well-organized paperless environment. Key results areas and associated performance measures were defined for these strategies (e.g., improve tracking of trailers, reduce price of non-conformance, etc.). Analysis of the transportation area and dispatching, in particular, had identified a number of areas of potential improvement. Dispatch had three different tracking systems (board, paper, and computer) that made it difficult for areas within transportation to communicate with each other. Furthermore, the steep learning curve for dispatchers created a serious problem when there was staff turnover. A traditional systems analysis and design project had identified a number of ways in which technology could be used to improve and speed up the existing work flows. Data flow diagrams that depicted the existing system had been developed and 50% of the user requirements had been defined for a new system.

However, a number of serious questions arose in the process of completing the process of this traditional systems design effort. One key concern in the dispatch area was the potential for the current work load to double, creating an estimated 10,000 incoming phone calls per week. Even with a new automated system for dispatching, it was unclear how the proposed system design would deal with this increase in call volume without a corresponding increase in personnel. Furthermore, the proposed system design had a number of features designed to improve the ability to track trucks and trailers. However, in order to gather the needed data for the system, even more phone calls would be generated and a heavy data entry burden would be imposed on dispatchers. After the team reviewed these issues in light of their proposed design, the team determined that automation alone would not solve these problems and that a business process redesign effort was required.

Business process redesign would entail radically rethinking the dispatching and transportation business processes in order to achieve dramatic performance improvements. As a result of the redesign project, unnecessary activities would be eliminated and the remaining value-added activities would be simplified prior to the application of enabling information technologies. By communicating the vision for where the Logistics organization wanted to go in the future, the Logistics area's five-year strategic plan provided a context for the redesign effort. A consulting firm was engaged to help provide the needed expertise. The following steps were then undertaken.

Documenting the Current Process

As a result of prior system analysis, the dispatching employees were already expecting to get personal computers with planning software. The dispatching staff were very disappointed when this purchase was put on hold until the new business process redesign project was completed in an estimated four months. The participants' concern about this change in approach, combined with some frustration stemming from prior automation efforts, resulted in strong emotions as the project team began to document the current process. As one participant related, "it took almost an hour to decide where the load for a truck came from and to get the first symbol on the page". However, the team, which consisted of MIS staff and dispatching employees, started out with the understanding that the dispatching staff had to agree with each item or it would not be put down on the current process model. The full commitment of the Logistics organization and the development of a close-knit team enabled them to complete the task of understanding and documenting the dispatching process in detail. When the current process model was complete, the team identified non-value added activities such as inspection and rework, and began to identify and prioritize opportunities for improvement.

Conducting an Envisioning Workshop

With the organizational requirements from the Logistics strategic plan, the current process model analysis, and the knowledge of state-of-the-art techniques provided by external consultants, the team proceeded with the next step: to envision a world class transportation system. Three separate cross-functional teams met to develop ideas for the new process. The experience demonstrated how difficult it can be to think "out of the box", especially when one is concerned with obtaining management approval to proceed. As one participant related, "going into the visioning process, I thought radio frequency tags were the best solution we could achieve. Previously, the idea of using satellites had come up, but I thought there was no way it could be cost justified and approved. So, when we left the meeting with a decision to use satellites, I was very upset". The experience also pointed up the value of bringing in "outsiders" who provide a unique perspective on the problem.

Building a New Process Model

Although the costs seemed high, satellite tracking seemed to be the only approach that could address all the issues that had made process redesign necessary in the first place. Without the ability to address issues such as the increasing number of phone calls that would require the hiring of additional dispatchers, the project team would not be able to identify sufficient benefits to justify the system development effort. With the direction provided by the envisioning meeting, the team rapidly developed a future process model. Overall, the new process model was designed to: a) streamline the process, b) build in flexibility to adapt to a changing environment, c) incorporate an equipment tracking system, d) give the process the ability to grow and change as needed, and e) enable FoodCo to meet the needs of the marketplace.

Buy-in was generated as the system analysts discussed with each dispatcher how the future process design specifically addressed the items on their original issue list. The team identified nine projects and did a cost/benefit analysis for each. In a meeting with top management, their plan was approved. The team was directed to carry out an equipment tracking pilot project using satellite transponders to handle all tracking and check calls. The pilot was designed so that if there were any serious problems, they would surface in the pilot.

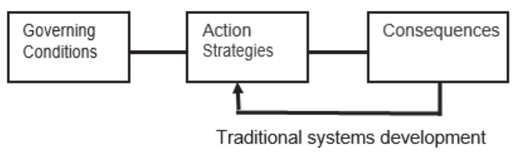
Institutionalizing Business Process Redesign

When the pilot equipment tracking project was completed, many of the desired benefits were realized. For example, dispatcher phone calls were reduced by approximately 30% and the ability to track and better utilize trailers resulted in a 7% reduction in trailer purchases. Dispatchers now have a better ability to make on-the-spot decisions with knowledge of the different trade-offs so that customer and organizational performance goals can be achieved. Changes to the transportation organization's work culture, organization and policy and procedures have been well-received. Approval to proceed with the full implementation of the new business processes in dispatching was given. Additional system development projects identified through the process redesign effort were initiated in the transportation organization.

DISCUSSION AND ANALYSIS

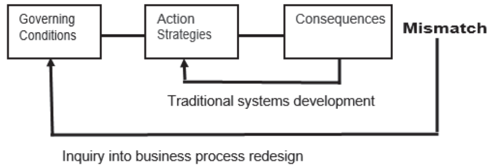
FoodCo used a traditional systems implementation process to identify and resolve problems in functional areas of the business. As shown in Figure 2, this activity, which is focused on a business problem, could be termed single loop learning since "the process enables the organization to carry on its present policies or achieve its objectives" (Argyris, 1977, p. 116).

Figure 2: Traditional systems development



However, in the case the dispatching organization at FoodCo, the MIS group found that the traditional systems analysis approach that they had used would not solve the problem situation that they encountered. Their analysis found that even if they installed an automated dispatching system, the anticipated increase in calls to dispatchers would not make the proposed systems solution approach cost effective. As shown in Figure 3, the fact that the normal systems analysis and design process would not result in an effective dispatching solution surprised the FoodCo employees, and the mismatch triggered a search for a new workable strategy. At this point, the organization was undertaking an inquiry, the type that occurs when a surprise/mismatch blocks the flow of spontaneous activity and gives rise to thought and further action aimed at re-establishing that flow" (Argyris and Schon, 1996, p. 11).

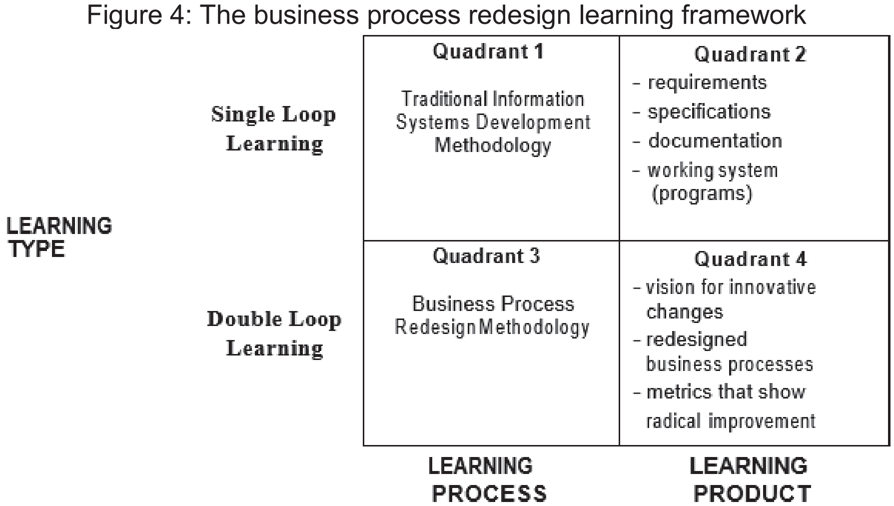
Figure 3: Inquiry into business process redesign



Until the team began to wrestle with the question - which had been unthinkable before - of whether their traditional systems implementation approach could resolve the dispatching problem, it was engaged in single loop learning. If FoodCo had uncovered a strategy for action that would have enabled it to resolve dispatching's systems dilemma within the framework of their established systems development methodology, that would have constituted single loop learning. However, since the governing values of the FoodCo organization favored incremental change (as evinced by their TQM program), the dispatching design team was initially at a loss as to how to proceed with a solution that would fit within the existing values and norms. When the inquiry led the dispatching design team to look to an external consultant to train them on business process redesign in order to achieve radical change, they were striving to change the governing values of the theory-in-use regarding systems change at FoodCo. This effort resulted in a double-loop learning process since the result was eventually a "change in the values of the theory-in-use, as well as in its strategies and assumptions" (Argyris and Schon, 1996, p. 21). When management agreed to the use of business processes redesign, they were espousing new values - an endorsement of intentional, radical, and risky organizational change.

A framework for analysis, derived from the work of Argyris and Schon (1996), can help further understanding of the organizational learning that occurred at FoodCo as they engaged in business process redesign. As discussed previously, Argyris and Schon define types of learning to be either single loop learning ("instrumental learning within a constant frame of values") or double loop learning ("learning to change the values that define 'improvement'") (1996, p. 4). Argyris and Schon (1996) also defined a generic schema of organizational learning that includes these items: "some information content, a learning product, a learning process which consists in acquiring, processing and storing information, and a learner" (p. 3). Combining the two types of learning (single and double loop) with two aspects of the generic schema of organizational learning (learning process and product) produces the framework for analysis shown in Figure 4.

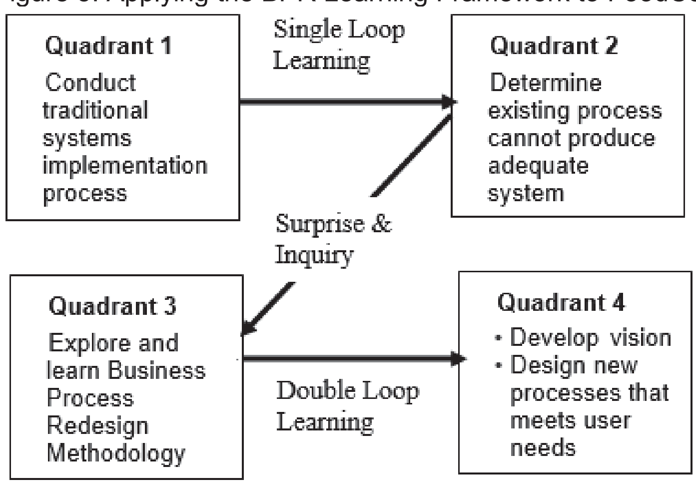
Figure 4: The business process redesign learning framework



From an overall perspective, the traditional systems development methodology can be seen as a learning process (quadrant 1) that produces a variety of information content/learning products (quadrant 2). At FoodCo the traditional systems development methodology was a well-understood process routinely used to automate existing operations. Within this framework, a business process redesign methodology/approach represents a learning process that should bring about double loop learning, a change in the values that define improvement from the incremental change associated with the traditional systems methodology to a new emphasis on innovative, radical improvements. Thus, quadrant 3 shows business process redesign methodology as a double loop learning process that produces a number of learning products (quadrant 4).

The following diagram (Figure 5) shows the process that took place at FoodCo in terms of the framework introduced in Figure 4. FoodCo was conducting a traditional systems implementation process (quadrant I) that could be termed single loop learning. When they determined that their existing methodology could not design a solution that would meet the users' requirements they could not produce the learning products associated with quadrant 2 (e.g., requirement documents leading to a working system). The surprise they encountered led to an inquiry process whereby they explored a new systems development approach, business process redesign.

Figure 5: Applying the BPR Learning Framework to FoodCo



The adoption of business process redesign constituted double loop learning since they were learning to change the values that defined improvement to include the risky, radical change called for by this new approach (quadrant 3). Business process redesign became the espoused theory-in-use for FoodCo for problems that appeared to require novel, innovative solutions outside the scope of the traditional systems development approach. These new values (e.g., of the importance of quantum, radical change) required a new way of carrying out the system analysis and design process. For example, employees who were exhorted to "think out of the box" in the dispatching envisioning workshop were engaging in a new strategy for problem solving. They were engaged in double loop learning related to instrumental business problem solving (quadrant 4). Indeed, without the urging of the consultant, some of the envisioning workshop participants would likely have selected a solution that they thought management would endorse (radio frequency networking) as opposed to the solution that was selected (satellite connections). When management accepted the risky and controversial satellite solution recommended by the dispatching team, it reinforced the espoused theory-in-use that it was OK to go for radical (and more risky) solutions. Thus, the teams search for dispatching solutions led them to rethink and adjust the governing conditions that affected system solutions at FoodCo. Business process redesign will require the ability to engage in double loop learning on an ongoing basis if teams are going to be able to challenge existing solutions and design innovative new processes, as opposed to falling back on tried and true traditional approaches to system development that simply automate existing processes.

FoodCo was able to engage in double loop learning and rethink the governing conditions that had guided systems/process change activities. Research in other organizations has indicated that activating double loop learning can lead to performance improvement (Jaaron & Backhouse, 2017). For FoodCo, which placed high value on efficiency and continuous, incremental change, it meant they had to be willing to experiment and learn a new radical and risky way of doing things. However, in order to successfully diffuse the new approach throughout the organization, teams engaged in actual business process redesign projects had to be able to continue to engage in double loop learning related to instrumental, functional business process issues. Argyris notes an organization is rarely "able to use double loop learning for its instrumental and policy issues if it cannot do so for the games and norms" (Argyris, 1977, p. 117). FoodCo was able to take the first steps toward this capability when it used double loop learning to adjust its norms (governing conditions) and then proceeded to engage in double loop learning in the dispatching business process redesign project itself. However, it remains to be seen if the development of FoodCo's own business process redesign methodology and the use of it in other projects will be sufficient to maintain this change and accommodate instrumental double loop learning on projects on an ongoing basis. Business process redesign requires identification of problems and a willingness to uncover unpleasant truths about how processes operate in order to envision a novel solution. If individual theories-in-use are geared toward avoiding embarrassment and blame, then it is unlikely that business process redesign will thrive even if it becomes an espoused organizational theory of action.

CONCLUSION

In the future, FoodCo must adapt to rapidly changing market conditions. Business processes will be continually examined in order to identify areas for improvement, and business process management is seen as key to FoodCo's future success. In some cases, process analysis will result in selected improvements in the existing work flows and incremental improvement in their efficiency and effectiveness (e.g., single loop learning). In other cases, business process redesign will be called for in order to implement a novel solution involving transformational learning on the part of process participants (e.g., double loop learning). Now, both adaptive learning from total quality management and transformative learning from business process redesign are part of the FoodCo culture, and the appropriate approach can be employed as needed. By incorporating process redesign tools and techniques along with the total quality management philosophy that pervaded the company, FoodCo positioned itself to undergo the continual learning that will be required for its ongoing success.

REFERENCES

Argyris, C. (1977). Double Loop Learning in Organizations. *Harvard Business Review, 55(5)*, 115-125.

Argyris, C. & Schon, D. A. (1996). *Organizational Learning II: Theory, Method, and Practice.* New York, NY: Addison Wesley Publishing Co.

Bostrom, R.P. & Heinen, J.S. (1977). MIS Problems and Failures: A Socio-Technical Perspective. *MIS Quarterly*, 1(3), 17-32.

Davenport, T. H. (1993). *Process Innovation: Reengineering Work Through Information Technology*. Boston, MA: Harvard Business School Press.

Davenport, T. H. & Beer, M. C. (1995). Managing Information about Processes. *Journal of Management Information Systems*, 12(1), 57-90.

Hammer, M. (2015). What is business process management?. In *Handbook on business process management,* Berlin: Springer, pp. 3-16).

Huang, S. Y., Lee, C. H., Chiu, A. A., & Yen, D. C. (2015). How business process reengineering affects information technology investment and employee performance under different performance measurement. *Information Systems Frontiers*, 17(5), 1133-1144.

Jaaron, A. A., & Backhouse, C. J. (2017). Operationalising “double-loop” learning in service organisations: a systems approach for creating knowledge. *Systemic Practice and Action Research*, 30(4), 317-337.

Hammer, M. (1990). Reengineer Work: Don't Automate, Obliterate. *Harvard Business Review*, 90(4), 104-112.

Lau, F. (1999). Toward a Framework for Action Research in Information Systems Studies. *Information Technology & People*, 12(2), 48-175.

McGill, M. E. & Slocum, J. W. (1994). *The Smarter Organization: How to Build a Business that Learns and Adapts to Marketplace Needs.* New York, NY: John Wiley & Sons.

Yin, R. K. (1994). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

Vanwersch, R. J., Shahzad, K., Vanderfeesten, I., Vanhaecht, K., Grefen, P., Pintelon, L., ... & Reijers, H. A. (2016). A critical evaluation and framework of business process improvement methods. *Business & Information Systems Engineering*, 58(1), 43-53.

DECISION SCIENCES INSTITUTE

Literature Review on Blockchain Research and Applications in Business

Prinsa Roka
Eastern Illinois University
Email: proka@eiu.edu

Chao Wen
Eastern Illinois University
Email: cwen@eiu.edu

Nan (Tina) Wang
Eastern Illinois University
Email: nwang@eiu.edu

ABSTRACT

Blockchain, well-known as the Distributed Ledger Technology (DLT) behind cryptocurrencies such as Bitcoin, has attracted plenty of attention from both practitioners and academics and has begun to revolutionize the way businesses handle their day-to-day operations. In this study, we applied the systematic mapping method to review existing articles related to blockchain applications and research in different business areas (accounting, finance and banking, information systems, marketing, and supply chain). Following the recommended steps of the systematic mapping method, we extracted fifty primary papers from several scientific databases. Research gaps and future research directions are discussed.

KEYWORDS: Blockchain, Systematic Mapping Methods, Business Applications, Literature Review, Distributed Ledger Technology

INTRODUCTION

Blockchain is revolutionizing the way businesses handle their day-to-day operations. Most businesses rely on centralized transaction systems where the information is controlled by third-party organizations, instead of the specific entities involved in the transaction. Blockchain provides a decentralized platform where transactions can take place without the control and involvement of third parties. It is an open, distributed database where different participating computers (or nodes) validate the transaction which then gets added to a continuously growing list of the ledger (Iansiti & Karim, 2017). The validated information is available to every node, which adds to the transparency of the transactions conducted.

A main application of blockchain is cryptocurrency which eliminates the problem of double spending the same currency and increases the transaction security with the use of hash numbers, timestamp, and append-only decentralized system (Miraz & Ali, 2018). However, blockchain can be applied to different business sectors due to its security, reliability, transparency, traceability, inherent data provenance and immutability. Blockchain also creates new possibilities and applications in business.

Like other emerging technology in its nascent stage, blockchain has limitations that need to be addressed before having a major impact on business. For example, since nodes in the blockchain network can be anywhere in the world, complex jurisdictional issues may arise. In addition, every transaction requires peer-to-peer verification, which can become time-consuming with more data added. As a result, blockchain might be unsuitable for transactions where speed is of the essence.

Blockchain has received and is likely to continue receiving attention from researchers and practitioners. It is important to synthesize existing progress related to blockchain applications and research in business, in order to identify areas needed future attention. Specifically, this paper seeks to understand the following research questions: *What is the current status of blockchain research in business?*

Understanding the above research question helps researchers and practitioners develop an overall understanding of existing blockchain applications and research done by different disciplines and points out gaps that require more future attention. To address the research question, a systematic mapping study is conducted in which 50 articles about blockchain are extracted from scientific databases. Several research gaps are identified, and future research directions are pointed out.

The remaining of the paper is organized as follows. The research methodology used for literature review is introduced first. After that, figures showing the status of existing blockchain research in business are presented. The last section of the paper discusses the research gaps identified from our review and point out several future research directions.

RESEARCH METHODOLOGY

The systematic mapping method

The systematic mapping study as described by Petersen et al. (2008) was adopted to identify, review, and summarize relevant research papers. The systematic mapping method is appropriate for this paper because our goal is to identify and map existing research and application of blockchain in different business areas (i.e., accounting, finance and banking, information systems, marketing and supply chain) and to identify possible research gaps. The systematic mapping method as well as our application in this paper is summarized in Table 1.

Table 1 Summary of systematic mapping steps used in this paper

Steps	**The application in this study**
Defining research questions	Defined the research protocol and narrowed research questions down to the two research questions mentioned earlier.
Searching for relevant papers	Used search strings related to blockchain and business areas to search scientific databases including Springer Link, EBSCO, JSTOR, and ScienceDirect. The papers chosen mainly focused on peer-reviewed, high quality research that was published in books, journals, conferences, workshop, and symposium. 97 papers were initially identified.
Screening of papers	• The first stage of inclusion and exclusion was based on the relevance of the titles of the selected papers to our research questions; the following types of papers were excluded: papers in a language other than English, duplicate papers, and posters. After first stage screening, 60 papers were kept.

	• In the second stage of screening, abstracts were screened to see if they focused on blockchain in business areas to decide whether the papers could made to the next step. 10 papers were excluded because they focused mostly on the technical workings behind blockchain or did not provide any actual or clear research findings or evidence. This resulted in the selection of 50 paper, which were included in this study as primary papers.
Keywording	Read all 50 papers and identified the keywords and concepts related to different business areas to develop a higher-level understanding of the papers. After that, clusters and categories were formed for each of the business area based on the year of publication, publication type, source, and field.
Data extraction and mapping process	A data extraction form was created (Table 2) to collect the information needed to address our research questions. The extracted data were in excel to make it easy to organize and analyze the retrieved information.

Table 2. Data Extraction Items

Title	Title of the paper
Authors' name	Name of the author(s)
Publication information	Name of the publication place
Publication type	Conference, book, journal, working papers, thesis/dissertation, industry article/ magazines
Publication source	Industry or Academia
Abstract	Abstract of the selected paper
Study aim	Aim of the selected paper
Field	Business area(s) focused in the paper
Study findings	Major findings of the selected paper

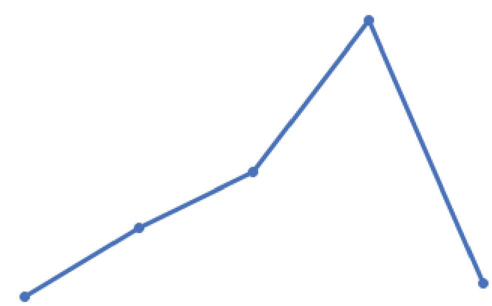
RESEARCH FINDINGS AND DISCUSSION

RQ2: What is the current status of existing blockchain research in business?

In this section, we will first provide an overview of the selected papers, with a focus on publication date, publication source, publication types and the related business fields. We will elaborate on the current research status, research gaps and future research directions in the discussion section.

Figure 1 shows the distribution of selected papers by publication year. All selected papers were published after the year 2015. This reflects that Blockchain is a fairly new concept for research in Business. However, the papers that focused mainly on Bitcoin and technical working of blockchain date (which were excluded from this study) was published as early as 2012. When we pay closer attention to the publication year, we can see that out of all the selected papers, only 3 papers (6%) were published in 2015, 8 papers (16%) in 2016, 12 papers (24%) in 2017, 23 papers (46%) in 2018, and so far 4 papers (8%) in 2019 [The data for 2019 is only for the first 4 months and it is likely that more papers will be published before the end of the year]. This shows an upward trend with an increasing number of papers each year, which suggests that there is a growing interest in blockchain technology in different business areas.

Figure 1. Publication Year of Selected Papers



	2015	2016	2017	2018	2019
Total	3	8	12	23	4

Figure 2 provides a summary of the publication sources (academia or industry). Academic sources are the ones written by experts in the field for the audience who are familiar with the topic in order to make the research available to the scholarly world. Industry sources are the ones written by free-lance or scholars for the general population with little to no familiarity with the topic in order to provide general/ informative information about the topic. Our result showed that 40 papers (80%) were published by academic sources, 10 papers (20%) were published by industry sources. Please note that the results might be due to the tendency for scientific databases to not include industry papers.

Figure 2. Publication Source of Selected Papers

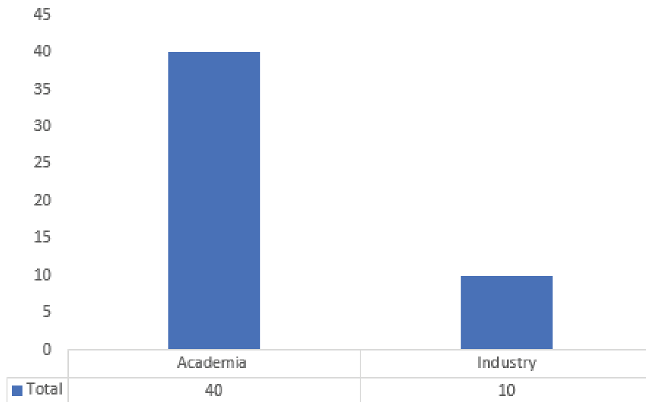
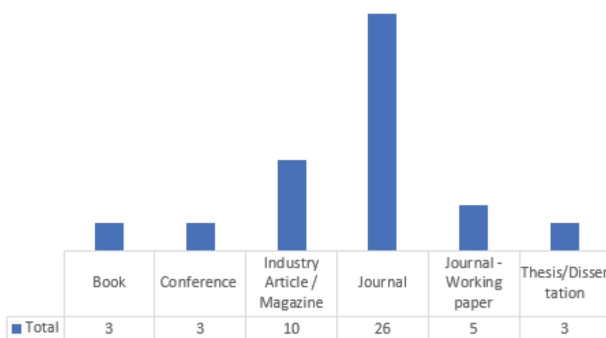


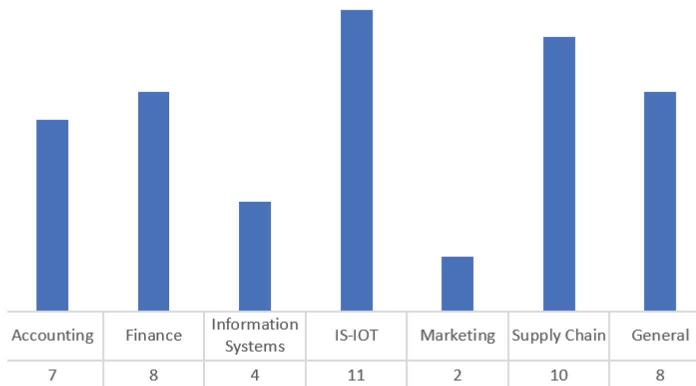
Figure 3 shows the different publication types (i.e., the channel where the paper is published, as listed in Table 2). Most of the papers were published in journals (52%) or industry article/magazine (20%). The other types include working paper (10%), book (6%), thesis/dissertation (6%), and conference proceedings (6%).

Figure 3. Publications Types of Selected Papers



The distribution of business fields related to selected papers is shown in Figure 4. The largest number of papers were published with the IS filed 15 papers (30%) with 11 papers focused on Internet of Thing (IoT), and supply chain with 10 papers (20%). The accounting field had 7 papers (14%), the finance field had 8 papers (16%), marketing had 2 papers (4%), and there were 8 papers (16%) that talked about the use of blockchain in general business setting.

Figure 4. Publications by Different Business Fields



DISCUSSION

Using the systematic mapping method, this paper seeks to understand the applications and existing research status related to blockchain in business areas (i.e., accounting, finance and banking, information systems, marketing and supply chain). Our study shows, first, that there are a number of journal articles on blockchain; in addition, there are quite a few industry articles and thesis/dissertation articles that focus on the application of blockchain in business areas. Most of these publications are made by academic (rather than industry) sources.

As to the distribution of papers across displaces in business, many research papers focuses on the application of blockchain in information systems (mainly IoT), followed by the application in the supply chain area. There are fewer research papers focusing on blockchain applications in accounting and finance and even fewer papers focusing on the marketing field.

As to the topics examined, supply chain research mainly focused on topics regarding the use of blockchain for supply chain sustainability, risk management, resilience, distributed data-driven application systems, and excipient supply chain. Information systems research, on the other hand, focused on topics of collaborative blockchain, operation management interface, proof of work, machine economy, and IoT dataset. Finance research mainly focused on topics such as financial blockchain technology, fintech, virtual currency, sustainable development, and banking industry transformation. Marketing research focused on the use of blockchain for programmatic marketing, data veracity, data-driven marketing, and consumer satisfaction. Finally, accounting research focused on blockchain accounting, triple-entry accounting, request network, real-time accounting, automated audit future accounting and assurance, and credit risk modeling.

RESEARCH GAPS AND FUTURE RESEARCH DIRECTIONS

One of the research gaps that we identify when conducting the review is the abundance of research papers on bitcoin and the relative lack of research focusing on the business areas, especially in the fields of marketing and accounting. Since both fields are likely to be heavily impacted by the implementation of blockchain, further research needs to be conducted. In addition, due to the complexity and fundamental changes that may be introduced by blockchain to different business fields, future research directions for blockchain in different business areas may be varied and it would be interesting to see where each business area is heading. Since the 2018 cryptocurrency crash, bitcoin's popularity has significantly declined. Academic and industry researchers are more interested in understanding the application of the fundamental technology behind blockchain from both business and technical perspectives. We believe that as more businesses start using blockchain in their day-to-day functioning, it will generate a significant amount of new research regarding the adoption of blockchain technology in different contexts. As the blockchain scope and user base increases, more attention needs to be paid to overcoming implementation limitations and challenges (e.g., the scalability issues) discussed earlier. The security and privacy issues would be especially important, as more and advanced attacks are likely to target the blockchain network with its increasing popularity.

The second research gap we identify is that most research focuses on blockchain application by companies (especially large companies). If the adopters were small-medium sized companies or teams or even individuals, would our existing understandings about blockchain hold or not? Accumulated knowledge in information systems indicate that factors affecting individual versus team versus organizational attitudes and behaviors (e.g., adoption) towards an innovation are likely to vary (e.g., Hwang 2005; Park, Lee, & Yi, 2011; Venkatesh et al 2013). Hence, future research should examine attitudes and behaviors towards blockchain at different levels of analysis.

The third research gap identified is that most of the current studies focused on the technical aspect of blockchain. The behavioral and regulatory aspects related to blockchain need more attention. For the behavioral aspects, researchers could study, for example, what factors (e.g., individual or organization or industry or cultural characteristics, blockchain characteristics, network externality) (e.g., Venkatesh et al 2003; Venkatesh and Zhang, 2010; Veiga et al 2001) influence the (initial and continued) (e.g., Bhattacherjee & Premkumar, 2004) adoption and implementation of blockchain? How the different concerns mentioned earlier (e.g., concerns for privacy) affect adopters' attitudes and behaviors towards blockchain? Does the blockchain technology exhibit the same bandwagon pattern as we have seen with other popular innovations (e.g., Abrahamson & Rosenkopf, 1993)?

In addition to the behavioral aspect, the regulatory aspect (i.e., industry or national regulations) also deserves more attention. For example, Chinese government imposes restrictions on the amount of money that individuals and companies could transfer to overseas. What kind of challenges would blockchain bring to such regulation? In addition, when disputes arise regarding transactions that happened on blockchain platform, what kind of authority (if any) could help to settle the dispute?

Finally, we noticed that the majority of blockchain applications in different business areas are still in the experimental or pilot phase. There is limited data published related to the success or failure of blockchain in different business fields. Future research should continue looking for such data. Also, researchers may examine market responses (e.g., stock market price change) to companies announcing the adoption and implementation of blockchain (e.g., Wang, 2010).

CONCLUSIONS

As an emerging and in-fashion technology, blockchain, a decentralized distributed ledger, where transactions are validated by the participating nodes and are visible to everyone on the network, has received quite some attention in academia and industry. Using the systematic mapping method and extracting 50 primary papers from various scientific databases, this paper conducts a systematic review to understand the applications and current research status related to blockchain in different business areas. Research gaps are identified, and future research directions are pointed out.

References

Abrahamson, Eric, & Rosenkopf, Lori. 1993. “Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion,” *Academy of Management Review*, 18(3), pp. 487-517.

Alphand O., & Amoretti M., & Claeys T., Dall’Asta S., & Duda A., Ferrari G., & Rousseau F., & Tourancheau B., & Veltri L., Zanichelli F. 2018. “IoTChain: A blockchain security architecture for the Internet of Things,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*. Barcelona, Spain. pp. 1–6.

Anascavage, R., & Davis, N. 2018. “Blockchain Technology: A Literature Review.” SSRN. doi: org/10.2139/ssrn.3173406.

Andersen, N. 2016. *Blockchain Technology A game-changer in accounting?* Deloitte.
Angelis, J., & Silva, E. R. 2018. “Blockchain adoption: A value driver perspective.” *Business Horizons*. doi:10.1016/j.bushor.2018.12.001

Apte, S., & Petrovsky, N. 2016. “Will blockchain technology revolutionize excipient supply chain management?” *Journal of Excipients and Food Chemicals*, pp. 75-78.

Atlam, H., & Alenezi, A., & Alassafi, M., & Wills, G. 2018. “Blockchain with Internet of Things: Benefits, Challenges and Future Directions.” *International Journal of Intelligent Systems and Applications*. 10(6).pp 40-48. doi: 10.5815/ijisa.2018.06.05.

Banerjee, M., & Lee, J., & Raymond C., K. 2017. “A blockchain future to Internet of Things security: A position paper.” *Digital Communications and Networks.* 4(3). pp 149-160 doi: 10.1016/j.dcan.2017.10.006.

Bansal, S., & Batra, R., & Jain, N. 2018. Blockchain and the future of accountancy. Institute of Chartered Accountants in England and Wales (ICAEW)., Retrieved from https://www.icaew.com/technical/technology/blockchain/blockchain-articles/blockchain-and-the-accounting-perspective.

Beck, R., Avital, M., Rossi, M, & Thatcher, J. 2017. “Blockchain Technology in Business and Information Systems Research.” *Business & Information Systems Engineering.* 59 (6). pp 381-384 Doi: 10.1007/s12599-017-0505-1.

Bhattacherjee, Anol, & Premkumar, G. 2004. “Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test.” *MIS quarterly,* pp 229-254.

Brandon, D. 2016. “The blockchain: the future of business information systems?” *International Journal of the Academic Business World,* 10(2), pp 33-40.
Casey, M., & Vigna, P. 2018. *The truth machine: The blockchain and the future of everything*. New York: St. Martins Press.

Dai, J. & Vasarhelyi, M. 2017. “Toward Blockchain-Based Accounting and Assurance.,” *Journal of Information Systems.* 31(3). pp 5-21 doi: 10.2308/isys-51804.
Dills, T. (2018). *The coming of blockchain.* Retrieved from https://www.overdriveonline.com/the-coming-of-blockchain/

Dorri ,A., & Kanhere, S, S., & Jurdak, R., & Gauravaram, P. 2017. “Blockchain for IoT security and privacy: The case study of a smart home,” in the *Proceedings of IEEE Int. Conf. Pervasive Computer Community.* pp 618–623.

Feng, T. 2017. “A supply chain traceability system for food safety based on HACCP, blockchain & Internet of things.” In the *Proceedings of International Conference on Service Systems and Service Management*, Dalian, pp. 1-6. doi: 10.1109/ICSSSM.2017.7996119.

Fernández-Caramés, T., & Fraga-Lamas, P. 2018. “A Review on the Use of Blockchain for the Internet of Things.” IEEE Access, 6, pp 32979-33001. doi: 10.1109/ACCESS.2018.2842685.

Ghose, A. 2018. “What Blockchain Could Mean for Marketing,” *Harvard Business Review*, 1-6.

Guo, Y., & Liang, C. 2016. “Blockchain application and outlook in the banking industry,” *Financial Innovation,* 2(24). doi:10.1186/s40854-016-0034-9.

Hambiralovic, M., & Karlsson, R. 2018. *Blockchain Accounting in a Triple-Entry System*. Lund University, 1-52.

Harvey, C. R., & Moorman, C., & Toledo, M. 2018. “How Blockchain Can Help Marketers Build Better Relationships with Their Customers.” Harvard Business Review, 1-8.

Hwang, Y. 2005. “Investigating enterprise systems adoption: uncertainty avoidance, intrinsic motivation, and the technology acceptance model,” *European Journal of Information Systems*, 14(2), pp 150-161.

Iansiti, M., & Karim R. 2017. “The Truth about Blockchain,” *Harvard Business Review*, 95, no. 1, pp 118–127

IBM. 2018. “Maersk and IBM to form joint venture applying blockchain to improve global trade and digitize supply chains.” Available at https://www03.ibm.com/press/us/en/pressrelease/53602.wss

Jaikaran, C. 2018. *Beyond Bitcoin: Emerging Applications for Blockchain Technology.* Congressional Research Service, 1-9.

Kshetri, N. 2017. “Can Blockchain Strengthen the Internet of Things?” *IT Professional.* 19. pp 68-72. doi: 10.1109/MITP.2017.3051335.

Kshetri, N. (2018). “Blockchain’s roles in meeting key supply chain management objectives,” *International Journal of Information Management*. 39. 80-89. doi: 10.1016/j.ijinfomgt.2017.12.005.

Kumar, S., & Mookerjee, V., & Shubham, A. 2018. “Research in Operations Management and Information Systems Interface,” *Production and Operations Management*, 27(11). doi:10.1111/poms.12961.

Leon, C. D., & Antonius, Q. S., & Ananth, J., & Haney, A. Michael & Sheldon, F.T. 2017. “Blockchain: properties and misconceptions,” *Asia Pacific Journal of Innovation and Entrepreneurship*. 11. pp 286-300. doi: 10.1108/APJIE-12-2017-034.

Lin, I, C., & Liao, T, C. 2017. “A survey of blockchain security issues and challenges,” *International Journal of Network Security*. 19. pp 653-659. doi: 10.6633/IJNS.201709.19(5).01.

Madhwal, Y., & Panfilov, P. 2017. “Blockchain And Supply Chain Management: Aircrafts’ Parts’ Business Case.” *International DAAAM Symposium* doi: 10.2507/28th.daaam.proceedings.146.

Mahbod, R., & Hinton, D. 2019. “Blockchain: The Future of the Auditing and Assurance Profession,” *Armed Forces Comptroller*, 64(1), 23–27.

Manta, O. 2017. “The virtual currency and financial blockchain technology Current trends in digital finance,” Research Gate.

Min, H. 2019. “Blockchain technology for enhancing supply chain resilience.” *Business Horizons,* 62(1), pp 35-45. doi:10.1016/j.bushor.2018.08.012.

Miraz, M. H., & Ali, M. 2018. “Applications of Blockchain Technology beyond Cryptocurrency,” *Annals of Emerging Technologies in Computing,* 2(1), 1-6. doi:10.33166/aetic.2018.01.001

Mylaavaram, S, C., & Mishra, R, K. 2018. “Blockchain Technology - An Exploratory Study on Its Applications,” *The Management Accountant.* June 2018. SSRN: https://ssrn.com/abstract=3194522

Norberg, H. 2019. “Unblocking the Bottlenecks And Making The Global Supply Chain Transparent: How Blockchain Technology Can Update Global Trade,” Research Gate. doi: 10.11575/sppp.v12i0.61839.

Panarello, A., Tapas, N., Merlino, G., Longo, F., Puliafito, A. 2018. “Blockchain and IoT Integration: A Systematic Survey,” *Sensors.* doi: 10.3390/s18082575.

Park, Sung-Hee, Lee, Lorraine, & Mun, Y Yi. (2011). “Group-level effects of facilitating conditions on individual acceptance of information systems.” *Information Technology and Management,* 12(4), pp 315-334.

Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M. (2008). “Systematic Mapping Studies in software Engineering,” in the *Proceedings of the 12th international conference on evaluation and assessment in software engineering*. UK: British computer society. 68-77.

Piscini, E., & Basir, M., & Datardina, M., & Stratopoulos, T. 2017. “Blockchain Technology and Its Potential Impact on the Audit and Assurance Profession.” Deloitte. 1-21.

Potekhina, A., & Riumkin, I. 2017. *Blockchain a new accounting paradigm*. Umeå School of Business and Economics. Retrieved from https://umu.diva-portal.org/smash/get/diva2:1114333/FULLTEXT01.pdf

Sarkintudu, S. M., & Ibrahim, H. H., & Abdwahab, A. B. 2018. “Taxonomy development of Blockchain platforms: Information systems perspectives,” in the Proceedings of AIP Conference, doi:10.1063/1.5055532.

Shah, T. & Jani, S. 2018. “Applications of Blockchain Technology in Banking & Finance,” Research Gate. doi: 10.13140/RG.2.2.35237.96489.

Subramanian, H. 2018. “Decentralized blockchain-based electronic marketplaces,” *Communications of the ACM,* 61(1), 78-84. doi:10.1145/3158333

Swan, M. 2015. *Blockchain: Blueprint for a new economy.* Beijing: OReilly.
Tapscott, D., & Tapscott, A. 2016. *Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world.* NY, NY: Portfolio/Penguin.

Tilooby, A. 2018. *The Impact of Blockchain Technology on Financial Transactions.* Dissertation, Georgia State University. Retrieved from “"https://scholarworks.gsu.edu/bus_admin_diss/103”

Trautman, Lawrence J. 2016. “Is Disruptive Blockchain Technology the Future of Financial Services?” *The Consumer Finance Law Quarterly Report* 232. SSRN.

Veiga, John F, Floyd, Steven, & Dechant, Kathleen. 2001. “Towards modelling the effects of national culture on IT implementation and acceptance.” *Journal of Information Technology*, 16(3), pp 145-158.

Venkatesh, Viswanath, Thong, James YL, & Xu, Xin. 2016. “Unified theory of acceptance and use of technology: A synthesis and the road ahead,” *Journal of the Association for Information Systems*, 17(5), pp 328-376.

Venkatesh, Viswanath, Morris, Michael G, Davis, Gordon B, & Davis, Fred D. 2003. “User acceptance of information technology: Toward a unified view,” *MIS quarterly,* pp 425-478.

Venkatesh, Viswanath, & Zhang, Xiaojun. (2010). “Unified theory of acceptance and use of technology: US vs. China,” *Journal of Global Information Technology Management*, 13(1), pp 5-27.

Wang, Ping. 2010. “Chasing the hottest IT: Effects of information technology fashion on organizations,” *MIS quarterly,* 34(1), pp 63-85.

Whitman, M., & Herman, M. 2016 *Principles of Information Security.* Fifth Edition, Boston MA: Cengage Learning.

Wikipedia contributors. 2018. Blockchain — wikipedia, the free encyclopedia. Retrieved from https://en.wikipedia.org/w/index.php?title=Blockchain&oldid=827331254. Online; accessed 27-February-2018.

Yingli, W., & Jeong, H. H., & Paul B. D. 2019. "Understanding blockchain technology for future supply chains: a systematic literature review and research agenda," *Supply Chain Management: An International Journal,* 24 (1), pp.62-84, doi:org/10.1108/SCM-03-2018-0148.

Zhang, Y., & Wen, J. 2016. "The IoT electric business model: Using blockchain technology for the internet of things," *Peer-to-Peer Networking and Applications*, 10(4), pp 983-994. doi:10.1007/s12083-016-0456-1.

Zhu X., & Badr, Y. 2018. "Identity Management Systems for the Internet of Things: A Survey Towards Blockchain Solutions," *Sensors.* 18(12), 4215. doi: 10.3390/s18124215.

DECISION SCIENCES INSTITUTE

Supply Chain Workforce Training: Addressing the Digital Skills Gap

Abbas Foroughi
University of Southern Indiana
8600 University Blvd., Evansville, Indiana 47712
(812) 465-1667 | aforough@usi.edu

ABSTRACT

New digital technologies are transforming supply chains, enabling end-to-end transparency, precision in planning and forecasting, enhanced customer relationships, and efficiency and cost savings. However, many employees lack the skills needed to use these technologies. This paper addresses the need for supply chain employee training and identifies currently available sources of training, including non-profit supply chain organizations, supply chain consultant companies, MOOC providers, and Micromasters degrees offered by colleges and universities.

KEYWORDS: Supply Chain Management, Digitalization, Corporate Training

INTRODUCTION

In the global economy, in which companies around the world have far-flung supply chain suppliers and widely diverse customers with demands for highly customized, rapidly delivered products, supply chain management has become front and center in efforts to cut costs and provide excellent customer service (MHI and Deloitte, 2018, 2019). From its beginnings as a mere operational function that focused on production lines and delivery to customers, the supply chain has evolved into a management function of its own that focuses on sophisticated planning processes like analytical demand planning and integrated sales and operations planning and is a crucial part of modern digital business (McKinsey & Company, 2018). The stature of supply chain management has risen, and supply chain leaders are now included in corporate strategic decision making (Mussomelli, Gish, & Laaper, 2016).

As supply chains begin to exploit digital technologies, more and more supply chain leaders are realizing the benefits—and the necessity--of transforming their supply chains for the digital business world (MHI & Deloitte, 2019). Firms now use a wide variety of technologies in their supply chains to increase visibility and efficiency, analyze data, and maximize the strategic nature of decision making (MHI & Deloitte, 2018; McKinsey & Company, 2018; Alicke, Rachor & Seyfert, 2016; Deloitte Development, 2017; Ellis & Santagate, 2018; EYGM, Limited, 2016; Gezgin et al., 2017; Hansen & Malone, 2019; Stank, Scott & Hazen, 2018). To keep abreast of new technology advances, workforce development is needed to update employee skills (Berry & Mok, 2018).

Ed Miller, CEO of NovoEd, an online corporate learning platform, said,

> *"As we enter the next Industrial Revolution, with some jobs being automated and entire new job categories being created, CEOs expect their employees to build new skills in areas needed for the future workplace. These include both technical and human skills in areas such as cloud architecture and data science as well as innovation, design thinking, collaborative team building, and learning how to learn. We have found that while many of these skills were once developed in face to face classrooms, they can now be delivered at scale and with greater impact in a collaborative learning platform…"* (Meister, 2018).

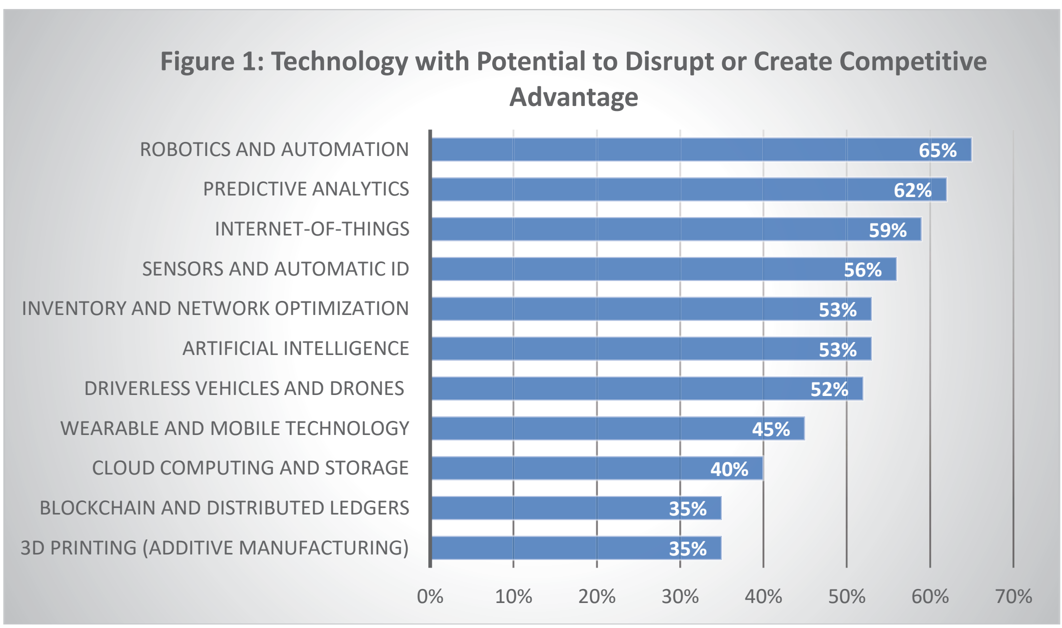
This paper explores digital technologies that are transforming supply chains, the need for skill upgrades and retraining, and sources of training that are available. In particular, answers are sought to the following questions:

1. What specific digital technologies are expected to have the greatest impact on supply chains?
2. What are currently available sources of the training needed to skill, reskill, and upskill supply chain employees to use these technologies?

The paper begins by describing the particular digital technologies that are transforming supply chains, followed by a discussion of the need for supply chain employee training, background on educational theories that are applicable in structuring workforce training, and results and insights about training options that are currently available for supply chain personnel.

DIGITAL TECHNOLOGIES THAT ARE TRANSFORMING SUPPLY CHAINS

New technologies are streamlining response times, cutting through complexity, reducing the assets needed to carry out operations—and transforming supply chains (Lyall, Mercier, & Gstettner, 2018). New sources of data and technological capabilities are enabling analytical decision-making and secure system interactions (Gezgin et al., 2017). Cognitive analytics, powered by AI and data captured through network automaton, can now do routine supply chain tasks, enabling employees to focus more on strategic and exceptional decisions. In a survey conducted for the MHI & Deloitte 2018 Annual Report (MHI and Deloitte, 2018), supply chain professionals identified the following eleven technologies as having potential to disrupt or create competitive advantage for supply chains (See Figure 1).



(Source: MHI & Deloitte, 2018, 5)

The following paragraphs describe the eleven technologies identified in Figure 1 and the role they play in digital supply chains.

Robotics and Automation

Robotics and automation are being employed for traditionally routine, manual tasks such as picking, packing, inspecting, classifying, and sorting orders; loading, unloading, and stacking; receiving and put-away; assembly operations; and processing such as welding, painting, and cutting (MHI & Deloitte, 2018; McKinsey & Company, 2018; Treblicock, 2018). This frees employees from dull, repetitive tasks in which mistakes can be made, and can operate round the clock. From automatic vehicles that could only travel along established pathways, robotics and automation now make it possible to use autonomous, self-driving vehicles that are equipped with sensors that provide information about the environment around them and can even work together to complete tasks (Alick, Rocher, & Seyfert, 2016; Treblicock, 2018; Deloitte Development, 2017; MHI &d Deloitte, 2018; Stank, Scott, & Hazen, 2018).

Predictive Analytics

Predictive analytics uses statistical models, data mining, and machine learning techniques to provide insights about data provided by digital supply chains and to identify the likelihood of future outcomes based on historical data. Predictive analytics uses big data, which are datasets of large size, high velocity (i.e., continuously growing via real or near-real time data acquisition), and variety (Ellis & Santagate, 2018). Big data are often unstructured and are derived from crowdsourcing, internet applications, direct from customers (i.e., point of sale, membership programs, etc.), and other sources. Models capture relationships among many factors to enable the assessment of risk or potential associated with a particular set of conditions. This is used to forecast consumer behaviors and risk, and also to gauge operations efficiency and maintenance needs (MHI & Deloitte, 2018) Businesses are using predictive analytics to maximize efficiency and predict anomalies across the IT infrastructure to prevent interruptions (Hanifan, Sharma, & Newberry, 2014).

The Internet of Things (IoT)

The Internet of Things refers to physical devices, vehicles, home appliances, and other items that are embedded with electronics, software, sensors, actuators, and network connectivity that enable them to connect and exchange data (Deloitte Development, 2017). Each object is uniquely identifiable through its embedded computing system and inter-operates within the internet infrastructure. IoT makes it possible to access or control remotely located objects across the network and creates opportunities for integration of the physical world into computer-based systems (Supply Chain Management Review, 2018). For supply chains, IoT is mostly used for real-time analytics, customer/market insights, customer and/or supplier collaboration, quality control, streamlined production, AI and machine learning, and cold chain temperature/humidity integrity (MHI & Deloitte, 2018; Lund et al, 2019; Supply Chain Management Review Staff, 2018; Stank, Scott, & Hazen, 2018).

Sensors and Automatic ID

Automatic sensory and identification technologies enable robots to sense, identify, and react to, other devices (Hanifan, Sharma, & Newberry, 2014; Lund et al, 2019). Radio Frequency Identification can identify a specific sensory-tagged item up to several meters away, without direct line-of-sight, and can also sense thousands of sensory-tagged items at the same time (Treblicock, 2018). Global Positioning Systems (GSPs) go even farther by identifying the exact location of tagged items. GPS receivers receive signals from GPS satellites that enable calculations of an object's position and time. Sensors on pallets track location, temperature, humidity, and tilt and shock (Hanifan, Sharma, & Newberry, 2014). Real-Time Locating Systems (RTLS), can identify and track the location of objects within a building or other enclosure, track products and supplies through an assembly line, and locate pallets in a warehouse. (Sainathan,

2018) Additionally, Bluetooth is now used as part of RTLS to track items, for indoor navigation, temperature track, and data generation. Bluetooth devices like smartphones can also monitor assets as they move around, and they send movement data to the supply chain cloud (Richardson, 2018).

Inventory and Network Optimization Tools

Inventory optimization tools can provide efficient and effective management of inventory throughout the supply chain, resulting in minimal cost for holding and storage (MHI and Deloitte, 2018; Excess stock level and shortage level are not static for every period, and inventory optimization's mathematical algorithms can determine the most probable excess stock level and shortage level. Inventory optimization tools can help a company achieve multi-echelon inventory optimization; cost-effective postponement strategies; stock keeping unit rationalization; optimization of all inventory components, including cycle stock, safety stock, raw material, work-in-process, and finished goods; strategic use of Vendor Managed Inventory (VMI); enhanced supplier intelligence for precise inventory planning and zero out-of-stock situations; demand forecasting and planning; and Just-in-Time (JIT) strategies (MHI & Deloitte, 2018; Stank, Scott, & Hazen, 2018).
Network optimization tools use mathematical modeling to reduce supply chain network complexities and improve responsiveness to customer needs by optimizing asset locations across the supply chain. Enterprise priorities and supply chain competencies in all areas of the supply chain network (sourcing, transportation, inventory, warehousing) are identified, prioritized, and mapped (Stank, Scott, & Hazen, 2018; EYGM, 2016). Network optimization makes supply chains more flexible, more efficient and more responsive to dynamic business scenarios and uncertainties, such as new product introductions, shifts in demand and consumption patterns, and changes in regulation. It enables companies to create accurate models of their end-to-end operations to identify major improvements in cost, service, sustainability and risk (MHI & Deloitte, 2018).

Artificial Intelligence (AI)

Artificial intelligence combines several types of other technologies in the simulation of human intelligence and the rapid solution of complex problems. AI can include machine and deep learning, reasoning, voice recognition, augmented reality, cognitive computing, natural language processing, and translation (Ellis & Santagate, 2018). These cognitive systems are adaptive, and they learn as information, goals, and requirements evolve, enabling them to resolve ambiguity and tolerate unpredictability; they interact easily with other processors, devices, and cloud services, as well as with people; and they are iterative, defining a problem by asking questions or finding additional source input if a problem statement is ambiguous or incomplete (Stank, Scott, & Hazen, 2018). They can also process multiple sources of information, including structured and unstructured digital information and sensory inputs (visual, gestural, auditory, or sensor-provided) to understand, identify, and extract meaning, syntax, time, location, appropriate domain, regulations, user's profile, process, tasks, and goals (MHI & Deloitte, 2017). AI technologies can enhance supply chain decision making, improve service elements, reduce cost, and shorten inventory turnarounds. In particular, AI technologies can improve demand prediction to better match supply with demand at lower cost, with less asset investment and with better response time; predict maintenance needs in manufacturing and transportation; check warehouse stocking levels to trigger re-orders; manage transportation mode and carrier selection by transaction; and manage and/or mitigate risk and disruption self-correction (MHI & Deloitte, 2017; Stank, Scott, & Hazen, 2018).

Driverless Vehicles and Drones

With enhanced vehicular sensor technology, driverless cars and trucks will be able to be more aware of their surroundings than is possible for human beings. They will be able to monitor a road continuously, judging the speed and distance of objects around them, and operating 24/7 (Wittemeier, 2017). This will necessitate the placement of logistics and service nodes along highways, so that trucks can notify an upcoming service station of their need for power/fuel, air for the tires, fluids or lubricants. Smart warehousing systems will direct the trucks to the correct inventory bay, where they will be unloaded by robot forklifts, while flying drones inspect and verify the load (Alicke, Rachor, & Seyfert, 2016; Wittemeier, 2017). Final Mile of delivery is already being performed by smaller vehicles or drones. Drones can fly up to 100 miles per hour and are faster than ground delivery, because they are not subject to traffic delays. They currently can only deliver limited-weight packages, but solutions to this are on the horizon. Additionally, it is possible that line-of-sight flying regulations may be lifted, making it possible for drones to fly and deliver packages in a wider span of locations (MHI & Deloitte, 2018; Treblicock, 2018).

Wearable and Mobile Technology

Factories are often very spacious venues. In the past, employees had to walk to and from stock to record/update data on stationery computers or to printout pallet labels, leaving a lot of room for error. Mobile devices now eliminate wasted “travel” time within a warehouse by enabling employees to input/update data and create and print labels right on the spot (Advantech, 2019). That real-time data is then available to everyone in the company, as well as to other outside stakeholders, such as supply and distribution partners. With mobile printers, workers can print out labels and reports on the spot, saving time and avoiding costly recording and reporting errors (Advantech, 2019). Wearable technologies are application-enabled computing devices, worn or attached to the body, that accept and process inputs from the Internet or from other devices. Workers equipped with wearable technology do not need to input information and can capture, hands-free, leads and sales updates via voice messages. Smart glasses graphically guide staff through the warehouse by virtual reality (Supply Chain 24/7, 2014; Michel, 2017; Van den Bossche, 2016). SmartWatch systems can enable managers to receive 24/7 emails, voice mails, text message, and notifications, giving them visibility into what is happening in their operations. Fitness bands can measure efficiency and ergonomics by tracking the steps required to complete particular operations. This information can be used to optimize storage locations to minimize the movements necessary to access and process inventory.They can also provide input about employee biometrics that identify activities that are overly stressful on employees, so that location adjustments can be made. Voice control headsets can also provide needed voice command updates and direct workers to issues or tasks (MHI & Deloitte, 2017).

Cloud Computing and Storage

Cloud computing and storage use a network of remote servers to access shared resources like data servers, storage, applications, and other services. The advantage offered by cloud computing and storage is that users can store and process data in a privately-owned cloud or a third-party server, making data readily accessible from anywhere. This has the effect of minimizing the cost of infrastructure and maintenance in information technology. Supply chain data stored on a cloud is available to anyone inside the company, as well as distributors and suppliers, so that all of the stakeholders can access the same accurate, real-time information (MHI & Deloitte, 2017; Alicke, Rachor, & Seyfert, 2016; MHI & Deloitte, 2017).

Blockchain and Distributed Ledger Technologies

Blockchain is a continuously growing list of digital records, like a spreadsheet of duplicated records that are linked, secured using cryptography, and continuously updated. There are thousands of duplicated records across networks, and they are public and easily verifiable. Blockchain can be used for transaction processing, records management, and other data-driven tasks. For supply chain clouds, blockchains can provide continuously updated and verified information from data inputs throughout the supply chain in shared ledgers that are available 24/7 to all enterprise stakeholders—internally and externally (MHI & Deloitte, 2018; Ellis & Santagate, 2018; Lund, et al., 2019; EYGM, 2016).

Additive Manufacturing

In additive manufacturing, structures are made by the addition of thousands of minuscule layers which combine to create a product. The process involves the use of a computer and special CAD software which can relay messages to the printer so it "prints" in the desired shape, one wafer-thin layer at a time. These layers are repeatedly printed on top of each other and are fused together during the process until the shape is complete. One of the greatest benefits is the greater range of shapes that can be produced, because. designs that can't be manufactured in one entire piece with traditional means can be created (Beyer, 2014). Additive manufacturing offers many advantages, including reduction in material inputs for leaner manufacturing, simplification and reduction in cost of production processes, lowered risk by providing a contingency plan, improvement in process flexibility and faster reaction to demand changes, and reduction in the cost of entry into new markets (Solel, 2018: Stank, Scott, & Hazen, 2018; MHI & Deloitte, 2018).

Workplace Transformation

The 2017 Gartner Workplace Survey predicts that digital technologies like those described above will transform the workplace of the future in a number of significant ways (Escherich & Jump, 2016; Hansen & Malone, 2019; World Economic Forum, 2018). Instead of having established roles within a company, groups of independent teams will be resized and reformed dynamically to facilitate the achievement of specific outcomes. Interestingly, as well as a good sign, a survey of high school students revealed that 4 out of 10 students prefer working for companies that encourage and support teamwork (Poitevin, 2016; Thurman, 2016). Algorithms will take over many managerial functions, functioning as "Robobosses" (Lee et al., 2015, 4; Cappelli, 2018), enabling fewer managers to oversee many workers in an optimized manner at a large scale (Lee et al., 2015; Poitevin et al., 2017). Upskilling and employee digital dexterity will be valued more than tenure and experience. "The Illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn and relearn" (Toffler, 1971, 414). The demand for employees with digital skills has already grown by 60% over several years, and this trend is expected to continue or even increase (Waller et al., 2017). Nonroutine work is expected to increase significantly by 2028, due to the growing automation of many routine jobs. (Autor & Price. 2013). Nonroutine work in the digital era, in which each day can present unexpected, dynamic issues and challenges, requires complex problem solving, effective teamwork, and digital dexterity. A conscientious habit of continuous, lifetime learning, as well as corporate support for it, will be necessary for employees to succeed in a challenging and evolving work environment like this (Manyuka et al., 2017; Griffin & Coleman, 2018).
It is also predicted that each employee will have at his/her disposal multiple smart machines, such as software, applications, avatars, artificial intelligence, and personal virtual assistants, which will be customized to suit each individual user's work needs (Buytebdijk, 2018). By 2018,

“robot resources” departments may become as common as Human Resource departments are today (Roth, 2018). Current high school students demonstrate a passion for volunteering in support of causes like education, health, social justice, and the environment (Thurman, 2016). They will bring the same commitments to the workplace and will view their jobs as more than just the opportunity to earn money. They will seek purpose in their professional lives and will support companies that have agendas for the causes that are important for them (Hurst, 2014; Hurst, 2017).

THE CRITICAL NEED FOR SUPPLY CHAIN EMPLOYEE TRAINING

At the same time that digital technologies like those described above are transforming supply chains and the workplace, employee expertise in mastering and using new technologies is more crucial than ever (Carnevakem, Smith & Strohl, 2014; Lund et al., 2018). It has been projected that by 2030, the work of approximately 375 million people, approximately 14% of workers worldwide, will be so disrupted by digital technologies that they will have to seek other types of work (McKinsey Global Institute, 2018; Berry & Mok, 2018). The following statement describes the need for workplace learning in general, and it applies to learning opportunities for supply chain employees, as well.

> “Leadership at all levels must embrace and require a continuous learning climate. Knowledge acquisition, sharing and dissemination across the organization must become a part of the day-to-day operations of the enterprise. The organization must provide an array of learning resources, tools, access and time devoted to gathering knowledge and using that knowledge to improve individuals as well as areas where it is needed in the organization” (Griffin & Coleman, 2018, 6-7).

The results of the *Industry 4.0 Paradox*, a survey conducted by Hanley & Holdowsky (2018), pointed out that, although strategic investment in digital supply chain efforts is critical, the supply chain is still not always recognized as a strong driver of innovation and transformation. As a result, the chief supply chain officer (CSCO) is often not included in digital investment decisions. There are signs, however, that this trend is changing--In 2004, only 8% of Fortune 500 companies had in place an executive in charge of the entire supply chain. By 2016, 68% of these companies had CSCOs (Mussomelli, Gish, & Laaper, 2016; O’Marah, 2016), indicating their recognition of the integral and vital role of supply chains in the achievement of digitalization.

For supply chain employers, in particular, the results of both the 2018 and 2019 MHI Annual Industry Reports, each of which surveyed more than one thousand manufacturing and supply chain leaders around the globe, identified the urgent need for a number of specific digital workforce skills (MHI & Deloitte, 2018; MHI & Deloitte, 2019).

Interestingly, the MHI Reports (2018; 2019) identified both technological skills as well as business, communication, and interpersonal skills, as important. The 2018 MHI report also noted that not all of these areas of expertise need to be possessed by one employee, but that they should be part of the collective expertise of a supply chain team (MHI and Deloitte, 2018). Although this current research recognizes the importance of business acumen, communication, and interpersonal skills, the current focus is technology skills needed by supply chain employees.

At the same time, supply chain managers have identified workforce challenges faced the most by supply chain managers (MHI & Deloittte, 2018). At the top of the list were finding (83%), hiring (76%), retaining talent (63%), and offering talented staff a compelling career progression (63%). Again, the need for training, uptraining, and retraining supply chain employees was underscored (Gexgin, et al., 2017; Lund et al., 2018; Lund et al, 2019; Stank, Scott, & Hazen, 2018; Alicke, Rachor, & Seyfert, 2016; MHI & Deloitte, 2017; Lyall, Mercier, & Gstettner, 2018). Similar results are reported in the 2019 Annual Industry Report (MHI and Deloitte, 2019), which also noted that companies in the early stages of digitalization are in a chicken-or-the-egg dilemma, in which they are aware that they seriously need digital supply chain talent, but cannot attract it because they have not yet adopted digitalization (Kane et al., 2016). The MHI 2019 report emphasized the importance of providing ongoing opportunities in the workplace for employees to develop digital skills; tying career growth to employee participation in training; using younger digitally-savvy employees to mentor older employees and assist in facilitating digital transformation; and establishing partnerships with STEM programs in universities that encourage the development of future employees with the skills companies need.

Competency-Based Learning and Social Learning

In the past, many companies provided employee training internally. However, the complexity of new technologies has often made it necessary to seek expertise outside the firm, from sources like non-profit supply chain professional organizations, consulting companies, MOOC providers, and online courses developed with universities. Workplace training programs increasingly use two pedagogical approaches—competency-based learning and social learning. With the mastery of specific skills in mind, competency-based learning is appropriate, because it can be used to define specific competencies or skills, and gives learners the flexibility to develop mastery of each competency or skill at their own pace and progress as they demonstrate mastery of academic content, regardless of time, place, or pace of learning (Exner, 2019). Competency-based learning emphasizes the customizability of content, according to the needs of the learner or a group of learners. Learners can develop a specific skill, for which they often receive a badge or other form of validated recognition. Such recognitions are often “stackable”, meaning that they can be combined to enable the learner to earn a full qualification, such as a certificate or diploma. Competency-based learning is increasingly delivered fully online, because many students taking such programs are already working or seeking work. Student engagement and outcomes improve, because the content is relevant to each student and tailored to their unique needs (Flesher & Carson, 2018). Currently, 41.7% of the most profitable companies in the world, in the Fortune 500, including Shell, Toyota, Paypal, and Booking.com, are using online tools for training (Wins, 29-019). Such solutions are helpful for training new employees and also help to make ongoing employee education both affordable and effective.
At the same time, social learning theory is very applicable for workplace learning, because it emphasizes the use of learning from others, mentorships, and online learning in place of formal learning methods like traditional curriculum-based education (Brown, 2017). With social learning, skill-building opportunities are relevant to specific employee needs, efficient, open, and accessible to all employees, while also offering them greater flexibility and control over their learning. Social learning also emphasizes how employees can learn from each other, either through participation in group learning activities or by observing others. Companies can create and maintain an atmosphere for ongoing learning, by rewarding employees who contribute to a culture of learning, creating corporate mentorship programs, and encouraging employees to connect, interact, and learn from each other. With platforms like Skype, Slack, Basecamp, and social media channels, employees can learn from each other by sharing informative content, links, ideas, and other information (BigThinkEdge, 2018).

Employee Training Preferences

Many firms are seeking a better understanding of the crucial need for employee skill development and of employee perceptions/preferences about workplace training (Coursera for Business survey, 2018). In winter, 2017, Coursera surveyed 750 managers and more than 1,000 employees about on-the-job training, yielding many valuable insights that can guide workplace education efforts. First, learning and development budgets are being increased, and

63% of those surveyed say that training available is now varied in content and variety. The greatest emphasis is on building a culture of continual learning (51%), followed by mentoring (45%), big data to analyze learning outcomes (35%), the use of virtual/augmented reality to improve training (33%), redesign of learning experiences (33%), new learning credentials (27%), gamification as a learning motivator (26%), artificial intelligence to guide employee development plans (24%), and the use of crowdsourcing content in learning (24%). Employers and employees agree that they both share responsibility for training—employers for providing it and employees for participating in it.

The survey continued by reporting that in the year to come (2018), 62% of employees plan to learn a new skill. Their reasons for wanting to learn are to improve performance in their current job (42%), build on existing skills in their current job (42%), learn skills for future jobs (36%), increase credibility and expertise (36%), stay up-to-date in industry (31%), feel more confident professionally (28%), learn new skills to complete a task (10%), and/or get a new job (9%). These employee motivators for seeking training were echoed in research by Egloffstein & Ifenthaler, 2016. Sixty-three percent of employees said that receiving high-quality training makes them more likely to continue to work for their current employer. Millennials, in particular, see training as an important perk in choosing a place of work. Hard skills receive more training funding (41%) over soft skills (23%), although employees rate soft skills higher. Training quality is valued the most (51%), over cost (29%), ease of implementation (22%), ability to assess employee learning (20%), and user friendliness (19%).

Only 48% of employees are satisfied with current learning experiences, and only 53% with the quality of learning content. Employees prefer content that is relevant and applicable (49%), learning that fits their schedule (32%), content that is valued by colleagues (28%), the ability to earn a valuable credential (23%), and lastly, content provided by the employer (20%). Sixty-five of employees say that lack of time is the number one obstacle to taking training, and 69% of employers agree. Employees prefer training that is 45 minutes or less in length, followed by 45-minute to 2-hour trainings, and then training that is half-day or longer. Employees also prefer more mobile training options, valued credentials, and online courses (as opposed to in-class training, studying training materials, attending conferences and webinars, and consulting co-workers and mentors). Employees (66%) show interest in certificates that can be earned quickly, and 76% of employers believe certificates are valuable.

It should be noted that the main tenets of competency-based training and social learning-based training—flexibility, online format, customized content, performance-focus, and recognitions—correspond closely to the employee workplace training preferences identified above in the Coursera survey. The following section evaluates sources of digital supply chain training and the extent to which they follow these training tenets.

SOURCES OF DIGITAL SUPPLY CHAIN SKILL TRAINING

An Internet search yielded a number of types of supply chain training options. These options fall into several general categories—professional supply chain organizations focused on certifications, supply chain consultants, MOOC providers, and university Micromasters programs.

In order to address the tenets of competency-based and social learning-based workplace training, and the employee preferences pointed out in the Coursera survey, the following criteria were used to evaluate supply chain training. The author began by choosing only those training programs that are available to complete online. Training providers who met this first criterion were included, and were subsequently evaluated according to the following: 1) supply chain content and/or digital technology content, 2) flexibility in terms of time frame and self-paced format, 3) customizable to the needs of particular employees or employee teams, and 4) credentials. It should be noted that training offered by corporations that focuses on the use of their specific products, such as Oracle, SAP, and IBM, were not included; nor were sites like Class Central, Udemy, Mooc-list.com, which are search engines for courses offered by other organizations and may include courses that are not validated. The supply chain training providers included in this study are described below.

Supply Chain Professional Organizations

Association for Supply Chain Management

The Association for Supply Chain Management (ASCM), formerly known as the American Production and Inventory Control Society (APICS), is the largest global nonprofit association for supply chain management, providing industry-leading APICS certifications and training (https://www.ascm.org/). Certifications include Certified in Production and Inventory Management (CPIM), the standard of professional competency in production and inventory management; Certified Supply Chain Professional (CSCP), which focuses on knowledge and organizational skills needed to develop streamlined operations; Supply Chain Operations Reference Professional (SCOR-P) Endorsement, which includes managing and measuring global supply chain performance using the SCOR model with the SCOR Professional program; and Certified in Logistics, Transportation and Distribution (CLTD), which focuses on -in-depth knowledge of a broad range of supply chain logistics topics. On-demand online courses are offered on Demand Management, Material Requirements Planning, Distribution and Logistics, Supply Chain Risk Management, and Inventory Accuracy. Individual employees or teams can study at their own pace, where and when it is convenient, thus avoiding the expense of off-site training and time away from work.

The Council of Supply Chain Management Professionals (CSCMP)

The CSCMP is a not-for-profit organization that works with academics and supply chain leaders to develop and offer degrees, certifications, and continuing education programs (https://cscmp.org/). In addition to SCPro™ Certification programs, CSCMP offers a variety of virtual and on-site educational options for supply chain professionals and organizations. An online Supply Chain Management Essentials (SCME) course covers supply chain management fundamentals, including supply chain planning, processes, and global operations. Self-paced Quick Courses are also taught via instructional web-based videos, each of which is 45-75 minutes in length.

The Institute of Supply Chain Management (IoSCM)

The IoSCM is a UK-based international organization and professional body that represents the interests of the supply chain industry (https://www.ioscm.com/).The organization encourages professional development, produces publications on supply chain management interests and offers accredited supply chain training courses and qualifications. The Institute of Supply Chain Management offers distance learning qualifications from entry level through to Senior strategic level to meet business and/or learner development requirements. Flexible distance learning packages enable students to adapt the course to study at their own pace and time. An easy to use online learning platform provides unlimited access to customized resources and course materials. Self-assessments, assignment preparation and tailored tutor feedback are available, as well as guidance and feedback from IoSCM's industry tutor team and individually assigned Student Support Liaisons. Students can connect with students and professionals across the globe, interact on forums, offer or receive peer support, and expand their network with IoSCM's online platforms. Accredited courses from levels 2 to 7 include Supply Chain Management, Manufacturing, Production and Planning, Logistics and Transport, Management and Leadership, Import and Export, Warehousing and Inventory, Ports and Shipping, Purchasing and Procurement, and Quality Management. Students can earn IOSCM awards, certificates, and diplomas.

The Institute for Supply Management® (ISM)

Institute for Supply Management® (ISM®) is the largest not-for-profit professional supply management organization worldwide (https://www.instituteforsupplymanagement.org/). ISM has over 50,000 members located in 100 countries. The ISM offers education, certification, leadership development and research, as well as ISM's Certified Professional in Supply Management (CPSM), and Certified Professional in Supplier Diversity (CPSD) designations. ISM features both self-paced and instructor-led online courses as well as short micro-learning videos that help supply management professionals master their craft. Each course focuses on specific competencies.

The ISM also maintains the ISM Mastery Model™, which includes 16 competencies and 69 sub-competencies that serve as an outline of the skills included in supply chain management. The sixteen competencies include business acumen, category management, corporate social responsibility, cost and price management, financial analysis, level, logistics management, negotiation, project management, quality management, risk, sale and operations, sourcing, supplier relationship management, supply chain planning, and systems capabilities and technology. For each competency, the skills needed at four levels (Essentials, Experienced, Leadership, and Executive) are detailed. The model was developed to guide supply chain professionals as they advance in their particular area of expertise. The competencies in the model are updated on a continuous basis to reflect trends and new developments in the supply chain field.

The ISM also maintains a list of institutions offering supply management-related degree programs, certificate programs, CPSM® review courses, and distance learning opportunities. Distance learning opportunities include courses offered via the Internet, video, satellite or other means. ISM’s Corporate Program addresses critical skills needed, delivered in flexible channels that can be adapted to meet the specific needs of an organization--category management, strategic sourcing, negotiating, supplier relationship management, contracting and legal aspects of procurement to risk management, strategy development and leadership.

Consulting Companies

Accenture

Accenture is a provider of supply chain management consulting and outsourcing services (https://www.accenture.com/us-en). Consulting courses include manufacturing and production management, purchasing and procurement, supply chain management, leadership and management, manufacturing and production management, purchasing and procurement, and supply chain management.

The Accenture Academy supports industry certifications provided by leading professional societies and provides an online study program to support the Certified in Production and Inventory Management (CPIM) review process and the Chartered Institute of Logistics and Transport (CILT) qualifications. Accenture Academy also offers an extensive, proven supply chain management curriculum, integrated with specialty and business management skills (in areas such as analytics, process excellence, project management, risk management, leadership and sustainability) that supply chain professionals must have to better partner with the business to achieve desired business outcomes. Courses are designed to help organizations achieve measurable business outcomes. Learning professionals work with organizations to keep learning plans aligned with changing strategies and business goals. Accenture Academy provides a supply chain management curriculum that is based on up-to-date functional competency models and job frameworks and is tailored to the goals of an organization and its employees. The academy assists in assessing an organization's current skills, analyzing skill gaps, and matching learning plans to those needs. The Accenture Academy's integrated supply chain management curriculum (https://www.accenture.com/us-en/service-supply-chain-operations) includes the following subject areas: Product Innovation and Lifecycle Management, Supply Chain Planning, Sourcing and Procurement, Manufacturing, Logistics, and Customer and Service Management.

Cognitive Class for Business

Cognitive Class, which is dedicated to addressing skills shortage in emerging technologies, is a source of courses on artificial intelligence (AI), machine learning, data science, blockchain, cloud computing, big data, analytics, and databases (https://cognitiveclass.ai/business/) The site offers a virtual lab in which students can practice what they learn. Customized learning paths can be tailored for specific training needs. Learners can earn digital badges, which can be shared as a recognizable form of achievement. Courses include Blockchain for Developers, Lightbend Reactive Architecture, Containers, microservices, Kubernetes, and Istio on the Cloud, Big Data with IBM, Applied Data Science with Python, Applied Data Science with R, Lightbend Reactive Architecture, Deep Learning, Text Analytics, Data Science Foundations, Big Data Fundamentals, Scala Programming for Data Science, Hadoop Fundamentals, Spark Fundamentals, Data Science for Business, Big Data Analytics, Hadoop Programming, Hadoop Administration, Hadoop Data Access, Data Management with IBM, and Cognitive Analytics with IBM.

Supply Chain Academy

The Supply Chain Academy is a Belgium-based consulting company that offers elearning courses on a variety of topics, including inventory control, demand planning, supply planning, order management, transportation, warehousing, manufacturing, procurement, legal issues, supply chain foundation, and project management (https://www.supplychain-academy.net/). The Supply Chain Academy combines best practice and current research to focus on competency development for corporations, providing supply chain course formats to meet an organization's specific objectives and challenges. Courses are scenario-based, using real life situations, and learners can work at their own pace. The academy also offers Virtual Application Labs, focused on implementing course content into practice, with focus on real-life implementation issues, knowledge development, and the analysis of theoretical concepts. Currently, 230 E-learning courses are available, covering nine supply chain competencies—supply chain basics, inventory control, supply planning, demand planning, transportation, order management, warehousing, manufacturing, procurement, legal issues, and digitalization and analytics in supply chains. Supply chain courses can be taken individually or in bundles, as Master Classes.

MOOC Courses

A growing trend is to facilitate employees in taking online MOOC (Massive, Open Online Courses) courses, some offered by universities and others through MOOC platform providers such as Cognitive Class, Coursera, edX, FutureLearn, iversity, NovoEd, and Udacity, MOOC courses are usually shorter than the usual university course and can be tailored to focus on a particular skill or technology which fits the needs of the company.

MOOCs (massive open online courses) were first developed by Canadian educators George Siemens (2005) and Stephen Downes (2008) as a way of implementing connectivist pedagogy, which views learning as totally learner-driven and consisting of the building of connections, with other people, information, and systems. In connectivism, learning is self-directed and not dependent on direction from a knowledgeable expert. This type of MOOC is mostly unstructured, with learners themselves creating a learning network as they share and build knowledge and connections. Connectivist MOOCs are often not associated with a particular university and are free and open to registration by anyone who cares to participate. A second major category of MOOCs is a type of mass enrollment course offered by major universities that features pre-taped lectures, a pre-defined body of topics to be covered, online quizzes, and other online assignments. Instructors at Stanford University held an xMOOC in 2011 and then formed Udacity, a for-profit company that provides a platform for xMOOCs. Coursera is another for-profit xMOOC platform that originated at Stanford. MIT and Harvard collaborated to form edX, an open source not-for-profit MOOC platform. Futurelearn works with a consortium of universities in the United Kingdom.
Recently, the MOOC concept has evolved into an alternative, viable pathway to education and training. MOOCs are no longer necessarily massive or open, and they now offer a number of types of credit/credentials for completion of courses. Credentialing can include verified certificates, completion certificates for all courses, college credit, learning pathways, specializations, and accredited degrees. These incentives have had a positive impact on course completion, an issue that has been proven to be a weakness of MOOCs in the past. With specific goals in mind, such as advancement at work, or upgrading undergraduate skills, students are more motivated to finish a MOOC course to earn the “carrot” of credit or credentials.

Coursera

Coursera is an online learning platform supported by Stanford professors that provides an online learning library in multiple subjects including sciences, marketing, engineering, and business (https://www.coursera.org/). The site currently offers over 2,000 courses and has more than 25 million registered users. Coursera offers 77 supply chain courses. Through Coursera, a student can take courses from instructors and universities around the world. Course certificates are given at the completion of a course. Specializations can be earned for mastery of a specific career skill. Professional certificates prepare students for passing an industry certification exam, Mastertrack™ certificates divides master’s degrees into online modules that can be used toward a full university master’s degree. Coursera for Business helps companies design a learning program for their employees, in topics like business, technology, leadership, social sciences, and health. Coursera offers a Specialization in Supply Chain Management that includes the following courses: Supply Chain Logistics, Supply Chain Operations, Supply Chain Planning, Supply Chain Sourcing, and Supply Chain Management Strategy.

edX

EdX, founded by Harvard University and MIT in 2012, is an online learning destination and MOOC provider that offers high-quality courses from 130 universities and institutions (https://www.edx.org/). Specializations include the MicroMasters program (standalone course or earn credit toward a Master’s Degree, Professional Certificates, and XSeries Programs (cutting-edge or in-demand fields), and XSeries Certificate programs, which offer recognition for the completion of a series of related MOOC courses, on the topics of Supply Chain Management, Education Technology and Aerodynamics.The MITx MicroMasters program in SCM program offers courses that include Supply Chain Analytics, Supply Chain Fundamentals, Supply Chain Design, Supply Chain Dynamics, and Supply Chain Technology and Systems.

FutureLearn

FutureLearn is a private company owned by UK's The Open University that offers courses from leading universities and cultural institutions from around the world that are accessible on mobile, tablet, and desktop (https://www.futurelearn.com/).FutureLearn for Business offers customized courses built around social learning pedagogy, creating collaborative online social learning where employees can learn from each other. Learning analytics can be used to track employee progress. Supply chain courses offered include Open Innovation, Business Futures: Understanding Omni-channel Retailing and Supply Chains, Farm to Fork: Sustainable Food Production in a Changing Environment, and Trust in Our Food: Understanding Food Supply Systems. FutureLearn courses are created by experts from leading universities and organizations. Learning can happen anytime, anywhere. Courses are customizable to fit specific organizational training needs. Online programs allow learners to deepen understanding of a subject and develop career relevant skills, with the chance to earn an academic or professional credential. FutureLearn Awards can be earned, which can include a professional or academic credential.

iversity

Iversity is a German-based online education platform that provides online courses and lectures in higher education, delivered through MOOCs (https://iversity.org/). iversity develops high-quality online courses in collaboration with corporate and academic experts, delivers them on a proprietary platform and distributes them to individual and corporate clients Students can choose between different study paths to earn certificates. iversity offers help to organizations in preparing customized courses for employees, with emphasis on the digital transformation of business. Ready-to-use online courses covering topics such as Agile Management, Predictive Analytics and Digital Marketing, can be used by companies. European Credit Transfer and Accumulation System (ECTS) credits can be earned by participants. iversity courses can be accessed anytime, anywhere, in an interative, online course community. iversity recently joined the brand portfolio of STEM publisher Springer Nature, to further enhance iversity's existing collaborations. Iversity's PRO courses help organizations prepare for the digital transformation of business. Companies can use a set of online courses on topics like Decent Work in Global Supply Chains, Critical Thinking, Predictive Analytics and Digital Marketing, or they can collaborate with iversity to create a customized, online course. Courses include Critical Thinking for Business, Presentation Skills for Business, Visual Thinking for Business, Digital Marketing – Strategies & Channels, Digital Marketing – Critical Success Factors, Predictive Analytics in Commerce, and Agile Management.

NovoEd

NovoEd, founded at Stanford University, specializes in social and collaborative learning for enterprises (https://www.novoed.com/). NovoEd's cloud-based learning platform creates and delivers global learning experiences to teach innovation and design thinking skills, and enhances learning through team-based projects, online team workspaces, social sharing and feedback, peer-to-peer feedback, and the design of customized platform content. NovoEd's philosophy is to advance the online learning experience by making courses more experiential, interactive, and collaborative. Students have access to lectures by leaders and professors from top universities and can also form teams with people around the world to work on class projects.

NovoEd also emphasizes the benefits of micro-learning, a type of learning with bite-sized and specific content, in answering the need for ad hoc learning from short and appealing content.

OpenLearn

OpenLearn is a free learning platform, delivered by The Open University, through which one thousand free courses are offered, featuring topical and interactive content, videos and online games (https://www.open.edu/openlearn/). Some OpenLearn courses carry a free digital badge, which is awarded in addition to a Statement of Participation. Courses are 24 hours in length and require learners to read every page of a course and pass compulsory online quizzes. OpenLearn offers numerous courses in supply chain, as well as the new digital technologies that are impacting supply chains--robotics, automation, predictive analytics, Internet of Things, sensors, automatic ID, inventory optimization, network optimization, artificial intelligence, driverless vehicles, wearable technology, mobile technology, cloud, blockchain, and 3-D printing.

Udacity

Udacity focuses on vocational courses for professionals, with emphasis on technology (https://www.udacity.com/). Most courses are free, but students can earn Nanodegrees, which are designed to build skills in a particular career area. Udacity partners with companies to develop a customized, end-to-end, managed solution to transform their workforce and build agility into the organization so they can operate at scale. Udacity offers nanodegree programs in critical areas where the need is high, and talent is hard to find. Udacity helps a company identify current levels of knowledge and generate custom learning paths for employees, which they can develop fluency and mastery on a timeline of the company's choosing. Employees learn and apply new techniques, analyze results, produce actionable insights, and build a dynamic portfolio of work. The Udacity for Enterprise program comes with an interactive dashboard (Enterprise Management Console), that allows administrators to manage employees, track course progress, perform bulk enrollments and more.

College and University MicroMasters Programs

A number of universities offer certificates in supply chain areas, some online and others on-site. The Institute of Supply Management maintains a database of universities that offer supply chain degrees, undergraduate and graduate, and certificates. in Supply Chain Management, Materials Management, Logistics, Operations, Purchasing, Contract Management, and Operations.These programs are not discussed here in detail, because most do not offer the ad hoc types of training that fit the needs of workplace learning.
Important exceptions to this are the MicroMasters programs offered on edX.org, which are very appropriate for workplace learning. MicroMasters programs have several major benefits for workplace learning—they are shorter in length than traditional Master's programs, they enable learners to engage in advanced study in an affordable and flexible manner, are career outcome-focused, recognized by top companies, affordable when compared to a standard master's degree program, and offer the flexibility of a self-paced program which workplace learners prefer (Gordon, 2018). The number of MicroMasters programs continues to grow, as it is recognized as a strong option for workplace skill development. Learners who successfully earn the MicroMasters Credential in Supply Chain Management are eligible to apply to a number of institutions across the globe to pursue a full Master's degree. MicroMasters can also be earned at Columbia University, University of Pennsylvania, Georgia Tech, Boston University, University of Michigan, UC San Diego, University System of Maryland, and Rochester Institute of

Technology (RIT) (https://www.edx.org/micromasters). In addition, the programs are offered at schools in other countries, including the University of British Columbia, Universitè Catholique de Louvain, and the University of Adelaide. MicroMasters programs are flexible and free to try, and they offer learners valuable knowledge to enhance their careers. Because these certificates are offered from prestigious universities, the programs are recognized by some of the top companies in the world, including Walmart, GE, IBM, Volvo, Bloomberg, Adobe, Fidelity Investments, Booz Allen Hamilton, Ford Motor Company, PricewaterhouseCoopers, and Equifax. MIT's MicroMasters in Supply Chain features five courses: Supply Chain Analytics, Supply Chain Fundamentals, Supply Chain Design, Supply Chain Dynamics, and Supply Chain Technology and Systems (https://scm.mit.edu/micromasters).

MicroMasters programs currently offered by several other schools have a strong focus on new technology skills areas—data science, artificial intelligence, Internet of Things, robotics—that are also important for supply chain professionals: MIT(Supply Chain Management; Data, Economics, and Development Policy), University of Maryland (Cloud Computing, Instructional Design and Technology, Bioinformatics, Software Testing, and Verification), University of Michigan (User Experience, Research and Design, Social Work, Leading Educational Innovation and Improvement), Columbia University (Business Analytics, Artificial Intelligence; UC San Diego, Data Science), Georgia Tech (Analytics: Essential Tools and Methods), University of Pennsylvania (Robotics), and Boston University (Digital Product Management) (https://www.edx.org/micromasters.)

AT&T's experience with skilling, reskilling, and retraining its employees provides a good example of a successful collaboration between a corporation, a supply chain training provider, and academia (Donovan & Benko, 2016). When AT&T moved to cloud services, it needed to upskill nearly 300,000 employees. Instead of searching for talent elsewhere, AT&T decided to retrain its employees. Human resources launched an online self-service platform, with tools and processes for performance management, career development, and talent planning. Udacity helped AT&T develop "nanodegree" programs to deliver training and certification in technical specialties like software engineering, coding, and web development. Nanodegrees can be completed in 6-12 months, at a cost of $200 a month for unlimited courses, and AT&T refunds tuition for all courses completed. Georgia Tech, Udacity, and AT&T collaborated on the development of an accredited online master's degree in computer science, which costs $6,600, compared to the $45,000 price of a campus-based program. The company offers up to $8,000 in annual tuition aid per employee for degrees and nanodegrees. Thousands of employees have earned nanodegrees and master's degrees. AT& T proved that a skills gap can be overcome by providing learning opportunities to employees that are flexible, relevant, and feature credentials, by using performance metrics to measure employees' achievements and value in the workplace, and by increasing financial compensation for employees with skills in high demand (Simon B, 2018).

DISCUSSION AND CONCLUSIONS

This paper has examined the current demand for rapid change in supply chains, due to new digital technologies being adopted that can transform a supply chain into a transparent, always-on, end-to-end intelligent network. Companies around the world are realizing the importance of supply chain in the digitalization of their enterprise and the benefits which this transformation can bring to them—end-to-end transparency, precision in planning and forecasting, tighter collaborations with suppliers and distributors, enhanced customer relationships, efficiency and cost savings, They also realize that none of this is possible without preparing employees for the new roles they will play in implementing a digital supply chain. Veteran employees have supply chain expertise built over years, understanding of the company and their particular function, and good business acumen. Newer employees grew up as digital natives who are more comfortable in learning and mastering new technologies, but lack knowledge and experience in the particular environment of an enterprise. Both veterans and novice workers need training to prepare them to use digital technologies to enhance the supply chain.

Fortunately, as this paper has shown, there are numerous sources of training that is customizable to each individual or team’s need. Employees can engage in training that begins where their knowledge and skills are and builds on this to bring them up to speed. All of the providers identified offer courses online, for both individuals and groups, with flexible start/finish parameters, credentials, and customizable formats. These course features are consistent with the employee preferences and the competency-based learning and social learning educational philosophies discussed above.

Table 1 shows the number of digital technology courses offered by each of the supply chain training providers included in this study.

Table 1: # of Digital Technology Courses Offered by Supply Chain Training Providers

Supply Chain Training Providers	Robotics/ Auto ID	Predictive Analytics	Internet of Things	Sensors/ Auto ID	Inventory Optimization/ Network	AI	Driver-less Vehicles	Wearable/ Mobile Tech	Cloud	Blockchain	3-D Printing
ASCM											
CSCMP											
IoSCM											
ISM											
Accenture	/1	24	15	/3	123/15	11	2		2		
Cognit Class	3/2	2	4	/3	/1	1		/3	33	8	
Supply Chain Academy	1	5	1		1	1					
Coursera	34/74	105	97	64/31	11/42	17/9	6	4/58	181	26	25
EdX	44/42	30	40	29/	7/12	41		/92	100	13	
Future Learn	9/2	4	6	5	/1	2		4/5	5	1	
iversity		2	2								
NovoEd						2					1
Open Learn			1			1			3	1	
Udacity	1/1	1				9			15	1	
Micro Masters	4/	3				4			4		

Table 2 shows the number of supply chain content courses offered, the number of digital technology content courses, and the coverage of employee preferences (flexibility, online/mobile format, credentials, customizable).

Table 2: Coverage of Supply Chain/Technology Courses and Employee Preferences						
	Supply Chain Content	Digital Tech Content	Flexible	Online, Mobile Format	Credentials	Customized
ASCM	5		yes	yes	yes	yes
CSCMP	6		yes	yes	yes	yes
IoSCM	10		yes	yes	yes	yes
ISM	8		yes	yes	yes	yes
Accenture	232	240	yes	yes	yes	yes
Cognitive Class	2	60	yes	yes	yes	yes
Supply Chain Academy	153	9	yes	yes	yes	yes
Coursera	61	937	yes	yes	yes	yes
edX	38	488	yes	yes	yes	yes
FutureLearn	12	56	yes	yes	yes	yes
iversity	2	4	yes	yes	yes	yes
NovoEd		3	yes	yes	yes	yes
OpenLearn		6		yes	yes	
Udacity		28	yes	yes	yes	yes
Micro Masters	5	15	yes	yes	yes	yes

Four sources of supply chain training were identified in this paper. First, supply chain training is available from non-profit professional supply chain organizations, such as the Association for Supply Chain Management (ASCM), the Council of Supply Chain Management Professionals (CSCMP), the Institute of Supply Chain Management (IoSCM), and the Institute for Supply Management (ISM). These organizations provide supply chain training, much of which is geared toward certifications, but tend to concentrate on more traditional aspects of supply chain management. They do offer flexible, customizable, credentialed course opportunities, but their courses do not feature the digital technologies that are so vital in organizational digitization. These providers are best recommended for employees who want to achieve a particular type of certification.

The second category of supply chain training providers is the consulting company, such as Accenture, the Supply Chain Academy, and Cognitive Class. The offerings of these providers are mixed. Accenture offers the most courses on traditional supply chain content (232) and also 240 courses that cover digital supply chain technologies. The Supply Chain Academy offers 153 courses in supply chain content and nine courses in digital technologies. Cognitive Class offers only two supply chain content courses but 60 digital technology content courses.

The third category is MOOC courses, available from Coursera, edX, FutureLearn, iversity, NovoEd, OpenLearn, and Udacity, with a wide variation of training content. Coursera offers 61 courses with supply chain content and 937 courses with digital technology content. edX offers 38 courses with supply chain content and 488 with digital technology content. FutureLearn offers 12 supply chain courses and 56 digital technology courses. Iversity offers 2 supply chain courses and 4 digital technology courses. NovoEd offers 3 digital technology courses. OpenLearn offers 6 digital technology courses, and Udacity offers 28 digital technology courses.

The fourth category of supply chain training providers includes colleges and universities that offer certificates and other advanced degrees in supply chain areas. These programs were not discussed in detail in this paper, because they are mostly long-term programs and not directly suitable for workplace learning applications. An exception to this is the MicroMasters programs that were developed with edX.org, the nonprofit online learning destination founded by Harvard and MIT. MIT offers a five-course MicroMasters degree in Supply Chain Management, and other universities offer MicroMasters programs in a number of technology areas that are important for employees to master for digitalizing supply chains. The number of MicroMasters programs continues to grow, as it is recognized as a strong option for workplace skill development.

This study is unique in addressing the specific needs of an organizational function—supply chains—as it moves to being digital, and in presenting an overview of the workforce development options available for employee skilling, reskilling, and upskilling. The results of this study reveal that, although many sources of supply chain training exist, they do not all cover the digital skills that supply chain employees need going forward. However, since digitalization of supply chains is, for the most part, still in the beginning stages in many companies, it is highly likely that formal certification standards like those offered by professional supply chain organizations like the Supply Chain Management, the Council of Supply Chain Management Professionals, the Institute of Supply Chain Management, and the Institute for Supply Management will update their certification requirements to reflect mastery of the new technologies discussed in the paper. The same is true for supply chain consultants, MOOC providers, and MicroMasters programs, who, as corporations begin to demand more upgrades to address the digital skill needs of their employees, will offer more advanced technology courses and/or will integrate them into their supply chain course offerings.

The results of this study are, of course, limited by the fact that the number of supply chain training providers identified is not exhaustive and because the study examined training opportunities available at a particular point in time. Course offerings discussed will change over time, certification requirements will be updated, and new providers will appear. Future research will most certainly reveal even more training opportunities for supply chain professionals. However, this study has provided a sense of the severity of the employee skill gap facing corporations as they move to digitize their operations, including their supply chains. By identifying sources of supply chain training that are currently available to help bridge this serious skill shortage, the results can serve as a guide to enterprises moving toward supply chain digitalization.

REFERENCES

Accenture website (https://www.accenture.com/us-en). Accessed 4.26.2019.

Accenture Supply Chain Curriculum, (https://www.accenture.com/us-en/service-supply-chain-operations), Accessed 4.18.2019.

Advantech (2019). How to select mobile pcs in an expanding supply chain technology climate. Advantech White Paper, January 4. https://buy.advantech.com/CMS/CmsDetail.aspx?CMSID=976770BA-9AE4-4F45-A18F-E59478115BDE&CMSType=White_Papers

Alicke, K., Rachor, J, & Seyfert, A. (2016). Supply chain 4.0—The next-generation digital supply chain. McKinsey & Company Operations Article.

Association for Supply Chain Management website (https://www.ascm.org/). Accessed 4.15.2019.

Autor, D.& Price, B. (2013). The changing task composition of the us labor market: An update of autor, levy, and murnane (2003)," *MIT Mimeograph*, June 2013, economics.mit.edu/files/9758.

B., Simon (2018). Three takeaways from at &t's talent overhaul, Aug 8. https://blog.sureskills.com/three-takeaways-from-atts-talent-overhaul

Berry, D. & Mok, L. (2018) Develop the competencies your workforce needs for the digital ecosystem. *Gartner.com*, Refreshed: 05 December 2018 | Published: 21 July 2017, G00325930. https://www.gartner.com/doc/3767836/develop-competencies-workforce-needs-digital

Beyer, C. (2014). "Strategic Implications of current trends in additive manufacturing." *Journal of Manufacturing Science and Engineering,* December, Vol. 136 (6).

Big Think Edge (2018). How Do You Apply Social Learning Theory in the Workplace? *Big Think Edge*, July 31. https://www.bigthinkedge.com/blog/apply-social-learning-theory-in-the-workplace

Brown, B. (2017). Social learning theory in the workplace. *Bizfluent.com*. Updated September 26. https://bizfluent.com/info-8405262-social-learning-theory-workplace.html

Buytebdijk, F., (2018). Digital connectivism tenet 1: We all have a digital identity, *Gartner.com*, November 1, G00314842. https://www.gartner.com/doc/3499417/digital-connectivism-tenet--digital

Cappelli, P. (2018) Are algorithms good managers? *Human Resource Executive,* February 20. http://hrexecutive.com/are-algorithms-good-managers/

Carnevakem A.P., Smith, N., & Strohl, J. (2014). Recovery: job growth and education requirements through 2020. Georgetown Public Policy Institute. https://cew.georgetown.edu/wp-content/uploads/2014/11/Recovery2020.FR_.Web_.pdf

Council of Supply Chain Management Professionals website (https://cscmp.org/). Accessed 4.16.2019

Cognitive Class for Business website. (https://cognitiveclass.ai/business/). Accessed 4.11.2019.

Coursera website. (https://www.coursera.org/). Accessed 4.18.2019.

Coursera for Business. (2017). Learning & development from both sides of the table: 21 Insights to maximize your l & d investment. file:///C:/Users/compaq/Desktop/new%20MOOC%20research/Coursera%20Enterprise%20L&D%20from%20Both%20Sides%20of%20the%20Table%20ebook%20v01.30.pdf

Deloitte Development, LLC (2017) Using autonomous robots to drive supply chain innovation: A series exploring Industry 4.0 technologies and their potential impact for enabling digital supply networks in manufacturing.

Donovan, J. & Benko, C. (2016). AT&T's talent overhaul. *Harvard Business Review*, October 16.

https://hbr.org/2016/10/atts-talent-overhaul

Downes, S. (2008). Places to go: Connectivism and connective knowledge. *Innovate: Journal of Online Education*, (5) 6, October/November.

edX website (https://www.edx.org/). Accessed 4.21.2019.

edX MicroMasters. (https://www.edx.org/micromasters). Accessed 5.1.2019.

Ellis, S. & Santagate, J. (2018). The path to a thinking supply chain. *IDC Custom Solutions Technology Spotlight*, August.

Escherich, M. & Jump, A. (2016), Market trends: The 24-hour user @ work — key trends for the digital workplace of the future. *Gartner.com*, December 1, G00315715. https://www.gartner.com/doc/3533117/market-trends-hour-user-

EYGM, Limited (2016) Digital supply chain: It's all about that data.

Exner, P. (2019). Why is competency-based training so effective? SkillBuilderLMS, May 1. https://www.skillbuilderlms.com/why-is-competency-based-training-so-effective/

Flesher, S. & Carson, S. (2018) Optimizing Competency-Based Learning. BaseCorp Learning Systems eBook. https://www.skillbuilderlms.com/wp-content/uploads/2018/05/Optimizing_Competency-Based_Learning.pdf

FutureLearn website. https://www.futurelearn.com/).Accessed 4.22.2019.

Gezgin, E., Huang, X., Samal, P., & Silva, I. (2017). Digital transformation: raising supply-chain performance to new levels. McKinsey Insights, McKinsey & Company.

Gordon, A. (2018). 'MicroMasters' surge as moocs go from education To Qualification. *Forbes*.com, Feb. 13.

Griffin, D. & Coleman, M. How will we work in 2018. *Gartner.com*, February 27, G00349777 http://planinonline.com.br/gartner-news/wp-content/uploads/2018/04/Pesquisa_GartnerDA.pdf

Hanifan, G., Sharma, A., & Newberry, C. (2014). Industry 4.0 and the rise of the digital supply chain: how digitalization can create a more transparent, efficient and responsive supply network. Arvato Bertelsmann White Paper.

Hanley,T & Holdowsky, J. (2018). The supply chain paradox. In Tim Hanley, Andy Doecher, Mark Cotteleer, and Jonathan Holdowsky, *The Industry Paradox*, Deloitte Development LLC.

Hansen, D. & Malone, T.W. (2019). The dawn of the intelligent enterprise: artificial intelligence and machine learning power the new workforce. *MIT Sloan Management Review*, Sponsor's Content (Oracle), January 7.

Hurst, A. (2014). The Purpose Economy: How Your Desire for Impact, Personal Growth and Community Is Changing the World. Boise, Idaho: Elevate Publishing.

Hurst, A. (2017). CEOs—2020 tipping point for purpose economy. *The Huffington Post*, December 6.

Institute for Supply Management website.(https://www.instituteforsupplymanagement.org/) Accessed 4.25.2019.

Institute of Supply Chain Management website. (https://www.ioscm.com/). Accessed 4.17.2019.

iversity website. (https://iversity.org/).Accessed 4.22.2019.

Kane, G.C., Palmer, D., Phillips, N., Kiron, D., & Buckley, N. (2016). Aligning the organization for its digital future. *MIT Sloan Management Review and Deloitte University Press*, July.

Lee, M.K., Kusbit, D., Metsky. E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. *CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Seoul, Republic of Korea, April 18 - 23, 2015, 1603-1612. file:///C:/Users/compaq/Downloads/2015-CHI_algorithmic_management.pdf

Lund, S., Mourshed, M., Rutherfoed, S., & Tyreman, M. (2018). Retraining and reskilling workers in the age of automation. McKinsey Global Institute, KcKinsey & Company, January.

Lund, S., Manyika, J., Woetzel, J., Bughin, J. Krishnan, M., Seong, J., & Muir, Mac. (2019). "Globalization in transition: The future of trade and value chains. McKinsey Global Institute.

Lyall, P., Mercier, P., & Gstettner, S. (2018). The death of supply chain management. *Harvard Business Review Blog*, June 15.

Manyuka, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., & Sanghvi, S. (2017). Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages.

MHI & Deloitte. (2017). The 2017 Annual Industry Report: Next-Generation Supply Chains: Digital, On-Demand and Always-On.

MHI & Deloitte (2018.) The 2018 Annual Industry Report: overcoming barriers to nextgen supply chain innovation.

MHI & Deloitte (2019). The 2019 MHI Annual Industry Report: Elevating Supply Chain digital consciousness.

McKinsey & Company (2018). The automation imperative survey, September.

McKinsey Global Institute Report (2018). November 17.

Meister, J. (2018). NovoEd Is poised for growth to lead a revolution in corporate learning. October 17. https://www.novoed.com/blog/6897/novoed-poised-growth-lead-revolution-corporate-learning/.

Michel, R. (2017). Information management: wearables come in for a refit. www.supplychain247.com, March 30.

Massachusetts Institute of Technology website. (https://scm.mit.edu/micromasters). Accessed 4.24.2019.

Mussomelli, A., Gish, D., & Laaper, S. (2016). The Rise of the Digital Supply Network: Industry 4.0 Enables the Digital Transformation of Supply Chains. Deloitte University Press, December 1, 2016.

NovoEd website. (https://www.novoed.com/) Accessed 4.23.2019.

O'Marah, K. (2016). What's the point of a csco. *Forbes*, May 26.

OpenLearn website. (https://www.open.edu/openlearn/. Accessed 4.23.2019.

Poitevin, H. (2016). Maverick research: When we are all freelancers without a boss, *Gartner.com*, September 27, G00306255.

Poiteven, H. Sicular, S., Lovelock, J.D., Rollings, M., Andrews, W., Roth, Cr., & Hetu, R. (2017). Predicts 2018: AI and the future of work. *Gartner.com*, November 28, G00342326. https://www.gartner.com/doc/3833572/predicts--ai-future-work

Richardson, A. (2018). 4 steps to digital success in 2019: Supply chain will lead digital transformation in 2019. *Kontakt.io Blog*, November 30.

Roth, C. (2018). High-Tech Webinar: To avoid working for robots, make robots work for you. *Gartner.com*, August 28, G00372925.

Sainathan, P. (2018). 6 Ways Bluetooth beacons can improve inventory & warehouse management. November 19. *Roambee blog*. https://blog.roambee.com/supply-chain-technology/6-ways-to-shrink-your-warehouse-space-and-cost-with-ble-beacons

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2, 3-10.

Solel, I. (2018). Additive manufacturing and the supply chain: opportunities and risks. *Supply Chain Brain*.com, October 17.

Stank, T., Scott, S., and Hazen, B. (2018). A Savvy Guide to the Digital Supply Chain: How to Evaluate and Leverage Technology to Build a Supply Chain for the Digital Age. University of Tennessee Global Supply Chain Institute's Technology in the Supply Chain Series.

Supply Chain 24/7 Staff. (2014). Next-Generation warehouse control system on the wearable smartwatch, March 16, *www.supplychain247.com*.

Supply Chain Academy website. (https://www.supplychain-academy.net/).Accessed 4.18.2019.

Supply Chain Management Review Staff (2018) How iot enables an intelligent and connected supply chain. SCMR White Paper, September 28. https://www.scmr.com/article/how_iot_enables_an_intelligent_and_connected_supply_chains

Thurman, S. Millennial career survey. (2016). *The National Society of High School Scholars*, April.

Toffler, A. (1971). *Future Shock*. Bantam Books: New York City, New York.

Treblicock, B. (2018). The warehouse of the future. *Logistics Management.com*, December 20. White Paper by Arizona State University, W. P. Carey School of Business.

Udacity website. (https://www.udacity.com/) Accessed 4.24.2019).

Van den Bossche, P. (2016). Wearable technology in the Warehouse. *www.supplychain247.com*, February 1.

Waller, G., Cain, Matthew, C., Olding, Elise, Papegaaij, Bard, Coleman, M., Poitevin, H., Berry, D., Raskino, M., Struckman, C., Scheibenreif, D., & Howard, C. (2017). Four ways for cios to cultivate digital dexterity in leadership and the workforce. *Gartner.com*, December 6, G00342145. https://www.gartner.com/doc/3836173/ways-cios-cultivate-digital-dexterity

Wins, M. (2019). How are top fortune 500 companies using eLearning to boost productivity? Supply Chain Academy Blog, April 12. Accessed 4.19.2019. https://www.supplychain-academy.net/top-fortune-500-companies-elearning-productivity/

Wittemeier, M. (2017). Drones, robots, and self-driving vehicles: reshaping the logistics industry. www.inform-software.com/blog/post, June 7.

World Economic Forum. (2018). *Global Risks Report 2018*.

DECISION SCIENCES INSTITUTE

Internet of Things and Artificial Intelligence for Smart Airports

Dothang Truong
Embry-Riddle Aeronautical University
Email: truongd@erau.edu

ABSTRACT

The Fourth Industrial Revolution, or Industry 4.0, promotes innovation and automation. In this new era, Internet of Things and Artificial Intelligence can drive the implementation of smart airports that will improve the airports' outcomes and productivity based on real-time and situation-specific conditions. The objective of this research is to propose a framework for implementing an Internet of Things based Artificial Intelligence (ITAI) system for smart airports and a research process for identifying and mitigating potential risks associated with ITAI implementation along with critical success factors.

KEYWORDS: Internet of Things, Artificial Intelligence, smart airport, innovation, machine learning, deep learning

INTRODUCTION

The Fourth Industrial Revolution, or Industry 4.0, promotes innovation and automation, in which machines are getting more advanced in predicting, learning, and managing themselves. In this new era, Internet of Things and Artificial Intelligence have driven the world to a new level of networked connection of people, systems, and smart products. They become the main drivers for innovation and automation to improve outcomes and productivity. Among public sectors, airports have a substantial contribution to economic growth. According to the FAA (2017), airports contributed $76 billion in total output to the U.S. economy in 2014. Airports maintain communications and interactions within their own internal operations, and among themselves, the airlines, and the passengers and guests who arrive and depart. Airports have numerous facilities and components that can be connected, but the number of components and parties involved often creates difficulties in ensuring safe and effective operations and providing satisfactory customer services. Industry 4.0 provides technological capabilities to create smart airports. There is no consensus definition of a smart airport in the current literature. Zmud et al. (2018) describe a connected airport as a system using a "variety of technologies through the Internet of Things, with the goal of improving the passenger experience and bringing monetary benefits to the host airport." Smart airports are more than just connected airports. Essentially, a smart airport is a system that uses an Artificial Intelligence platform to collect and analyze real-time data in a connected airport ecosystem and automatically solves airport business problems through generating optimal solutions in a real-time manner to enhance the safety and security, operational optimization, environmental sustainability, and financing solutions of the airport. In a smart airport, manual processing and human interventions are minimized to avoid human errors and delays and ensure the automation or airport safety and efficiency.

Internet of Things and Artificial Intelligence are critical components of a smart airport. Internet of Things is the third wave of information-technology-driven competition and promises more innovation, productivity gain, and economic growth (Porter & Heppelmann, 2014). The *Internet of Things* (IoT) is defined as “an infrastructure of interconnected objects, people, systems, and information resources together with intelligent services to allow them to process information of the physical and the virtual world and react.” (ISO, 2015). The IoT is characterized by monitoring capability (monitoring a product’s condition, operation, and external environment through sensors and external data sources), control capability (controlling through remote commands of algorithms built into the device or in the cloud), optimization capability (optimizing the performance through data analytics to improve output, utilization, and efficiency), and autonomy capability (incorporating autonomous product operation and enhancement, self-coordination, and self-diagnosis and service) (Porter & Heppelmann, 2014). An IoT system can create automatic and smart communications among objects, people, and systems, and IoT devices collect data in real-time and store them in the cloud.

The rapid growth of Artificial Intelligence (AI) creates a perfect combination with the IoT to enhance the automation. A combination of AI and IoT is a prerequisite for success in IoT-based digital ecosystems (Hwang, 2019). AI can be defined as a system that simulates human intelligence for either solving a problem or making a decision (Chouwdhury & Sadek, 2012). AI can learn, predict, improve, and solve. Through training, an AI application is able to generate optimal solutions and apply them to novel situations not encountered before. Thus, AI can process a large amount of data provided by IoT devices to make a prediction and automatically produce an optimal solution and even take actions. An *Internet of Things based Artificial Intelligence (ITAI)* system can be defined as a system that senses, monitors, connects, controls multiple components, and processes the data to automatically produce an optimal solution and take actions to solve business problems.

An ITAI system can enhance innovation and automation in airport ecosystems to create smart airports by connecting airport components, passengers, and stakeholders in a synchronized system. In this system, machine learning algorithms can be used to analyze data to detect unknown patterns, based on which ITAI can automatically produce solutions to solve business problems and take actions. Such a smart system allows airports to maintain safe operations at the highest level of effectiveness and, at the same time, provide passengers with satisfactory services. In order to implement this system successfully, an operational framework is needed. The current literature on applications of IoT and AI in airport research seems to focus on the use of a specific IoT device for a specific task rather than a comprehensive framework that shows the linkage between IoT devices, data storage, machine learning, and AI applications. Such a framework is needed to enable airport authorities to develop, deploy, and implement ITAI to enhance airport safety and operations.

Despite its advantages, ITAI is not without challenges. Airports need to meet strict technical, process, and managerial requirements to be able to successfully implement the ITAI. The biggest challenges are to connect all components and store a very large amount of data, deal with the complexity of interconnecting multiple components and data sources, and process big data with the aging infrastructure. Additionally, the ITAI is often considered a black box, which may pose some serious threats, such as technical (system and application failures), security, and privacy risks (Lykou et al., 2018). The risks may cause serious issues for airports, such as system failure, unavailable customer services, cyber-attacks, security vulnerabilities, and identity thefts. Airport executives need to gain sufficient understanding of the ITAI’s applications and components, implementation processes, and requirements. They must also learn about potential risks associated with ITAI implementation and know how to manage and mitigate these risks, so they can successfully implement the ITAI while ensuring the security and privacy of the airport and passengers.

The objective of this research is twofold: 1) to propose a framework for implementing ITAI for airports, including steps and requirements; 2) to propose a research process for identifying and mitigating potential risks associated with ITAI implementation at airports along with critical success factors.

LITERATURE REVIEW

IoT Studies for Airports

Despite the rapid growth of IoT in industries, IoT research studies in the airport environment are very limited. Cao et al. (2013) proposed a design for airport perimeter security system based on the IoT. The authors described the features and applicability of the airport perimeter security system using IoT. The authors conclude that IoT sensors provide a novel, low-cost, low maintenance choice to improve the capability to detect, locate, and classify intruders at airports. Wang et al. (2013) conducted a field survey that adopted IoT-based emergency monitoring and warning models to estimate the disaster losses and to prevent a secondary disaster from occurring. The authors used relative displacement sensing technology and Global System for Mobile Communications (GSM) technology to remotely monitor ground cracks in the landslide. Troiano and Pasero (2014) developed sensors to detect the presence of ice or water on the airport surface. They did experiments both in a laboratory and in the field to evaluate the repeatability, stability, and reliability of the sensor. Three sensors were embedded in the runway to check the state of the surface. Each sensor was connected to a GPRS modem to allow the data to be collected and stored. The state of the runway surface based on the collected data from the sensors is virtually represented the IoT features. Finally, Liu and Lu (2015) developed a perimeter intrusion detection system (PIDS) using sensors that detect illegal intrusion to an airport. The PIDS can accurately detect an intrusion, identify the type of intrusion, locate the intrusion target, and reduce false alarms due to unexpected interference. The PIDS has been successfully deployed and implemented by Shanghai Pudong International Airport. While these studies provide some interesting designs and applications of IoT systems for airports, they mainly focus on developing single systems using IoT sensors at airports. They lack the holistic view of airport systems and how AI can be used to support airport safety and effectiveness.

In order to determine the changes that airports need to make in a new era, Robinson (2017) explores the opportunities and challenges faced by airport authorities. The author indicates that the new digital age presents “the opportunity to not only reinvent the travel experience but also to develop solutions that result in cost-effective capital expenditure and optimize existing infrastructure.” The main challenge to the adoption on a global scale is the non-uniformity of the solutions. While digital solutions can be developed at an airport-by-airport level, the global air travel system will require strong collaborative initiatives across airports and stakeholders in different countries. In a more recent study, Zmud et al. (2018) developed a primer for airport operators and stakeholders in the Internet of things (IoT) within the airport environment. The authors described airport systems as a whole consisting of multiple connected components inside and outside of its boundaries. Additionally, the study described IoT enabling technologies that can be used for airports and obstacles to the implementation of connected devices and data sharing. The study also examined the passenger experience with IoT solutions at Atlanta International Airport (ATL), San Francisco International Airport (FSO) and Dallas Fort Worth International Airport (DFW).

Overall, the literature in IoT for airports seems to focus mainly on using a specific IoT device to accomplish a specific task, and the focus on a connected airport ecosystem is missing. As stated by Robinson (2017) and Zmud (2018), airport is a complex system consisting of multiple components that are connected to passengers and stakeholders. However, these studies are conducted at the conceptual levels and have not explored the role of AI in IoT systems for smart airports. A more comprehensive and robust system should be used at airports to take advantage of IoT and AI capabilities in the Industry 4.0 era. ITAI is a system that can automatically collect data, predict, learn, and generate solutions to help airport authorities solve business problems with minimal human intervention. Given the important role of this system for developing and deploying smart airports, further exploration and investigation are required.

IoT Components and Process

The objectives of IoT are to connect objects through smart technologies to be able to monitor and control these objects through a robust process. Table 1 presents the IoT components described in the literature. In an IoT system, physical objects are interconnected through a reliable network (Zmud et al., 2018). To connect physical objects and collect data in real-time, sensors, and communication devices are required (Madakam et al., 2015, Zmud et al., 2018, Deloitte Insights, 2018). Additionally, in order to ensure the consistency, validity, and usability of the data, uniform standards are needed for these devices (Deloitte Insights, 2018). Once the data are collected, they will be stored in cloud data storage, which should be large and powerful enough to manage the data and allows accessing data securely and effectively. Last, but not least, IoT requires analytics tools that can analyze the data to support the decision-making process (Madakam et al., 2015, Zmud et al., 2018, Deloitte Insights, 2018).

Zmud et al. (2018) and Meyers et al. (2015) also describe the IoT process, which typically starts with collecting data by smart sensors. Then data will be aggregated and analyzed to provide business insights, which in turn will lead to recommendations for decisions or acts. This process is to ensure the connected smart things can connect and operate smoothly in a connected system to support the decision-making process. Madakam et al. (2015) suggest some prerequisites for an IoT solution, including dynamic and real-time demand, access to data, end-user applications, security, and privacy.

Table 1: IoT components and process

Madakam et al. (2015)	***Zmud et al. (2018)***	***Meyers et al. (2015)***
IoT Components • Hardware (sensors, IP cameras, communication devices) • Middleware (demand storage and data analytic tools) • Presentation (visualization and interpretation tools)	IoT components • Physical objects (person, luggage, or boarding pass) • Instrumentation (smart components such as sensors or data collection system) • Connectivity (network-based device facilitating the interconnection between an object, its	IoT components • Sensors • Network • Standards • Augmented intelligence • Augmented behavior

	environment, and data management system) • Analytics (information gained from data analysis)	
<u>IoT Prerequisites</u> • Dynamic resource demand • Real-time need, exponential growth of demand • Availability of applications • Data protection and user • Privacy • Efficient power consumptions of applications • Execution of the applications near to end-users • Access to an open and interoperable cloud system	<u>IoT Basic framework</u> Data captured by sensors → Sensor data communicated and aggregated □ Sensor data analyzed to modify future acts	<u>IoT process</u> Business activities □ Sensors produce data □ Data is analyzed in the cloud □ Analysis leads to insights □ Decisions and actions

Roles of Machine Learning, Deep Learning, and Optimization Modeling in Artificial Intelligence Applications

Artificial Intelligence (AI) can be defined as a system that simulates human intelligence for either solving a problem or making a decision (Chouwdhury & Sadek, 2012). AI can process data to learn, predict, improve, and solve problems. Through training, an AI application is able to generate optimal solutions and apply them to novel situations not encountered before. In the airport context, faced with the challenges of growing air traffic, resource demands, increasing uncertainties, and operational complexity, AI can automatically predict issues in an airport ecosystem and provide real-time optimal solutions and decisions under uncertainties, hence solving the problem. However, AI is not without limitations and challenges. Since AI is often regarded as a black box, there are concerns about AI's capabilities to generalize to new situations or how AI determines the best decisions, hence some skepticism on AI's capabilities. It is also a challenge to integrate AI to current airport systems. Furthermore, AI could also be a potential liability, given its autonomous nature (Chouwdhury & Sadek, 2012). Some key components of AI include machine learning, deep learning, and optimization modeling (Sadek, 2007; Tien, 2017).

<u>Machine Learning</u>

Machine learning is a key component of an AI system. SAS defines machine learning as "a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention." (https://www.sas.com/en_us/insights/analytics/machine-learning.html). Machine learning is an important component of an AI system since it allows the AI to learn and uncover unknown patterns from large and noisy data and predict the desirable outcome (Sadek, 2007). The model can be adjusted and improved with real-time data feeding to the algorithms. Machine learning is usually used in a data mining process, which uses various machine learning algorithms to analyze large and noisy data to uncover new patterns that would provide information for business decision making (Nisbet et al., 2009). Popular machine learning algorithms include, but not limited to, regression, decision tree, subject vector machine, memory base reasoning, Bayesian network, neural network, random forest, and gradient boosting tree. Machine learning can be categorized into supervised learning and unsupervised learning. Supervised learning requires identifying a variable as a target variable and determines the relationships between the predictors and the target variable, hence, predicting this variable (Sarma, 2013). For example, prediction models can be used to predict the probability of an aviation incident. Truong et al. (2018) used decision tree and Bayesian networks learning algorithms to predict the probability of a flight delay between two cities. Additionally, Truong (2018) uses various machine learning algorithms to predict the risk of small unmanned aircraft systems (sUAS) violations in the National Airspace System. Unsupervised learning discovers probabilistic relationships among a large number of variables, without having to specify predictor and target variables (Sarma, 2013). For example, unsupervised learning can be used to detect anomalies in pilot behaviors during an unstable approach or passengers' perception of airport service performance.

Deep Learning

Deep learning is a growing field in the past few years and plays an important role in AI systems (Arel et al., 2010). Essentially, deep learning is an advanced class of machine learning algorithms that use Artificial Neural Network (ANN) to train the model. Unlike machine learning, which requires structured data through a data cleaning, coding, and variable identifying process to train the model, deep learning can automatically handle unstructured data without preprocessing. Deep learning learns the same way as the biological neural network of the human brain by using multiple hidden layers in ANN. Deep learning can automatically extract information from large and noisy data and detect patterns at a high level of accuracy (Smidhuber, 2015). Convolutional neural network (CNN), a commonly used algorithm in deep learning, is a multi-layer neural network designed for use on two-dimensional data, such as images, audios, and videos. Image recognition and classification are good examples of how CNN can be used to detect physical and human subjects in images or videos (Arel et al., 2010). Image or audio recognition can be used to improve passenger services at airports or to ensure airport security. Similarly, audio or text recognition can be used to detect anomalies in runway safety. While machine learning requires structured data, identified variables, and interferences of the data analyst, deep learning automates the entire process. In essence, deep learning makes AI work. Nonetheless, this does not necessary means machine learning is not useful in an AI system. In a system where a specific target variable is identified, a supervised learning model works best to predict the target variable and predictors contributing to this prediction. Such a prediction model would allow the system to predict the risk of a certain aviation incident, which can lead to a solution to mitigate the risk (Truong et al., 2018). Accordingly, when to use machine learning and when to use deep learning depends on the business problem and the organizational objectives.

Optimization Modeling

While machine learning and deep learning are great methods to discover unknown patterns which help detect a problem, an AI system needs to automatically generate a solution and even take actions. In order to do so, optimization modeling is needed to find the optimal solutions for the problem (Van Zuylen, 2012; Zhang and Xie, 2012). Optimization modeling is a decision science method that determines the optimal values of decision variables to achieve the objective function, given the constraints of resources. Methods such as linear programming, integer linear programming, goal programming, non-linear programming, and dynamic programming have been used extensively in many areas. In an AI system, such an optimization model can be developed based on the outputs of machine learning or deep learning, business resources, and business objectives. The optimal solution of this model along with sensitivity analysis provide the AI system necessary information about what actions should be taken to solve the problem.

PROPOSED INTERNET OF THING BASED ARTIFICIAL INTELLIGENCE (ITAI) FRAMEWORK FOR AIRPORTS

The capabilities of IoT can be combined with AI to create a smart system that can connect, sense, monitor, and control objects, collect data in real-time, learn, predict, produce optimal solutions, and take actions to improve the safety and operational efficiency of an airport. Figure 1 presents the proposed framework for Internet of Things based Artificial Intelligence (ITAI) implementation for smart airports.

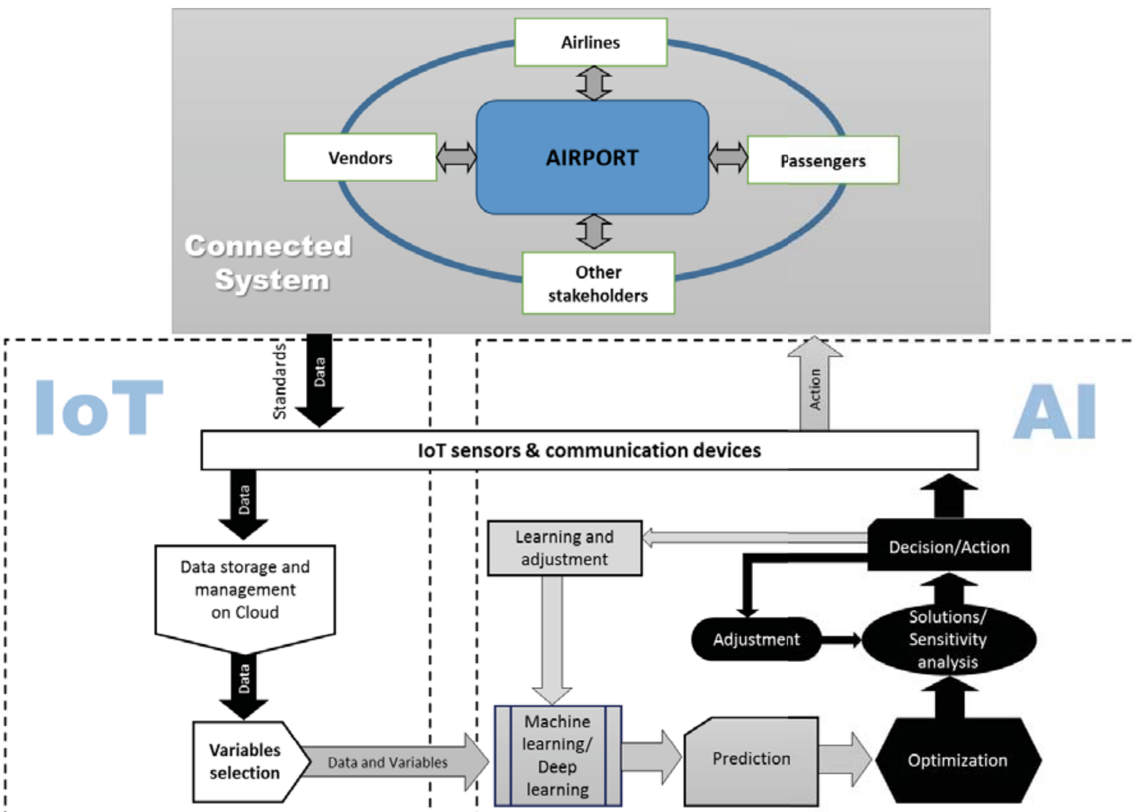


Figure 1: Internet of Thing based Artificial Intelligence (ITAI) framework for smart airports

The framework consists of the following components.

- **Connected airport system**: Airport is a complex transportation hub serving aircraft, passengers, cargo, and surface vehicles. An airport typically consists of airside facilities, landside facilities, and the terminal building (Office of Technology Assessment, 1984). In a connected airport ecosystem, the airport is connected with airlines, passengers, vendors, and other stakeholders, to ensure the visibility and effectiveness of airport operations.

- **IoT system**: A wide variety of IoT sensors and communication devices are installed and deployed at the airport. These devices are able to sense, monitor, and control objects. They communicate with each other to share and collect data throughout all airport-related operations. Uniform standards are needed to make sure sensors and devices are seamlessly communicated, and data are consistently collected in real-time. The data are collected in real-time and stored in the cloud through an effective big data management system. This system plays a critical role in ITAI, given the tremendous amount of data is collected continuously (Tien, 2017). The data need to be stored effectively and securely, and at the same time, must be accessible. Cloud computing provides a scalable solution for big data management and allows accessibility (Truong, 2010). Nonetheless, given the sensitivity of the data, the system must guarantee the security and privacy for the stored data. Furthermore, the system must be reliable to ensure uninterrupted operations. Finally, various data analytic methods can be used to explore the data and identify relevant variables.

- **Artificial Intelligence:** ITAI framework takes the data analytics to the next level by automating the entire process so airport operations can be done with minimal human interference. The first component of the AI in the ITAI framework is machine learning, in which machine learning or deep learning algorithms are used to detect unknown patterns. The process requires to develop, train, and validate prediction models. Then, these models are evaluated and compared based on pre-determined selection criteria, and the best prediction model with the highest prediction accuracy is selected. Machine learning algorithms can handle large and noisy data to build a prediction model with high predictive power (Sarma, 2013). This prediction algorithm is used to make a prediction with real-time data feeding to the algorithm. The prediction model can be automatically adjusted to improve its performance with new data coming in. Supervised machine learning can be used to predict a specific target variable, such as runway safety incident, maintenance issue, flight delay, or security issue. If unstructured data, such as images, audios, videos, and texts are provided, deep learning can be used to detect anomalies or patterns (Arel et al., 2010). Image and audio recognition can be used to enhance airport security, provide passengers with directions, or help them find lost items or needed services. Similarly, audio and text recognition can be used to detect anomalies related to runway safety or maintenance errors. Finally, an optimization module in the AI system uses the results of machine learning or deep learning to find an optimal solution to the airport business problem, such as to either maximize the performance or minimize the cost of operations. The optimal solution usually includes the values of decision variables. Additionally, sensitivity analysis can be conducted to evaluate various what-if scenarios to determine the best course of actions needed to solve the problem. Then, the AI makes the decision and take actions by sending commands to IoT devices, which control the connected objects to take corrections. The information about the decision and results is sent back to the AI system to allow this system to learn and make an adjustment. The machine learning will be continuously improved.

In order to implement this ITAI framework, specific steps should be followed. Figure 2 presents a process consisting of specific steps. These steps are to allow airport authorities to correctly deploy and implement the system to meet the business objectives.

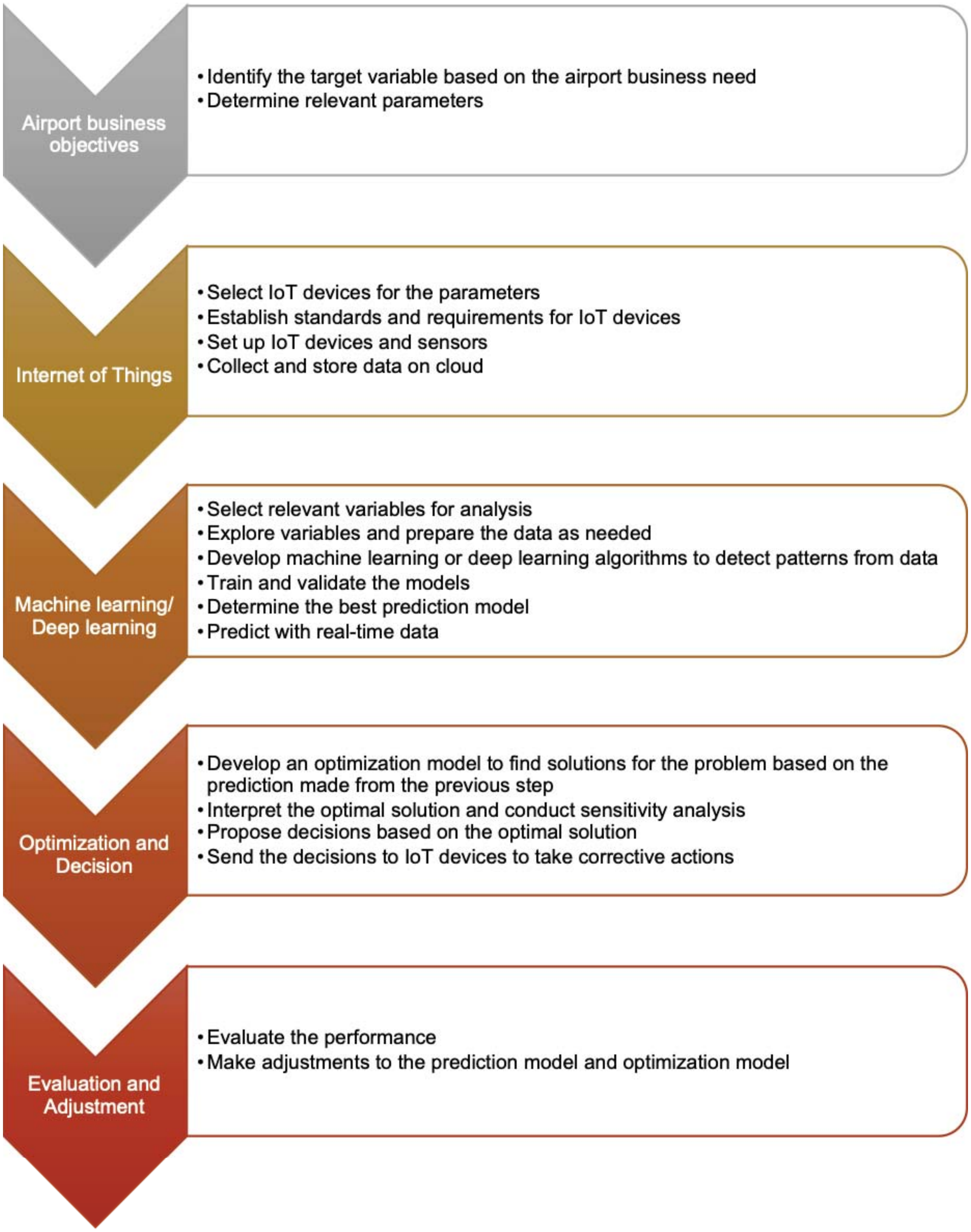


Figure 2: Steps for implementing the ITAI framework

PROPOSED RESEARCH PROCESS FOR IDENTIFYING AND MITIGATING RISKS

Since ITAI is considered a black box, there are risks associated with deploying and operating this system. In order to identify the risks and develop mitigation strategies, a combination of research synthesis and case study methods is proposed. In phase one, the research synthesis method will be used to collect information about the IoT, AI, and airport operations. The goal is to present the state of knowledge concerning innovation at airports through the implementation of ITAI and associated risks. This research will use seven steps proposed by Cooper (2010).

1. *Formulating the problem*. In this step, the specific objectives of the research will be identified. Important research questions include: What are the benefits and challenges of the ITAI? What are major applications of the ITAI at airports? What are technical and managerial requirements for ITAI deployment? How can airports successfully implement the ITAI to improve productivity and outcomes? What are the risks associated with using the ITAI, and how can they be managed and mitigated?
2. *Searching the literature*. Studies relevant to the objectives will be searched and collected from various resources: Embry-Riddle Aeronautical University's Hunt Online Library (which contains thousands of academic and professional databases), American Association of Airport Executives, Airport Services Association (ASA), The Internet Society, Internet of Things journals, Internet of Things Community IEEE, Artificial Intelligence journals, and Artificial Intelligence IEEE. Literature, including research articles, professional reports, technical reports, and cases will be collected and reviewed.
3. *Gathering information from studies.* The relevant information being gathered includes not only characteristics of the study, but also how the study was conducted and the results. The collected information will be stored and categorized for easy lookup and analysis. Bibliography and glossaries will be developed as well.
4. *Evaluating the quality of studies*. Quality of studies will be evaluated by examining the methodology to determine whether the data will be reliable and valid for addressing the research questions. Bad data will be discarded or given minimum credibility.
5. *Analyzing and integrating the outcomes of studies*. Collected data will be analyzed, summarized, and integrated into a unified picture.
6. *Interpreting the evidence*. The researchers will interpret the cumulative evidence and determine which conclusions are warranted by the data.
7. *Presenting the results*. Results will be presented to describe the investigation that answers the research questions.

In phase two, case studies will be conducted to provide best practices and lessons learned in implementing the IoT at airports. The results of phase one, along with case studies in phase two, will allow other airports to apply successfully the ITAI. Case studies will be conducted following Yin's (2009) process.

1. *Defining and designing case studies*. This research will use the two-case, embedded design to provide analysis of ITAI applications in airports (Yin, 2009). Since ITAI is new to airports, case studies will focus on collecting inputs from airport operators regarding smart airports and potential implementation of ITAI. US airports will be reviewed in order to identify two airports that have applied IoT sensors, big data analytics, and AI applications in their operations.
2. *Preparing case study evidence*. The case study protocol will be developed to ensure the quality of the data collection process. Interview questions and the survey questionnaire will

be designed to achieve the research objectives. The questions will be shared with subject matter experts for feedback to ensure the face and content validity.

3. *Collecting case study evidence*. Case study information will be collected through multiple sources, including interviews, survey, and archived records. To collect needed information and evidence, researchers will contact each airport manager, director of information technology, and director of operations. Data should include details of the ITAI applications, associated risks, risk management strategies, and critical success factors.
4. *Analyzing case study evidence*. Both quantitative and qualitative data analyses will be performed. Techniques such as pattern matching, explanation building, logic modeling, and cross-case synthesizing will be used in this research.
5. *Reporting case studies*. Results and findings of case studies will be reported to provide the audience with best practices in ITAI implementation and risk management. These results will be integrated and aligned with phase one results.

CONCLUSIONS

The rapid growth of IoT and AI in Industry 4.0 will drive airports to embrace automation and innovation to be able to improve and succeed. The results of this research will shed light on the applications of ITAI systems for smart airports. The requirements for ITAI implementation and deployment provide airport authorities and stakeholders with useful information about the investment needed to transform airports to smart airports to become more innovative and effective in the new era. The critical success factors, risk management strategies, and lessons learned are critical to airports' success in implementing and deploying the system. The findings of this research will also be useful for companies in other industries adopting an ITAI system.

REFERENCES

Arel, I., Rose, D.C., & Karnowski, T.P. (2010). Deep machine learning-a new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine*, November, 13-18.

Cao, X. M., Jing, C. M., & Zheng, X. D. (2013). Design of the airport perimeter security system based on the internet of things. *Applied Mechanics and Materials*, 361-363, 2276-2281.

Chowdhury, M., & Sadek, A.W. (2012). Advantages and Limitations of Artificial Intelligence. *Transportation Research Circular E-C168: Artificial Intelligence Applications to Critical Transportation Issues*, November, 6-8.

Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Los Angeles, CA: Sage Publications.

Federal Aviation Administration (2017). The Economic Impact of Civil Aviation on the U.S. Economy. *Federal Aviation Administration Report,* September. Retrieved from https://www.faa.gov/about/plans_reports/media/2017-economic-impact-report.pdf

Hwang, S. (2019). A network clock model for time awareness in the internet of things and artificial intelligence applications. *The Journal of Supercomputing*, February 9, 1-20.

International Organization for Standardization (ISO) (2015). Internet of Things (IoT). *ISO/IEC JTC1 Preliminary Report.* Retrieved from http://www.iso.org/iso/internet_of_things_report-jtc1.pdf

Liu, C. X., & Lu, K. (2015). A perimeter intrusion detection system based on sensor network for airport application. *Applied Mechanics and Materials*, 738-739, 50-55.

Lykou, G., Anagnostopoulou, A., & Gritzalis, D. (2018). Smart airport cybersecurity: Threat mitigation and cyber resilience controls. *Sensors (Basel, Switzerland)*, 19(1), 19.

Madakam, S., Ramaswamy, R. and Tripathi, S. (2015) Internet of Things (IoT): A Literature Review. *Journal of Computer and Communications*, 3, 164-173.

Meyers, M., Niech, C., & Eggers, W.D. (2015). Anticipate, sense, and respond - Connected government and the Internet of Things. A GovLab report in the Deloitte Future of Government series, *Deloitte University Press*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/iot-public-sector.pdf

Nisbet, R., J. Elder IV, and G. Miner. 2009. *Handbook of Statistical Analysis and Data Mining Applications* (1st Ed.). Burlington, MA: Academic Press/Elsevier.

Office of Technology Assessment (1984). Airport System Development. *Office of Technology Assessment, OTA-STI-231*, August. Washington, D. C.: U.S. Congress.

Porter, M.E. & Heppelmann, J.E. (2014). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review*, November. Retrieved from https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

Robinson, J. (2017). Passenger terminal development in the digital age. *Journal of Airport Management*, 11(4), 355-368.

Sadek, A.W. (2007). Artificial Intelligence Applications in Transportation. *Transportation Research Circular E-C113: Artificial Intelligence in Transportation Information for Application*, January, 1-6.

Sarma, K.S. 2013. *Predictive Modeling with SAS Enterprise Miner – Practical Solutions for Business Applications (*2nd Ed.). Cary, NC: SAS Press.

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61, 85–117.

Tien, J. M. (2017). Internet of things, real-time decision making, and artificial intelligence. *Annals of Data Science*, 4(2), 149-178.

Troiano, A., & Pasero, E. (2014). A runway surface monitor using internet of things. *Journal of Electrical Engineering*, 65(3), 169-173.

Truong, D. (2010). How Cloud Computing Enhances Competitive Advantages: A Research Model for Small Businesses. *The Business Review, Cambridge*, 15(1), 59-65.

Truong, D. (2018). Risk Assessment of Small Unmanned Aircraft System (sUAS) Operations. In C. Di Mauro (Ed.), *Decision Sciences Institute: 2018 Proceedings* (pp. 1-8). Chicago, IL: Decision Sciences Institute.

Truong, D., Friend, M, & Chen, H. (2018). Applications of Business Analytics in Predicting Flight On-time Performance. *Transportation Journal*, 57(1), 24-52.

Van Zuylen, H. (2012). Difference Between Artificial Intelligence and Traditional Methods. *Transportation Research Circular E-C168: Artificial Intelligence Applications to Critical Transportation Issues*, November 3-5.

Wang, H., Tuo, X., Zhang, G., & Peng, F. (2013). Panzhihua airport landslide (oct. 3rd 2009) and an emergency monitoring and warning system based on the internet of things. *Journal of Mountain Science*, 10(5), 873-884.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Los Angeles, Calif: Sage Publications.

Zhang, Y., & Xie, Y. (2012). Traffic Signal Timing and Optimization. *Transportation Research Circular E-C168: Artificial Intelligence Applications to Critical Transportation Issues*, November, 11-21.

Zmud, J., Miller, M., Moran, M., Tooley, M., Borowiec, J., Brydia, B., Sen, R., Mariani, J., Krimmel, E., & Gunnels, A. (2018). A Primer to Prepare for the Connected Airport and the Internet of Things. *Airport Corporative Research Program (ACRP)*. Washington, DC: The National Academies Press.

DECISION SCIENCES INSTITUTE

Performance impact of collaborative working capital management approaches: Empirical evidence from triadic supply chain sections

Philipp Wetzel
University of St. Gallen (HSG), Institute of Supply Chain Management
Email: philipp.wetzel@unisg.ch

Erik Hofmann
University of St. Gallen (HSG), Institute of Supply Chain Management
Email: erik.hofmann@unisg.ch

ABSTRACT

The purpose of this paper is to analyze the performance impact of collaborative working capital management (WCM) approaches. The results indicate that a relationship between collaborative WCM approaches and supply chain performance exists, but the direction and power of this relationship varies between the used proxies and the estimated performance metrics. The paper at hand contributes to SCF-related research by validating several proxy variables. Several practical implications are derived, demonstrating how companies can increase the efficiency and effectiveness of the financial supply chain network in which they operate.

KEYWORDS: Collaborative working capital management, Supply chain finance, Corporate performance, Financial supply chain management

INTRODUCTION

The relationship of working capital management (WCM) and corporate performance has received attention from researchers through a multitude of publications. Evidence has been brought forward that supports various shapes and directions of this relationship between WCM and firm performance: Jose, Lancaster and Stevens (1996) support the notion of a negative relationship between the amount of NWC and performance, Padachi (2006) found evidence for the opposite, whilst Singhania and Mehta (2017) found evidence that this relationship is of non-linear nature. However, those researchers have looked at WCM from a single firm perspective, and thus their findings only concern single firm performance. Hofmann and Kotzab (2010) were amongst the first authors to take a different approach to WCM by considering the effect of working capital decisions made by a firm on its suppliers and customers. From a conceptual standpoint they argue that in contrast to the classical paradigm, decision making in WCM should not be undertaken by a firm alone but should be aligned with the partner firms of the respective supply chain. In the long run, such supply-chain oriented collaborative WCM practices would be more performance enhancing than a myopic approach to WCM which focuses on keeping maximum liquidity within the own firm. A few authors subsequently took up this notion and identified leverages how collaborative WCM can contribute to supply chain and firm performance (e.g., Viskari & Kärri, 2012; Randall & Theodore Farris, 2009). In this paper, the claim that collaborative WCM is beneficial for performance is put to test. To do so, after a review of extant literature, proxies for collaborative WCM are developed and then applied in regression models to assess whether supply chains that exhibit a higher degree of collaboration in their working capital practices also demonstrate superior performance. Moreover, attention is given to the performance of the member firms within the supply chains, by testing whether the degree of WCM collaboration of a supply chain has an influence on performance of its member firms. The method used to identify and measure supply chains has first been introduced by Lanier, Wempe and Zacharia (2010) and is adapted to the use within the context of this paper.

THEORETICAL BACKGROUND

Single Company Perspective on Working Capital and Performance

The cash conversion cycle (CCC) is a WCM's key indicator for assessing the performance of individual companies or industries. According to Shin and Soenen (1998), CCC is calculated as CCC = (accounts receivable/sales)∗365+(inventories/sales)∗365−(accounts payable/sales)∗365. When looking at the existing literature on the relationship between working capital levels and profitability of a firm, extant evidence allows for multiple explanations of the nature of this relationship. Shin and Soenen (1998) propose a linear negative relationship using a sample of 58,985 firms and investigating a time span from 1975-1994. They demonstrated a significant negative relationship between CCC length and firm profitability which is supported by numerous studies such as Deloof (2003), Wang (2002) or Lazaridis and Tryfonidis (2006), which confirmed the findings of Shin and Soenen (1998). There are several theoretical explanations for these results. Exemplary, lower working capital implies lower financing and interest costs and thus improved shareholder value (Kieschnick, Laplante, & Moussawi, 2013; Brandenburg, 2016; Kim & Chung, 1990). In addition, firms may be forced to draw capital away from alternative value-enhancing projects if they have too much capital tied up in working capital positions (Deloof, 2003; Baños Caballero et al., 2014)
Alternatively researches propose a positive linear relationship. Higher investments in working capital is thought to increase firm’s profitability (Tauringana & Afrifa, 2013; Afrifa & Padachi, 2016). Reasons cited for the alternative view are rooted in considerations on the impact of higher working capital levels on business operations. For example, by extending payment periods (i.e., accepting higher DSO), firms may differentiate themselves from competitors in offering a unique selling proposition when there is low differentiation in the market, which in turn translates into better profitability measures (Shipley & Davies, 1991; Deloof & Jegers, 1996; George & Kirk, 2003). Such extensions of payment periods may be perceived as a price cut (Petersen & Rajan, 1997; Emery,1984), and reduce the risk for the buyer in using payment period to inspect the quality of the purchase (Biais & Gollier, 1997). Furthermore, according to Blinder and Maccini (1991) and Schiff and Lieber (1974) larger inventories reduce the risk of price fluctuations, distortions in the production process and of out-of-stock situations, which are costly because they limit a firm’s sales opportunities.
Ultimately, research papers tried to conciliate the two aforementioned theories of a negative and a positive relationship between working capital and performance. According to these recent studies, the relationship between levels of working capital and firm performance is a non-linear inverted u-shape, which implies that an optimal level of working capital exists. Previously proposed by Rafuse (1996), working capital is an investment to keep a business running, so that each firm has to manage a trade-off between too much capital tied up and too little working capital available for its operational daily business. This theory is supported by different research papers (Afrifa & Padachi, 2016; Baños-Caballero et al., 2014; Mun & Jang, 2015; Altaf & Shah, 2017; Pais & Gama, 2015). According to Baños-Caballero et al. (2014), the optimal level of working capital can differ between different firms, arguing that financially constrained firms have lower optimal working capital levels. It becomes apparent from the review of extant studies, that there is yet no consensus on the relationship between working capital and firm profitability.

Supply Chain-Oriented View on Working Capital and Performance

Studies which apply a multi-firm perspective have only been published more recently and emerged at the interface between the field of logistics and finance (Pfohl & Gomm, 2009). The SCF-orientation takes into account that management of working capital directly influences affiliated supply chains through accounts payable, inventories and accounts receivable along the supply chain partner firms. In particular, self-centered working capital improvements are made at the expense of affiliated supply chain partners. Hofmann and Kotzab (2010) used model calculations, to show that the optimization of working capital by a single firm in a supply chain does not add value to supply chain members. An illustrative example is that a DPO reduction by one supply chain member results in an increase of another members' DSO (zero-sum game). Grosse-Ruyken, Wagner and Jönke (2011) studied CCC's of different industry sectors and detected a strong negative correlation between average industry CCC and average industry ROCE. They thus argue that industries which collaborate in their WCM activities signified by a lower average industry CCC, are rewarded with a higher average industry ROCE.

Implications and Proposition Building

The aforementioned research outputs conclude that the application of a supply chain-oriented perspective to WCM will be beneficial for the performance of the supply chain. In their scenario, Viskari and Kärri (2012) found that the shift of inventory to upstream partners in a supply chain can reduce overall costs for the supply chain due to the increasing value of a product as it passes through the value chain (Porter, 1985). In addition, after a supply chain transaction is undertaken shifting the associated financial burden towards the supply chain partner with the lowest WACC poses an opportunity to lower incurred total supply chain costs and helps generate higher profits (Randall & Theodore Farris, 2009; Aberdeen Group, 2006).

Proposition 1: *Supply chains which shift working capital components between its members with the goal to minimize the cost associated with these components exhibit superior performance.*

Grosse-Ruyken et al. (2011) illustrate how the CCC is linked to profitability and provide evidence that industries with a shorter average CCC exhibit on average higher profitability. However, intense collaboration of all supply chain members, involving upstream and downstream partners, is necessary to achieve a sustainable reduction of the CCC. Hofmann and Kotzab (2010) showed that an isolated reduction of a CCC by a firm within a supply chain does not affect the collaborative CCC length at all. Such practices might even be detrimental to a supply chain in the long run. Weaker members will incur higher financing costs if their CCC is stretched by stronger supply chain members (Protopappa-Sieke & Seifert, 2017). This can lead to downstream delays, quality issues and ultimately higher costs of goods sold for the affected supply chain members, diminishing the competitive advantage of this supply chain (Hofmann & Kotzab, 2010).

Proposition 2: *Supply chains that manage their chain CCC collaboratively will exhibit superior performance.*

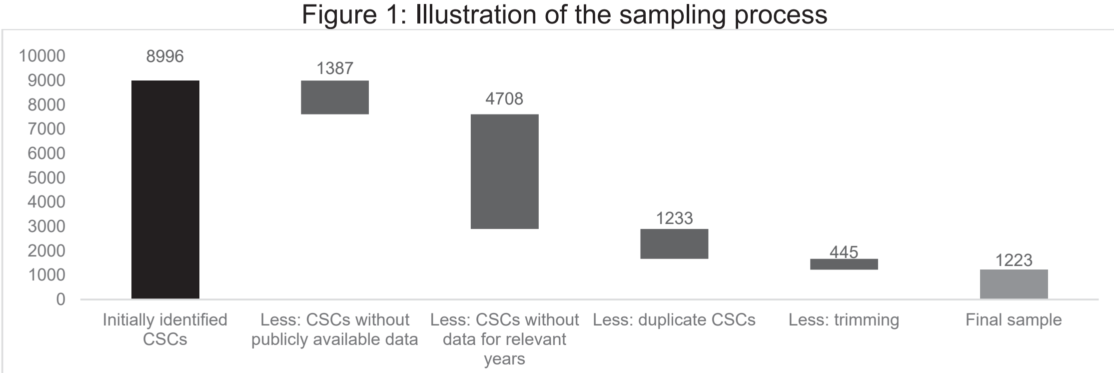
According to Rafuse (1996), working capital is an investment and comes with associated costs and benefits, and thus a reduction of working capital will enable the use of capital in a more productive way. Consequently, Hofmann and Kotzab (2010) identify the optimal collaborative CCC to be "the one that minimizes the cost of tied up capital […] across all collaboration members". Furthermore, simulations showed that a shared working capital pool in a supply chain overall is less costly than if members each keep their separate working capital stock, which is partly caused by the reduced overall requirement of working capital in case of a joint pool (Protopappa-Sieke & Seifert, 2017).

Proposition 3: *Supply chains which harmonize their members' activities to reduce the overall amount of net working capital in a supply chain exhibit superior performance.*

RESEARCH METHODOLOGY

The approach introduced by Lanier et al. (2010) is used to investigate financial performance of 3-firm supply chains adapted for the investigation on WCM. The Statements of Financial Accounting Standard No. 14 (1976) and No. 131 (1997), which require firms to disclose major customers, i.e. customer that account for 10% or more of a firm's sales in one year, is used for the construction of supply chains. If in the same year, a firm reports a major customer and appears as a major customer in the disclosures, then that firm is a focal company of a concentrated supply chain (CSC) (Lanier et al., 2010). The 3-firm CSCs were chosen as the subject of this study because the building of 3-firm CSCs allows for the creation of a larger sample as compared to larger CSCs, which will ceteris paribus lead to more meaningful results in a study (Littler, 2015). This will result in an increased robustness and validity of the results.

Figure 1: Illustration of the sampling process



The COMPUSTAT database delivered 137,063 major customer disclosures, which covered a time span from 1978 to 2018. This database is reduced to an initial dataset of 8,996 possible CSCs for the study (Figure 1) by identifying FCs that reported a major customer and were named as a major customer likewise. In the next step, identifiers are applied to screen the information that is needed for the calculation of the proxies resulting in a sample size of 7,609 CSCs. Subsequently, as not all data necessary for the calculation was available the number of CSCs was reduced to 2,901. Then, duplicate CSCs were removed in a screening within the reduced dataset resulting in 1,668 CSCs. Finally, to diminish the effects of statistical outliers and to improve robustness of the results, the dataset was trimmed using the 1.5x interquartile step fence, applied for all used proxy variables (Wilcox, 2001). This step delivered the final dataset, consisting of 1,223 CSCs, aggregated of 704 suppliers, 188 focal companies and 162 customers. Seeing that the database consists of an international sample of companies, which do not necessarily all calculate their income statement after the cost-of-sales method (some use the full-cost method), the inventory and payables turns are based on a company's sales.
The CSC performance metrics (dependent variables) EBITDA margin, EVA, profit margin, asset turnover and ROA are used to measure CSC performance. To diminish the risk of a regression

outcome that is biased by other influencing factors, the regression models are controlled using the following variables adapted from the study of Baños-Caballero et al. (2014). Size (natural logarithm of sales), leverage (ratio of total debt to total assets) and growth opportunities (ratio of book value of intangibles assets to total assets) represent the control variables. The calculation of metrics and variables on chain level follows the procedure outlined by Lanier et al. (2010). The five performance metrics (ROA, EBITDA margin, profit margin, EVA and asset turnover) are calculated on a chain level, interpreting the CSC as a quasi-entity. For the measurement of chain level oriented propositions one to three, proxy variables (formula (1)-(5)) are calculated for the supply chain in form of a supply chain as a quasi-entity (Lanier et al., 2010; Blois, 1972). This procedure allows the formulation of regression models, which then are used to assess the propositions for significance and power.

In order to evaluate proposition 1, proxy variable Inv_{shift} is used to measure the degree to which inventory is stored upstream within the supply chain as the associated costs for inventory are lower in the earlier stage of a good's production cycle (Viskari & Kärri, 2012):

$$Inv_{shift} = \frac{\left(Inv_{Supplier} + Inv_{Focal\ Company}\right)}{\left(Inv_{Supplier} + Inv_{Focal\ Company} + Inv_{\text{Customer}}\right) * \left(\mathrm{S}_{\text{supplier}} + S_{Focal\ Company}\right)} \quad (1)$$

where Inv_x means inventory held by the position within the supply chain (Supplier, Focal Company, Customer) defined in the index. $\mathrm{S_x}$ represents the relative sales of chain sales by the position within the supply chain (Supplier, Focal Company, Customer) defined in the index.
The second proxy variable NWC_{Costs} to assess proposition 1, builds on the contributions of Randall and Theodore Farris (2009) as well as Hofmann and Kotzab (2010) who recognize that firms within a supply chain may be exposed to different costs of capital. They argue that supply chains as a whole may benefit from such differences. NWC_{Costs} is used to calculate to which degree supply chain members with a lower WACC undertake financing activities for chain members with a higher WACC:

$$NWC_{costs} = \frac{\sum_{i=1}^{3}(NWC_i * WACC_i)}{\sum_{i=1}^{3}(NWC_i) * WACC_{chain}} \quad (2)$$

where $WACC_{chain}$ is calculated as the sum of $WACC_x$ of the members weighted by their relative sales (S_x) of the chain members. The first proxy variable CCC_{chain} to assess proposition 2 builds on the contribution of Hofmann and Kotzab (2010) who showed that harmonized WCM within the supply chain lead to shorter CCC. It is further supported by Grosse-Ruyken et al. (2011), who provided evidence on the negative relationship between CCC_{chain} and ROCE on industry level. Thus, CCC_{chain} is calculated as:

$$CCC_{chain} = CCC_{supplier} + CCC_{focal\ company} + CCC_{customer} \quad (3)$$

The second proxy variable $CCC_{balance}$ to assess proposition 2, builds on the contribution of Grosse-Ruyken (2011) who states that CSCs can increase their performance if they aim for a collaborative CCC which is balanced across member firms. According to Farrar and Glauber (1967), a log transformation in the calculation of $CCC_{balance}$ is necessary to prevent issues of multicollinearity with CCC_{chain}. An increasing $CCC_{balance}$ is perceived as a decreasing balance within the CCC. Thus, the proxy variable $CCC_{balance}$ attempts to measure the degree to which a CSC achieves an equal distribution of the collaborative CCC. $CCC_{balance}$ is calculated as follows:

$$CCC_{balance} = ln\left(\left(\frac{CCC_{chain}}{3}\right) - CCC_{supplier}\right)^2 + \left(\left(\frac{CCC_{chain}}{3}\right) - CCC_{Focal\ company}\right)^2 + \left(\left(\frac{CCC_{chain}}{3}\right) - CCC_{customer}\right)^2 \quad (4)$$

The proxy variable $NWC_{difference}$ to assess proposition 3 builds on the contribution of Protopappa-Sieke and Seifert (2017) who showed that a joint point of working capital in supply chains may reduce the overall need for working capital in supply chains. Thus, adapted from Vazquez, Sartal and Lozano-Lozano (2016) the proxy variable $NWC_{difference}$ attempts to measure the degree to which invested capital could have deployed in a more productive way (Rafuse, 1996).

$$NWC_{difference} = \frac{NWC}{Sales} - \frac{\overline{NWC}}{\overline{Sales}} \quad (5)$$

where $\overline{NWC}$ and $\overline{Sales}$ express the particular means. The following models are used to estimate effects of the before described proxy variables on the CSC performance metrics. Support for proposition 1 is expressed if regression (6) results deliver positive value for β_1 of Inv_{shift}. Support for proposition 1 is expressed if regression (7) results deliver negative value for β_1 of NWC_{costs}.

$$CSC_{metric} = \beta_1 Inv_{shift} + \beta_2 size + \beta_3 leverage + \beta_4 growth \quad (6)$$

$$CSC_{metric} = \beta_1 NWC_{costs} + \beta_2 size + \beta_3 leverage + \beta_4 growth \quad (7)$$

Support for proposition 2 is expressed if regression (8) results deliver negative value for β_1 of CCC_{chain}. Support for proposition 2 is expressed if regression (9) results deliver negative value for β_1 of $CCC_{balance}$.

$$CSC_{metric} = \beta_1 CCC_{chain} + \beta_2 size + \beta_3 leverage + \beta_4 growth \quad (8)$$

$$CSC_{metric} = \beta_1 CCC_{balance} + \beta_2 size + \beta_3 leverage + \beta_4 growth \quad (9)$$

Support for proposition 3 is expressed if regression (10) results deliver negative value for β_1 of $NWC_{difference}$.

$$CCSC_{metric} = \beta_1 NWC_{dfference} + \beta_2 size + \beta_3 leverage + \beta_4 growth \quad (10)$$

Each of the introduced models is used to estimate performance metrics, ROA, EBITDA margin, Profit margin, EVA and Asset turnover. Having five models and five performance metrics to be estimated results in a total of 25 regressions.

RESULTS

Table 1 shows the descriptive statistics for the independent and dependent variables giving insights into the distribution of variables. Exemplary, two variables are proven with corresponding literature. The mean CCC_{chain} is slightly higher than results of corresponding studies on firm level (Jose, Lancaster, & Stevens, 1996; Afrifa & Padachi, 2016), which is due to the created CSC with an increasing CCC along the upstream partners. The mean ROA of 0.050, is corresponding to the values found in similar studies (Baños-Caballero et al., 2014). Table 2 shows the estimation results of the derived models and assesses to what extent the proxies for collaborative WCM may explain CSC performance metrics.

Table 1: Descriptive statistics; N=1,223 for all coefficients

	Mean	Std. Dev.	Perc. 25	Perc. 50	Perc. 75
$\boldsymbol{Inv_{shift}}$	0.983	0.408	0.783	0.956	1.233
$\boldsymbol{NWC_{costs}}$	1.068	0.357	0.871	1.010	1.240
$\boldsymbol{CCC_{chain}}$	466.003	131.939	372.722	448.153	545.225
$\boldsymbol{CCC_{balance}}$	4.956	0.732	4.444	5.057	5.506
$\boldsymbol{NWC_{difference}}$	0.002	0.079	-0.051	-0.016	0.056

Table 2: Standardized coefficient output and standard error in parenthesis; *, **, *** represent significance below the 10%, 5% and 1% significance level

	ROA		EBITDA margin		Asset turnover		EVA		Profit margin	
$\boldsymbol{Inv_{shift}}$	0.110 (3.968)	***	-0.024 -(.853)		0.098 (3.424)	***	0.064 (2.596)	**	0.061 (2.113)	**
$\boldsymbol{NWC_{costs}}$	0.102 (3.703)	***	0.234 (8.513)	***	-0.221 -(7.941)	***	0.170 (7.121)	***	0.171 (6.034)	***
$\boldsymbol{CCC_{chain}}$	0.028 (1.023)		0.279 (10.495)	***	-0.333 -(12.645)	***	0.046 (1.914)	*	0.121 (4.321)	***
$\boldsymbol{CCC_{balance}}$	0.102 (3.744)	***	0.070 (2.512)	**	-0.026 -(.905)		-0.047 -(1.932)	*	0.103 (3.641)	***
$\boldsymbol{NWC_{difference}}$	0.019 (.589)		0.226 (7.034)	***	-0.473 -(15.744)	***	0.182 (6.554)	***	0.071 (2.138)	**

Inv_{shift}: The results indicate a significant influence of Inv_{shift} on ROA as well as asset turnover on the 1% significance level for profit margin as well as EVA on the 5% significance level. No significant influence of Inv_{shift} was found on EBITDA margin. The performance metrics where a significant influence was found, the direction is for all metrics positive, i.e. a higher value of Inv_{shift} would predict a higher value of ROA, Asset turnover, Profit margin and EVA. Regarding the power of the influence of Inv_{shift} in the statistically significant relationships, the highest standardized coefficient is found for the predictor of ROA, whilst the smallest power is found for the predictor of Profit margin.

NWC_{costs}: The results indicate a significant influence of NWC_{costs} on all performance metrics of CSCs on the 1% significance level. However, the directions of the relationships differ among the metrics. Whilst the performance metrics ROA, EBITDA margin, EVA and profit margin are positively influenced by NWC_{costs}, asset turnover is negatively influenced by NWC_{costs}. Regarding the power of the obtained regression results, the strongest statistically significant influence observed was the one of NWC_{costs} on EBITDA margin, whilst the least power was found for the influence of NWC_{costs} on ROA

CCC_{chain}: The results indicate a significant influence of CCC_{chain} on EBITDA margin, asset turnover, and profit margin on the 1% significance level and for EVA on the 10% significance level. However, the directions of the relationships differ among the metrics. Whilst the performance

metrics EBITDA margin, EVA and profit margin are positively influenced by CCC_{chain}, asset turnover is negatively influenced by CCC_{chain}. Thus, one out of five models support proposition 2.
$CCC_{balance}$: The results indicate a significant influence of $CCC_{balance}$ on ROA, as well as profit margin on the 1% significance level and for EBITDA margin and EVA on the 5% significance level. However, the directions of the relationships differ among the metrics. Whilst the performance metrics ROA, EBITDA margin and profit margin are positively influenced by $CCC_{balance}$, EVA is negatively influenced by $CCC_{balance}$. Regarding power of results, the strongest significant influence of $CCC_{balance}$. was found on Asset turnover, whilst the least powerful influence was found for ROA.
$NWC_{difference}$: The analysis of $NWC_{difference}$ yielded mixed results. The level of working capital present within a supply chain seems to have a significant influence on chain level performance, as all but the model that attempted to estimate ROA with $NWC_{difference}$ found a significant relationship between $NWC_{difference}$ and chain performance, for Asset turnover, EBITDA margin as well as EVA on the 1% significance level and for performance metric Profit margin on the 5 % significance level. There is no common direction of the influence of $NWC_{difference}$ on the performance metrics where statistical significance was detected. Whilst $NWC_{difference}$ is positively correlated with EBITDA margin, Profit margin, and EVA the regression that estimated Asset turnover using $NWC_{difference}$ found a negative correlation Regarding the power of the obtained results, the standardized coefficient in models with statistical significance was largest for Asset turnover, and smallest for Profit margin.

DISCUSSION

Regarding proposition 1, the results strongly indicate a relationship of shifting working capital components between its members and chain performance, as the regression models that test the influence of Inv_{shift} and NWC_{costs} on the performance metrics exhibit significant results. The effect of shifts of inventory towards upstream supply chain partners was investigated with models that employ proxy Inv_{shift}. These models showed that supply chains which exhibit a high level of Inv_{shift} are rewarded with a higher ROA, Asset turnover, EVA and Profit margin. When looking back at the interpretation of these financial ratios, it thus can be concluded that such supply chains are deploying their assets more efficiently to generate sales (positive relationship between Inv_{shift} and Asset turnover), returns on their assets (positive relationship between Inv_{shift} and ROA), economic value (positive relationship between Inv_{shift} and EVA) and turning sales into net profit (positive relationship between Inv_{shift} and Profit margin). The strongest effect, measured using the standardized regression coefficients, is found in the relationship of Inv_{shift} on chain ROA. This strongly supports proposition 1. The analysis therefore complements previous research on WCM in the interorganizational context (Viskari & Kärri, 2012). It is also in line with previous work that has shown that inventory stored upstream, when product value and associated inventory costs are lower, a supply chain may reduce its overall inventory costs (Randall & Theodore Farris, 2009). According to Hofmann and Kotzab (2010) the synchronization of the goods/material and financial flows within the supply chain is a key to improve CSC performance. The effect of shifting net working capital within CSCs on the performance metrics was investigated with models that employ proxy NWC_{costs}. Most performance metrics seem to be affected in the opposite direction as postulated by proxy NWC_{costs}. A lower level of NWC_{costs} is related with a lower ROA, EBITDA margin, profit margin as well as EVA and is related with a higher asset turnover. Thus, the models that predict influence of NWC_{costs} on chain performance metrics do not support proposition 1 since only one out of five assessed performance metrics support proposition 1.

Regarding proposition 2, the results do not indicate a strong relationship between a collaborative CCC and chain performance, as the regression models that test the influence of CCC_{chain} and $CCC_{balance}$ on the performance metrics exhibit mixed results. Supply chains which strive to balance their collaborative CCC cannot expect to achieve superior performance, in fact they face a worsening of chain level performance: A negative effect of such efforts can be observed on a supply chain's ability to generate return on assets (positive correlation of $CCC_{balance}$ and ROA*)*, its ability to minimize operating expenses needed to produce sales (positive correlation of $CCC_{balance}$ and EBITDA margin) and its ability to turn its sales into net profit (positive correlation of $CCC_{balance}$ and Profit margin). However, the results indicate that CSCs may increase chain EVA, if they strive for a more balanced collaborative CCC. Thus, regarding proposition 2, the models using $CCC_{balance}$ reject proposition 2 for four of the assessed performance metrics and offer support for one performance metric. The effect of the length of the collaborative CCC on the performance metrics was investigated with models that employ proxy CCC_{chain}. Most performance metrics seem to be affected in the opposite direction as postulated by proxy CCC_{chain}. A higher level of CCC_{chain} is related with a higher EBITDA margin, EVA and profit margin and is related to a lower asset turnover. Thus, the models that employed CCC_{chain} to predict chain performance metrics do not support proposition 2. The effect of supply chains that strive to balance their collaborative CCC on the performance metrics was investigated with models that employ proxy $CCC_{balance}$. Thus, the models that employed $CCC_{balance}$ to predict chain performance metrics reject proposition 2 for four of the assessed performance metrics and offer support for one performance metric.

Regarding proposition 3, the results do not indicate a clear relationship between a collaborative CCC and chain performance, as the regression models that test the influence of $NWC_{difference}$ on the performance metrics exhibit results for both supporting and rejecting proposition 3. The effect of collaboration in WCM on the overall working capital level was investigated with models that employ proxy $NWC_{difference}$. The applied regression models found correlation between $NWC_{difference}$ and four out of five assessed chain performance metrics. However, the direction of this relationship varies between the assessed performance metrics. Having larger amounts of net working capital seem to enable supply chains to improve their capability to turn generate economic value (positive relationship between $NWC_{difference}$ and EVA), minimize operating expenses needed to produce sales (positive relationship between CCC_{chain} and EBITDA margin) and to achieve a higher Profit margin. On the other hand, having lower levels of net working capital correlates within efficient use of chain assets to produce sales (negative relationship between $NWC_{difference}$ and Asset turnover). Thus, evidence to support proposition 3 is offered solely by the model that estimated Asset turnover with $NWC_{difference}$. Thus, the models that predict influence of $NWC_{difference}$ on chain performance metrics seem to reject proposition 3. Although four out of five assessed performance metrics measured a significant influence, evidence to support proposition 3 is only offered for one out of five performance metrics.
Asset turnover compares sales to total assets, thereby revealing to the analyst how efficiently a business has used its assets to generate sales revenue (Warren & Reeve, 2006). Since asset turnover is a coefficient to measure the efficiency of a business entity one might assume that a business that uses working capital more collaboratively uses its assets more efficiently to generate sales revenue. However, there is no clear evidence of the effect on ROA. It is calculated as the fraction of profit margin by asset turnover. Thus, the root cause might be in the profit margin, since there is no clear evidence for profit margin likewise as shown in Table 1. Profit margin is an indicator of the overall profitability of an entity as it shows the net income relative to sales, which shows the percentage of sales that remained in the company as net profit (Metha, 2016).

CONCLUSION AND OUTLOOK

This study makes several theoretical contributions to SCF-related literature. First, a relationship between collaborative WCM practices and the performance of supply chains and their member firms has been proven to exist. Second, there is strong evidence that collaborative WCM influences the efficiency of the asset usage in a positive direction. Third, the measured contradicting direction of the relationships between WCM and chain performance mirrors the current state of knowledge on the relationship between collaborative WCM practices and chain performance. Future studies should apply more sophisticated modeling approaches to isolate the performance impact of collaborative WCM approaches. This would potentially lead to more consistent results because special effects arising during a certain business period would no longer have a major impact on the considered performance metrics.

The exploratory character of the paper leads to certain limitations. Using COMPUSTAT as a secondary source to retrieve the necessary data, there is no certainty that the data used is fully accurate and that the companies worked collaboratively on purpose to increase chain performance. The proxies used to measure the degree of collaboration have been, with the exception of $NWC_{difference}$, untested constructs derived based on extant literature. Hence, the proxies could contain errors as the analysis gives partially indistinct results.

In future studies researchers who take on the challenge of improving the database could gather primary data. According to Lanier et al. (2010), the receipt of survey-based measures could add greater validity to any deductions made. Additionally, future SCF research should aim to validate or improve the used proxies for collaborative WCM.

REFERENCES

Aberdeen Group. (2006). *Supply Chain Finance Benchmark Report: The New Opportunity to Improve Financial Metrics and Create a Cost-advantage Supply Chain*. Boston, MA: Aberdeen-Group.

Afrifa, G. A., & Padachi, K. (2016). Working capital level influence on SME profitability. *Journal of Small Business and Enterprise Development, 23*(1), 44-63.

Altaf, N., & Shah, F. (2017). Working capital management, firm performance and financial constraints. *Asia-Pacific Journal of Business Administration, 9*(3), 206-219.

Baños-Caballero, S., García-Teruel, P. J., & Martínez-Solano, P. (2014). Working capital management, corporate performance, and financial constraints. *Journal of Business Research, 67*(3), 332–338.

Blinder, A. S., & Maccini, L. J. (1991). The resurgence of inventory research: what have we learned? *Journal of Economic Surveys, 5*(4), 291-328.

Blois, K. J. (1972). Vertical quasi-integration. *The Journal of Industrial Economics, 20*(3) 253-272.

Biais, B., & Gollier, C. (1997). Trade credit and credit rationing. *The Review of Financial Studies, 10*(4), 903-937.

Deloof, M., & Jegers, M. (1996). Trade credit, product quality, and intragroup trade: some European evidence. *Financial management, 25*(3), 33-43.

Deloof, M. (2003). Does Working Capital Management Affects Profitability of Belgian Firms? *Journal of Business Finance and Accounting, 30*(3), 573-587.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics, 49*(1), 92-107.

Blazenko, G. W., & Vandezande, K. (2003). The product differentiation hypothesis for corporate trade credit. *Managerial and Decision Economics, 24*(6/7), 457-469.

Grosse-Ruyken, P. T., Wagner, S. M., & Jönke, R. (2011). What is the right cash conversion cycle for your supply chain? *International Journal of Services and Operations Management, 10*(1), 13-29.

Hofmann, E., & Kotzab, H. (2010). A supply chain-oriented approach of working capital management. *Journal of Business Logistics, 31*(2), 305-330.

Jose, M. L., Lancaster, C., & Stevens, J. L. (1996). Corporate returns and cash conversion cycles. *Journal of Economics and Finance, 20*(1), 33-46.

Kieschnick, R., Laplante, M., & Moussawi, R. (2013). Working capital management and shareholders' wealth. *Review of Finance, 17*(5), 1827-1852.

Lanier, D., Wempe, W. F., & Zacharia, Z. G. (2010). Concentrated supply chain membership and financial performance: Chain-and firm-level perspectives. *Journal of Operations Management, 28*(1), 1-16.

Littler, S. (2015). The Importance and Effect of Sample Size. Retrieved from https://select-statistics.co.uk/blog/importance-effect-sample-size/

Metha, B.K. (2016). *Cost and Management Accounting*: Latest Edition. Agra, India: SBPD Publications.

Mun, S. G., & Jang, S. S. (2015). Working capital, cash holding, and profitability of restaurant firms. *International Journal of Hospitality Management*, 48, 1-11.

Pais, M. A., & Gama, P. M. (2015). Working capital management and SMEs profitability: Portuguese evidence. *International Journal of Managerial Finance, 11*(3), 341-358.

Petersen, M. A., & Rajan, R. G. (1997). Trade credit: theories and evidence. *The review of financial studies, 10*(3), 661-691.

Pfohl, H. C., & Gomm, M. (2009). Supply chain finance: optimizing financial flows in supply chains. *Logistics research, 1*(3-4), 149-161.

Porter, M. E. (1985). *Competitive advantage: creating and sustaining superior performance*. New York: FreePress

Protopappa-Sieke, M., & Seifert, R. W. (2017). Benefits of working capital sharing in supply chains. *Journal of the Operational Research Society, 68*(5), 521-532.

Rafuse, M. E. (1996). Working capital management: An urgent need to refocus. *Management Decision, 34*(2), 59-63.

Randall, W. S., & Theodore Farris, M. (2009). Supply chain financing: using cash-to-cash variables to strengthen the supply chain. *International Journal of Physical Distribution & Logistics Management, 39*(8), 669-689.

Shin, H. H., & Soenen, L. (1998). Efficiency of working capital management and corporate profitability. *Financial Practice and Education, 8*(2), 37-45.

Shipley, D., & Davies, L. (1991). The role and burden-allocation of credit in distribution channels. *Journal of Marketing Channels, 1*(1), 3-22.

Singhania, M., & Mehta, P. (2017). Working capital management and firms' profitability: evidence from emerging Asian countries. *South Asian Journal of Business Studies, 6*(1), 80-97.

Tauringana, V., & Afrifa, G. A. (2013). The relative importance of working capital management and its components to SMEs' profitability. *Journal of Small Business and Enterprise Development, 20*(3), 453-469.

Vazquez, X. H., Sartal, A., & Lozano-Lozano, L. M. (2016). Watch the working capital of tier-two suppliers: a financial perspective of supply chain collaboration in the automotive industry. *Supply Chain Management: An International Journal, 21*(3), 321-333.

Viskari, S. & Kärri, T. (2012). A model for working capital management in the interorganisational context. *International Journal Integrated Supply Management, 7*(1-3), 61-79.

Wang, Y. J. (2002). Liquidity management, operating performance, and corporate value: evidence from Japan and Taiwan. *Journal of Multinational Financial Management, 12*(2), 159-169.

Warren, C., & Reeve, J. (2006). *Financial & managerial accounting*. Boston MA: Cengage Learning.

Wilcox, R. R. (2001). *Fundamentals of modern statistical methods*. New Haven CT: Springer.

DECISION SCIENCES INSTITUTE

Asymmetry and Leverage Effect between US Institutional Investors' Funding Liquidity and Index Return - Using Data of 25 ETF Index and 22 Future Index

Junmao Chiu
Yuan Ze University
Email: jchiu@saturn.yzu.edu.tw
Le Van Hung
Yuan Ze University
Email: lehungtcnh@gmail.com
Kunchi Tsai
National Chiao Tung University
ken.iof99g@nctu.edu.tw

ABSTRACT

This study investigates the effect of U.S. institutional investor's funding liquidity on various nations' securities returns using data from 25 countries in the exchange-traded funds market (ETFs) and 22 countries in the financial future markets. Our results show that when U.S. investors face insufficient liquidity issues, leading to decrease on stock returns in other countries. We also find evidence that US government monetary policies have significant impact to money market and funding market. In sum, our study provides a better overall understanding of the effect of the liquidity–supplier funding constraint on returns.

KEYWORDS: Funding liquidity, Daily returns, 25 EFTs index, 22 Future index, Monetary policies

INTRODUCTION

The issue of funding has received considerable attention from numerous recently scholars. As a fundamental concept in finance, liquidity can be defined as the ability to buy or sell large quantities of an asset quickly and at low cost. Liquidity is risky and has commonalty: it changes over time not only individual stocks, but also whole market (Chordia et al., 2000; Hasbrouck & Seppi, 2001; Huberman & Halka, 2001). Liquidity plays an important role in financial markets, especially during a financial crisis. The continuous arrival of bad or uncertainty news within the market can cause in redemption pressure from retail investors on liquidity suppliers (e.g., intermediaries, speculators, and arbitrageurs). They therefore may be faced with funding constraints the risk of higher margins. Albert S Kyle and Xiong (2001), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2008) all argued that market declines and adverse shocks can result in funding constraints, leading fragile equity liquidity and a liquidity spiral effect. Liquidity control is one of the cores of banking industry. The Basel Accords therefore requires banks or large institutional investors maintain a certain level of funding liquidity per day to meet the capital adequacy ratio (CAR). Under typical circumstances, investors may apply for loans from the interbank market to adjust their funding liquidity. However, when a negative shock occurs, market uncertainty increases, banks will reduce their level of lending willingness, leading to the lending channels are more tightened, causing loan interest rates by interbank markets is significantly increased (Comerton-Forde et al., 2010). At this time, borrowing from banks becomes more difficult, large institutional investors who need funding must sell their liquid assets to maintain required necessary CAR ratio, this action is called pull back liquidity.

A significant example of the liquidity pull-back phenomenon is the 2008 financial crisis, the impact of subprime crisis on margins has led to the instability in the financial system, impacted many countries around the world. Brunnermeier (2009) argued that liquidity has commonality across securities because shocks to funding liquidity impact all securities in which speculators are marginal investors. Moreover, the US dollar has high liquidity in the international market, transactions by US dollars are the most often and highest frequency in the banking system. Therefore, the withdrawal of liquidity occurs not only in the US, but also in other countries. As large institution investors implement that behavior, stock returns decrease, a connection between interbank market and stock market is created (Nyborg & Östberg, 2014).
This paper therefore explores whether US institutional investors' funding liquidity affects the securities returns of other countries. Based on above literature, we hypothesize that a higher funding liquidity leads an increase in returns.
Our study adds several findings to the extant literature on the ways in which funding liquidity affects returns. First, past studies have explored how liquidity impacts financial market prices (Amihud & Mendelson, 1986; Jacoby et al., 2000), the interaction of liquidity and returns in stock markets(Hasbrouck, 1991), time-varying liquidity in Treasury bond markets (Krishnamurthy, 2002), liquidity predicts expected returns in the time series (Pástor & Stambaugh, 2003). To our knowledge, there is little empirical evidence emphasize the role of funding liquidity. Gromb and Vayanos (2010) who undertake explore the effect of funding liquidity and market liquidity, Hameed et al. (2010)find an indirectly effect of funding liquidity on equity liquidity during a period of market decline, Chiu et al. (2012) used index and financial exchange-traded funds (ETFs) to explored the relation between funding liquidity and equity liquidity during the subprime crisis period. We argue that liquidity shocks, market uncertainty, the announced of bad news forced the speculators to reduce their leverage and provide less liquidity to the markets, reduces asset liquidity (Brunnermeier, 2009). An important implication of their study is liquidity has commonality across securities because funding liquidity shock impact all securities in which speculators are marginal investors. However, the interaction between funding liquidity and in stock markets has not been directly deeply examined.
Second, prior studies have analyzed the effects of monetary policy and fund flows on financial markets. Monetary shocks are associated with large changes in bond and stock prices (J. Fleming & Remolona, 1997), fund flows affect price formation in equity markets (Goetzmann & Massa, 2002). The impact of monetary policy to liquidity emphasize in recent literature. Albert S. Kyle (1985), Admati and Pfleiderer (2015) argue from the speculation paradigm that strong find that market-wide changes in liquidity could closely related informational events such as Federal announcements about the state of the economy, Garcia (1989) examine the ways in which Fed's monetary stance impact liquidity by adjusting the margin lending terms and reducing dealers' borrowing restrictions; J. Fleming and Remolona (1997) and Fair (2002) argue that monetary shocks are associated with large changes in bond and stock prices. However, the impact of funding liquidity on returns in stock markets in the periods of monetary policies implementation has not been studied much. Past studies mainly focused on the tightening monetary policies; we focus on the quantitative easing methods. We argue that government applied the bailouts or funding injections have a direct impact on banking system and mutual funds (Lin et al., 2014), increases the bank's interest margin and decreases the default risk (Chen & Chang, 2015).
Third, Goldsmith-Pinkham and Yorulmazer (2010) describe the spillover effect under the bank run and the bailout announcement. The subprime crisis occurred, investors' willingness to enter the market decrease. The notice of bankruptcy or bailout of an individual bank will effect on the rest of the banking system. Nyborg and Östberg (2014) describe the phenomenon about "liquidity pull-back" which involves selling financial assets by banks directly or by levered investors in the interbank market. The typical example of "liquidity pull-back phenomenon" is the 2008 subprime crisis has led to the instability in the financial system. The 2008 funding crisis

emphasize the role of the liquidity spirals affect all interbank market, causing central banks around the world have to intervene in the monetary market at a level never seen. Many countries around the world are affected during the crisis. Laeven and Tong (2012) explore how US monetary policy affects to overseas stock market outside U.S. They discover that monetary policy has a negative relationship with the other countries' stock price. They also discover directly the impact of U.S monetary tighten or loosen to other stocks. We extend our studies how funding liquidity effects on returns in the international market. This paper also examines whether funding liquidity affects returns at varying degrees when the US market is tightening or loosening.

We use ETFs and future market data for the following reasons. First, both markets are usually more liquid and therefore more suited to our research question. Second, we also focus on the ETFs and future market because the financial industry is the sector most directly affected by the subprime crisis. Third, ETFs data is as known represent for the equity market, while future data of the capital market, we explore whether the impact of funding on these two markets is in different levels.

Our empirical findings are summarized as follows. First, our results show that higher funding liquidity leads to an increase in returns, which indicates that an increase in funding liquidity can improve returns. However, these affects to various countries are in different levels. Second, we find that with a decline in funding liquidity, investors tend to prioritize withdraws money at the higher liquidity countries first.

Third, our results generally reveal that funding liquidity is higher significant impact to returns when the market is loosening. Finally, we find that when US government implemented the monetary policies will have the effects to monetary and capital markets.

The remainder of this paper is organized as follows. Section 2 provides a description of the data and the research methodology. Section 3 presents and analyzes our empirical results. Finally, Section 4 offers the conclusions drawn from this study.

DATA SOURCE AND RESEARCH METHODS

Data source and sample selection

We use financial ETFs market from 25 countries and Global Future market from 22 countries in the period from January 1, 2004 to December 31, 2015 to explore the relationship funding liquidity and daily returns [List countries in financial ETFs market and global Future market is shown in Appendix A]. We collect the daily transaction data for individual stocks, daily trading volume, and daily trading volatility from DataStream database. Because GDP are macro-economic data with quarterly frequency, that is complicated to calculate. We consider that the information quantity in quarterly frequency will reduce the credibility and affect the explanatory ability. Therefore, we follow Allen et al. (2012) to adjust the frequency to monthly data information to avoid the insufficient number of sample situation.

Funding liquidity measures

We follow Brunnermeier (2009) to construct our funding liquidity measure. We use the spread between the 3-month US Treasury bills and Eurodollar Libor rate (i.e., Ted ratio). US Treasury bills are considered risk-free interest rates, while Libor rate reflects the British bank lending rates, implied risk premium. The Ted spread measure the degree of attention for the credit risk and dollar liquidity. We argue that with an increase in the Ted spread, leading to expanded market risk premiums, capital markets and interbank borrowing costs are tightened, leading to increase company's borrowing costs, therefore U.S. investors face the liquidity insufficiency situation, reducing invest to the securities markets, impact on stock market returns.

Measure daily returns

We use daily returns in the individual countries, which calculated as *ln* (P_d /P_{d-1})*;* where P_d is the stock price at the day d, P_{d-1} is the stock price at the prior day. To control for the factor that may be important in determining daily returns, we adopt daily trading volume (Duffee, 1992; Gallant et al., 1992; Karpoff, 1987), daily trading volatility (French et al., 1987; Schwert, 1989), returns in last period (Chan et al., 1996; Fama, 1990; Lee & Swaminathan, 2000), GDP (Binswanger, 2000; Schwert, 1989) to investigate the following regression model:

$$Ret_{i,d} = \alpha + \beta_{i,d} Ted_d + \gamma X_{i,d} + \varepsilon_{i,d}$$

Where $Ret_{i,d}$ is the daily return for ETFs and Future market for country i on day d. Ted_d is the Ted spread in day d. $X_{i,d}$ is control variable, consists of daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). Ted_d is the spread between the 3-month US Treasury bills and Eurodollar Libor rate, $Vol_{i,d}$ which is collect from DataStream database, $Volatility_{i,d}$ is calculated form past 30th $Ret_{i,d}$, Ret_{d-1} is measure by lagged $Ret_{i,d}$, , $GDP_{i,m}$ measured by *ln*($GDP_{i,m}/GDP_{i,m-1}$).

Basic Statistic

Table 1 reports the descriptive statistic for daily returns, funding liquidity and the other control variables. The number observations are 98,793 and 56,339 in financial ETFs and Future market, respectively. The min of *Ret* are -0.340 and -0.162 with a standard deviation of 0.019 and 0.015, respectively. The min of *Ted* is -0.309 with a standard deviation of 1.194 and 1.309 in both markets, respectively. ETFs market has the min value of *Vol*, *Volatility*, Ret_{d-1} are 100, 0.001, -0.340, respectively; while the min value of *Vol*, *Volatility*, Ret_{d-1} are 0.001, 0.001, -1.378, respectively in the future market.

Table 1. Descriptive statistic

Variables	Obs.	Min	Max	Average	S.D.
Panel A: Global ETF data					
Ret	98793	-0.340	0.256	0.001	0.019
Ted	98793	-0.309	5.973	1.419	1.194
Vol	98793	100.000	396583900.000	2555082.000	6956881.000
Volatility	98793	0.001	0.011	0.001	0.001
Ret_{d-1}	98793	-0.340	0.256	0.001	0.019
Turn	98793	0.001	1000.333	3.328	7.430
GDP	98793	-7.359	15.240	2.670	2.945
Panel B: Global Future data					
Ret	56339	-0.162	0.237	0.001	0.015
Ted	56339	-0.309	5.973	1.345	1.309
Vol	56339	0.001	828528.000	61060.000	80637.000
Volatility	56339	0.001	0.064	0.001	0.001
Ret_{d-1}	56339	-1.378	0.237	0.001	0.016
GDP	56339	1.762	4.488	2.891	0.796

Notes: This table provides the descriptive statistics for the dependent and control variables, with the data covering the period of 1 January 2004 to December 31, 2015.

EMPIRICAL RESULTS

Main results

This section explores how the funding liquidity affects daily returns using the following regression model:

$$Ret_{i,d} = \alpha + \beta_{i,d}Ted_d + \gamma X_{i,d} + \varepsilon_{i,d}$$

Where the dependent variable $Ret_{i,d}$ is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns *(Ret_{d-1})* and growth rate ($GDP_{i,m}$).

We separated the results using data from 25 different countries in financial ETFs market in Panel A and 22 different countries in global future market in Panel B of Table 2. Our findings show that the coefficients of *Ted* are significantly negative at -0.001, with significant at the 1% level for both equity market and future market sample. These results suggest that when the large U.S. institutional investor lack of funding liquidity leading to a decrease on stock returns in other countries.

Vol has a significant negative impact on *Ret,* ranging from -0.001 to -0.003, significant at the 1% level in ETFs market and future market, respectively. These results posit that a higher on daily trading volume can lead to decline on daily returns (Campbell et al., 1993; Hutson et al., 2008). Table 2 shows that the coefficient of *Volatility* is significantly negative at -0.071, with significant at the 1% level in the Future market (Campbell & Hentschel, 1992; Duffee, 1995). While, *Volatility* has a significant positive impact on *Ret* at 0.763, with significant at the 1% level in financial ETFs market (as shown in Panel A), this results are line with (Duffee, 1995; GRULLON et al., 2012; Veronesi, 1999), it has important theoretical implications because it is inversely related with "leverage effect" hypothesis (Christie, 1982; Fischer, 1976). These empirical results imply that higher daily trading volatility can lead to an increase (decrease) in securities returns. Our results also shows that Ret_{d-1} has a negative significant impact on *Ret*, with the coefficients of Ret_{d-1} range from -0.075 to -0.117, significant at the 1% levels in both ETFs and Future markets. In addition, the coefficients *GDP* have insignificant positive impact on *Ret* in both markets.

We also analyze the effects of funding liquidity on index returns in each individual country. The coefficients of *Ted* in 14 countries in Panel A and 17 countries in Panel B are significant negative, with significant at least the 10% level. The most impact of funding liquidity to daily returns clearly in South Korea, Thailand and Turkey on ETFs market (as shown in Panel A), and Germany, Hong Kong, India, Italy, Malaysia, Netherlands, Singapore, Switzerland, Taiwan on future market (as shown in Panel B) at the 1% significant level. While, the coefficients of Ted are negative significant at -0.001, with significant at the 10% levels in China, Switzerland (as shown in Panel A); and Brazil, South Korea, Spain and the United Kingdom (as shown in Panel B). Those results imply that when large institutional investors have insufficient liquidity, index returns will be affected, lead to a decrease returns will occur. In addition, our results posit that the effects of funding liquidity on index returns in different countries are in different levels. The majority of control variables also show significant coefficients that are mostly consistent with our conjectures and prior literatures. These results show that the control variables used in this study are appropriate and play an important role in our model.

Table 2. Funding liquidity and daily returns

Panel A: Global ETF data													
Variables	Ret_{AUS}	Ret_{AUT}	Ret_{BEL}	Ret_{BRA}	Ret_{CAN}	Ret_{CHN}	Ret_{FRA}	Ret_{DEU}	Ret_{HKG}	Ret_{IRL}	Ret_{ITA}	Ret_{JPN}	Ret_{MYS}
Ted	-0.001 (-1.59)	-0.001 (-0.94)	-0.001 (-0.83)	-0.001 (-0.86)	-0.001^{**} (-2.18)	-0.001^{*} (-1.67)	-0.001 (-1.35)	-0.001^{**} (-2.19)	-0.001^{**} (-2.49)	-0.005 (-1.58)	-0.001 (1.32)	-0.001^{**} (-2.53)	-0.001^{**} (-2.00)
Vol	-0.001 (-3.00)	-0.003 (-1.48)	-0.001 (-1.40)	-0.001 (-1.50)	-0.001^{***} (-3.10)	-0.001 (-0.05)	-0.002 (-1.60)	-0.001^{***} (-3.07)	-0.001^{***} (-2.85)	-0.009 (-1.03)	-0.001 (-1.44)	-0.001^{***} (-4.33)	-0.001^{**} (-2.40)
Volatility	0.518 (0.98)	-0.846 (-1.54)	-2.057^{***} (-3.61)	0.589 (1.22)	0.172 (0.23)	1.258^{**} (1.98)	0.440 (0.66)	0.278 (0.46)	1.646^{***} (3.10)	5.882 (1.22)	0.526 (0.77)	1.642^{**} (2.23)	1.547^{***} (2.68)
$\mathrm{Ret}_{\mathrm{d-1}}$	-0.127 (-8.75)	-0.057^{***} (-3.84)	-0.054^{***} (-3.65)	-0.014 (-0.82)	-0.037^{**} (-2.53)	-0.154^{***} (-7.74)	-0.072^{***} (-4.94)	-0.061^{***} (-4.23)	-0.134^{***} (-9.28)	-0.099^{***} (-3.22)	-0.075^{***} (-5.03)	-0.103^{***} (-7.09)	-0.043^{***} (-2.94)
GDP	0.001 (0.43)	-0.001 (-0.04)	0.001 (0.83)	-0.001 (-0.76)	-0.001^{***} (-3.32)	0.001 (0.04)	0.001 (1.03)	0.001 (2.28)	0.001^{**} (2.41)	0.001 (0.06)	-0.001 (-0.54)	0.001^{***} (4.22)	0.001^{***} (4.21)
Intercept	0.001^{**} (2.08)	0.002^{***} (3.24)	0.002^{***} (3.14)	0.002^{***} (2.40)	0.003^{***} (4.25)	-0.002 (-0.77)	0.001 (1.41)	0.001^{**} (2.04)	0.001 (0.81)	0.004^{**} (2.13)	0.001 (1.62)	0.001 (0.63)	-0.001 (-1.36)
Variables	Ret_{MEX}	Ret_{NLD}	Ret_{NZL}	Ret_{SGP}	Ret_{KOR}	Ret_{ESP}	Ret_{SWE}	Ret_{CHE}	Ret_{TWN}	Ret_{THA}	Ret_{TUR}	Ret_{GBR}	Ret_{TOTAL}
Ted	-0.001^{**} (-2.41)	-0.001^{**} (-2.00)	-0.001 (-0.38)	-0.001^{**} (-2.17)	-0.001^{***} (-2.71)	-0.001 (-0.63)	-0.001^{*} (-1.65)	-0.001 (-1.53)	-0.001 (-0.86)	-0.003^{***} (-4.23)	-0.004^{***} (-4.99)	-0.001^{**} (-2.07)	-0.001^{***} (-7.09)
Vol	-0.001^{***} (-3.59)	-0.005^{*} (-1.94)	-0.034 (-1.42)	-0.001^{**} (-2.15)	-0.001^{**} (-2.32)	-0.001 (-0.86)	-0.003^{*} (-1.75)	-0.001 (-0.93)	0.001 (0.18)	-0.012^{***} (-2.90)	-0.022^{***} (-5.72)	-0.001^{***} (-2.91)	-0.001^{***} (-3.55)
Volatility	1.224^{**} (2.26)	0.179 (0.26)	-0.464 (-0.15)	1.992^{***} (3.20)	1.244^{***} (2.88)	0.835 (1.23)	0.958^{*} (1.75)	0.279 (0.33)	0.306 (0.40)	1.942^{*} (1.90)	2.159^{***} (2.74)	0.117 (0.19)	0.763^{***} (6.91)
$\mathrm{Ret}_{\mathrm{d-1}}$	0.001 (0.04)	-0.084^{***} (-5.75)	-0.135^{***} (-4.88)	-0.108^{***} (-7.48)	-0.079^{***} (-4.30)	-0.060^{***} (-4.10)	-0.076^{***} (-5.22)	-0.114^{***} (-7.84)	-0.098^{***} (-6.13)	-0.132^{***} (-5.81)	-0.062^{***} (-2.73)	-0.120^{***} (-8.28)	-0.075^{***} (-23.60)
GDP	0.001^{***} (2.57)	0.001 (1.32)	0.001 (1.20)	0.001^{**} (2.15)	0.001 (0.95)	0.001^{**} (2.31)	0.001 (1.28)	0.001 (0.03)	-0.001 (-1.14)	-0.001 (-0.04)	0.001 (3.24)	0.001^{**} (2.06)	0.001 (-1.20)
Intercept	0.002^{**} (2.10)	0.001^{**} (2.51)	0.002 (0.98)	-0.001 (-0.10)	0.003^{**} (2.29)	-0.001 (-0.21)	0.001 (1.27)	0.001^{*} (1.92)	0.001 (0.82)	0.004^{***} (3.85)	0.006^{***} (4.73)	0.001^{**} (2.15)	0.001^{***} (4.87)
Panel B: Global Future data													
Variables	Ret_{AUS}	Ret_{BRA}	Ret_{CAN}	Ret_{DNK}	Ret_{FRA}	Ret_{DEU}	Ret_{GRC}	Ret_{HKG}	Ret_{HUN}	Ret_{IND}	Ret_{ITA}	Ret_{JPN}	Ret_{MYS}
Ted	-0.001 (-1.16)	-0.001^{*} (-1.88)	-0.001 (-1.34)	-0.001 (-1.01)	-0.001^{**} (-2.20)	-0.001^{***} (-2.85)	-0.001 (-1.48)	-0.001^{***} (-2.79)	-0.001 (-1.24)	-0.001^{***} (-2.94)	-0.001^{***} (-3.06)	-0.001^{**} (-1.99)	-0.001^{***} (-4.05)
Vol	-0.008 (-1.40)0	-0.028^{***} (-2.98)	-0.028^{**} (-2.05)	-0.088 (-1.56)	-0.014^{***} (-5.50)	-0.035^{***} (-8.70)	-0.076 (-1.25)	-0.021^{**} (-2.25)	-0.118^{***} (-2.88)	-0.010 (-1.49)	-0.107^{***} (-4.88)	-0.072^{***} (-8.50)	-0.096^{***} (-3.19)
Volatility	-2.895^{*} (-1.78)	1.106 (0.97)	-0.874 (-0.94)	-2.668^{**} (-2.20)	0.714 (0.66)	3.011^{***} (2.85)	-2.449^{***} (-2.80)	1.917^{**} (1.98)	-0.148 (-0.16)	2.044^{*} (1.82)	1.782 (1.62)	1.601^{**} (2.27)	6.996^{***} (3.14)
$\mathrm{Ret}_{\mathrm{d-1}}$	-0.029 (-1.51)	-0.016 (-0.80)	-0.092^{***} (-4.78)	0.050^{**} (2.19)	-0.070^{***} (-3.66)	-0.028 (-1.44)	0.020 (1.01)	-0.029 (-1.48)	0.065^{***} (3.37)	-0.025 (-1.21)	-0.003 (-0.15)	-0.099^{***} (-5.06)	-0.077^{***} (-3.96)
GDP	0.001	0.001	-0.001	-0.001	0.001	0.001^{**}	0.001	0.001	0.001	0.001	0.001	-0.001	0.001

	(0.19)	(1.16)	(-0.13)	(-0.28)	(0.32)	(2.08)	(0.79)	(1.26)	(0.72)	(0.91)	(0.28)	(-0.28)	(0.93)
Intercept	0.001 (1.01)	0.001 (0.69)	0.001 (1.30)	0.002 (1.32)	0.002** (2.05)	0.004*** (3.35)	0.001 (0.41)	0.001 (0.51)	0.001 (0.40)	0.001 (0.23)	0.002* (1.75)	0.007*** (4.71)	0.001 (1.10)
Variables	Ret_{MEX}	Ret_{NLD}	Ret_{SGN}	Ret_{KOR}	Ret_{ESP}	Ret_{SWE}	Ret_{CHE}	Ret_{TWN}	Ret_{GBR}	Ret_{TOTAL}			
Ted	-0.001** (-2.32)	-0.001*** (-3.10)	-0.001*** (-3.15)	-0.001* (-1.72)	-0.001* (-1.65)	-0.001** (-2.15)	-0.001*** (-2.58)	-0.001*** (-2.99)	-0.001* (-1.72)	-0.001*** (-3.03)			
Vol	0.029 (0.69)	-0.054*** (-5.81)	0.006 (0.39)	-0.028*** (-8.05)	-0.016 (-1.19)	-0.014* (-1.71)	-0.006** (-2.38)	-0.048*** (-5.73)	-0.013*** (-4.79)	-0.003*** (-3.94)			
Volatility	2.301** (1.98)	0.060 (0.07)	1.175 (0.92)	3.546*** (3.08)	0.544 (0.51)	0.001 (0.01)	1.668 (1.26)	1.627 (1.07)	1.355 (1.26)	-0.071** (2.02)			
$\mathrm{Ret}_{\mathrm{d}-1}$	0.066*** (3.42)	-0.041** (-2.16)	-0.043** (-2.24)	-0.058*** (-2.95)	0.010 (0.51)	-0.004 (-0.36)	-0.088*** (-4.33)	-0.024 (-1.19)	-0.057*** (-2.99)	-0.117*** (-31.78)			
GDP	0.001 (1.30)	0.001 (0.47)	0.001 (1.23)	-0.001 (-1.16)	0.001 (0.45)	0.001 (0.19)	-0.001 (-0.54)	-0.001** (-1.98)	-0.001 (-0.13)	0.001 (0.88)			
Intercept	-0.001 (0.67)	0.003** (2.52)	0.111 (0.01)	0.008*** (5.24)	0.001 (0.33)	0.001 (0.90)	0.002* (1.68)	0.007 (4.52)	0.002 (2.13)	0.001 (0.37)			

Notes: This table provides details of the funding liquidity effect on the daily returns. The regression model is as follows:
$Ret_{i,d} = \alpha + \beta_{i,d} Ted_d + \gamma X_{i,d} + \varepsilon_{i,d}$
Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use the panel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Classification Discussion

This section explored whether funding liquidity effect on daily returns in different groups of countries is in different level. To reduce the characteristics of individual countries, we used five different classified methods to divide countries into groups; we then use panel regression to analysis which group is deeply impacted when large-scale U.S. institutional investors face the liquidity insufficiency situation.

<u>Classify by Volume</u>

Information flow is the main cause of price or return volatility, however the reality indicates that information flow cannot be observed. Previous literatures showed that trading volume responds to information flowing (Karpoff, 1987), can be considered as a transmission information in the financial market (Lamoureux & Lastrapes, 1990a, 1990b). We consider that when U.S. investors face insufficient liquidity issue, they will inevitably consider the market liquidity to facilitate for withdrawing their funds. We next classified countries into three groups: large, medium and small trading volume countries [More detail classified countries by trading volume as shown in Panel A of Appendix B]. Table 3 presents separate results using data from financial ETFs market in Panel A, and Global Future market in Panel B.
Our results show that the coefficients of *Ted* are significantly negative at -0.001, with significant at the 1% level for all groups both in Panel A and Panel of Table 3. These results suggested that when the liquidity of U.S. investors is insufficient can lead to decrease daily returns in our all group countries. In addition, *Ted* has a deeply significantly negative impact on *Ret* in the medium and trading volume countries in both equity and future market. These results are consistent with our conjectures; it implies that when the U.S. investors face insufficient liquidity issues, they likely to withdraw their funds from better liquidity markets.

Table 3. The panel regression results which classified by volume

Variables	Small groups	Medium groups	Large groups
Panel A: Global ETF data			
Ted	-0.001*** (3.98)	-0.001*** (-4.71)	-0.001*** (-4.35)
Vol	-0.002*** (-3.74)	-0.001*** (-5.41)	-0.001*** (-3.06)
Volatility	-0.137 (-0.56)	1.005*** (4.85)	0.884*** (5.38)
Ret_{d-1}	-0.080*** (-13.22)	-0.069*** (-12.92)	-0.078*** (-14.95)
GDP	0.001 (0.81)	0.001*** (2.86)	-0.001** (-1.98)
Intercept	0.001*** (4.31)	0.001 (0.36)	0.001*** (3.61)
Panel B: Global Future data			
Ted	-0.001*** (-4.10)	-0.001*** (-5.91)	-0.001*** (-5.61)
Vol	-0.013 (-1.58)	-0.018*** (-6.21)	-0.017*** (-13.58)
Volatility	-1.309*** (-3.50)	0.008 (0.17)	1.115*** (2.94)
Ret_{d-1}	0.001 (0.20)	-0.011* (-1.80)	-0.056*** (-7.62)
GDP	0.001 (1.21)	0.001 (1.47)	-0.001 (-0.02)
Intercept	0.001 (0.86)	0.001 (0.34)	0.006** (2.42)

Notes: This table provides details of the funding liquidity effect on the daily returns in small, medium, and high daily trading volume countries in the period from 1 January 2004 to December 31, 2015.Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use the panel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Classify by Economic Freedom

Economic freedom reflects the nation's long-term economic growth and prosperity. To measure economic freedom, Wall Street Journal and the American Heritage Foundation established indexes to measure the economic freedom level, covers 179 countries and regions around the world based on 50 indicators. The score ranges from 0 to 100. The higher the score of an indicator, the lower the government intervention in the economic and the degree of economic freedom is higher. We then classify countries into different group based on level of economic freedom and use the panel the regression to exam the relationship between funding liquidity and daily returns in individual group [More detail classified countries by Economic Freedom as shown in Panel B of Appendix B]. Table 4 presents separate results using data from equity market in Panel A, and future market in Panel B.

Table 4 shows that the coefficient of *Ted* is significantly negative at -0.001, with significant at the 5% level in the group of countries has almost no economic freedom (as results shown in Panel B of Table 4). In addition, the coefficients of *Ted* are significant negative at -0.001, with significant at the 1% level for other three groups of countries in both ETFs and Future market. These results suggest that when the liquidity of U.S. investors is insufficient, investors will withdraw funds from the countries with higher economic freedom level, lead to negative effect on their countries returns.

Table 4. The panel regression results which classified by economic freedom index

Variables	Mostly Unfree	Moderately Free	Mostly Free	Free
Panel A: Global ETF data				
Ted	-0.001 (-1.09)	-0.001*** (-4.52)	-0.001*** (-4.29)	-0.001*** (-4.24)
Vol	-0.001 (-0.29)	-0.001*** (-4.48)	-0.001* (-1.91)	-0.001*** (-4.52)
Volatility	0.829** (2.23)	0.747*** (3.59)	0.522*** (2.78)	1.138*** (4.43)
$\mathrm{Ret}_{\mathrm{d}-1}$	-0.065*** (-5.39)	-0.051*** (-9.14)	-0.081*** (-15.47)	-0.108*** (-17.13)
GDP	-0.001* (-1.73)	0.001*** (2.94)	-0.001 (-1.37)	0.001 (0.99)
Intercept	0.001 (0.85)	0.001*** (2.67)	0.001*** (3.75)	0.001 (0.95)
Panel B: Global Future data				
Ted	-0.001** (-2.31)	-0.001*** (-5.51)	-0.001*** (-5.48)	-0.001*** (-5.11)
Vol	-0.017*** (-2.93)	-0.013*** (-5.79)	-0.024*** (-13.74)	-0.007*** (-3.45)
Volatility	-0.900 (-1.63)	1.226*** (2.55)	-0.005 (-0.10)	0.660 (1.36)
$\mathrm{Ret}_{\mathrm{d}-1}$	0.001 (0.12)	-0.007 (-0.82)	-0.022*** (-3.83)	-0.055*** (-6.28)
GDP	0.001 (1.50)	0.001 (1.62)	-0.001 (-0.54)	0.001 (1.48)
Intercept	-0.007 (-1.16)	-0.001 (-0.40)	0.007 (1.62)	-0.001 (-0.10)

Notes: This table provides details of the funding liquidity effect on the daily returns in mostly unfree, moderately free, mostly free and free countries in the period from 1 January 2004 to December 31, 2015. Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use the panel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Classify by Volatility

Changes in the level of trading volatility can have important effects on capital investment, consumption, business cycle and returns (Schwert, 1989). In this section, we explore the impact of funding liquidity on returns on different volatility groups. We classified countries into different

groups consist of: group of countries which small volatility, medium volatility and large volatility [More detail classified countries by trading volatility as shown in Panel C of Appendix B].
Table 5 presents separate results using data from financial ETFs market on Panel A and Global Future market on Panel B. Our findings show that the coefficients of Ted are significantly negative, at -0.001, with significant at the 1% level for all groups of countries in both ETFs and Future market. In addition, our results show that Ted have more significant impact on large volatility group in Panel A; and small volatility and medium groups in Panel B of Table 5. These results imply that when the liquidity of U.S. investors is insufficient, all groups of countries which small, medium, large volatility are impacted by withdraw of U.S. investors, lead to a decrease on their daily securities returns.

Table 5. The panel regression results which classified by volatility

Variables	Small groups	Medium groups	Large groups
Panel A: Global ETF data			
Ted	-0.001*** (-3.62)	-0.001*** (-3.80)	-0.001*** (-5.05)
Vol	-0.001 (-1.50)	-0.001*** (-4.31)	-0.001*** (-2.84)
Volatility	0.419 (1.52)	0.424** (2.08)	1.090*** (6.20)
Ret_{d-1}	-0.087*** (-15.30)	-0.085*** (-16.51)	-0.060*** (-10.73)
GDP	-0.001 (-1.64)	0.001* (1.95)	-0.001 (1.25)
Intercept	0.001*** (2.67)	0.001*** (3.28)	0.001*** (2.78)
Panel B: Global Future data			
Ted	-0.001*** (-5.90)	-0.001*** (-5.95)	-0.001*** (-4.92)
Vol	-0.013*** (-5.95)	-0.021*** (-12.23)	-0.011*** (-5.02)
Volatility	0.394 (0.95)	0.895** (2.18)	-0.014 (-0.25)
Ret_{d-1}	-0.032*** (-4.36)	-0.015*** (-2.17)	-0.016** (-2.52)
GDP	0.001 (1.57)	0.001 (0.67)	0.001 (1.53)
Intercept	-0.001 (-0.58)	0.002 (0.53)	-0.001 (-0.25)

Notes: This table provides details of the funding liquidity effect on the daily returns in small, medium, and daily trading volatility countries in the period from 1 January 2004 to December 31, 2015.Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use thepanel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Classify by the change of Ted spread

Next, we classified our sample to three different groups based on the change of Ted spread to examine whether funding liquidity affects to daily returns in different groups is in different levels. We collect the change of Ted spread in a given day compared with prior day (ΔTed_d), we considered if ΔTed_d is positive means that the monetary market is tighter, $\Delta Ted_d$is negative means the monetary market is looser, and ΔTed_d equals zero means the monetary market is unchanged. Table 6 presents the asymmetric relationship between Ted and daily returns using data from equity market (in Panel A) and future market (in Panel B).
Our empirical results show that the coefficients of *Ted* are significantly negative at -0.001, with significant at the 1% level when the U.S. monetary market is looser for both Panel A and Panel B of Table 6. These results suggest that when U.S. investors have abundant funds, they will invest spill into the global market, driving the increase market returns. While, the coefficients of *Ted* have a negative insignificant impact on daily returns when the U.S. monetary market is tightening for both markets. These results posit that when U.S. funding market is tighter, U.S. investors face insufficient liquidity problem, they will prioritize recover the required funds from the domestic capital market first. To the international market, this process takes a long time to wait and sell, leading to solve the capital shortage problem and avoid the risk of liquidity gaps are ineffective, therefore its impact to returns cannot be reflected immediately. Panel A of Table 6 also shows the coefficients of *Ted* is at -0.001, with significant at the 1% when U.S. monetary market is unchanged. These empirical results are line with our expectation. They indicating that higher *Ted* spread can lead to a decrease in market returns.

Table 6. The panel regression results which classified by the change of Ted spread

Variables	$\Delta Ted_d > 0$	$\Delta Ted_d < 0$	$\Delta Ted_d = 0$
Panel A: Global ETF data			
Ted	-0.001 (-1.31)	-0.001*** (-8.40)	-0.001*** (-2.61)
Vol	-0.001*** (-2.69)	-0.001** (-2.45)	0.001 (0.35)
Volatility	0.971*** (5.41)	0.277* (1.93)	11.501*** (15.23)
Ret_{d-1}	-0.067*** (-14.27)	-0.077*** (-17.25)	-0.133*** (-6.60)
GDP	-0.001 (-0.27)	-0.001 (-1.22)	-0.001** (-2.52)
Intercept	0.001 (0.14)	0.001*** (6.82)	-0.001 (-0.63)
Panel B: Global Future data			
Ted	-0.001 (-0.50)	-0.001*** (-4.98)	0.034 (0.70)
Vol	-0.001 (-1.27)	-0.004*** (-3.87)	-0.033** (-2.13)
Volatility	-0.272*** (-4.26)	0.039 (0.92)	18.798*** (3.58)
Ret_{d-1}	-0.105*** (-21.02)	-0.136*** (-24.10)	-0.047 (-0.37)
GDP	0.001 (0.18)	0.001* (1.67)	-0.030 (-0.60)

Intercept	0.001 (0.16)	0.001 (0.12)	0.051 (0.59)

Notes: This table provides details of the funding liquidity effect on daily index returns by the change of TED spread in d day and d-1 day (ΔTed_d). We classify ΔTed_dinto three groups are bigger, smaller and equal to zero. Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns *(Ret_{d-1})*, and growth rate ($GDP_{i,m}$). In Panel A, we use thepanel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

As the results show in Panel A of Table 6, *Ted* also have strongly negative significant to market returns when U.S. monetary market is unchanged. We next deeply explore the relationship between *Ted* and *Ret* in this group. We continue collect the change of Ted spread in day d-1 and day d-2 (ΔTed_{d-1}). We classified three difference subgroups based on the unchanged of ΔTed and the change of ΔTed_{d-1}. Our three subgroups consist of unchanged ΔTed_d and positive ΔTed_{d-1}; unchanged $\Delta Ted_d$and negativeΔTed_{d-1}; unchanged ΔTed_d and unchanged ΔTed_{d-1} [The number observation in the subgroup: unchanged ΔTed_d and unchanged ΔTed_{d-1} is only three observations, so we ignored this subgroup]. Table 7 shows that the coefficient of Ted is significant negative at -0.002, with significant at the 1% in group unchanged $\Delta Ted_d$and negative ΔTed_{d-1}. It implies that when the U.S. funding market looser than previous period, it not only has positive impact to market returns in this period, but also leading to an increase returns in the next period. Our results also pointed out the coefficient of *Ted* does not have a significantly negative impact on *Ret* in subgroup unchanged$\Delta Ted_d$and positive ΔTed_{d-1}, this result is similar to our expectations and consistent with prior findings on Table 6. These results strongly imply the asymmetric relationship between U.S. tighter (looser) monetary market and market returns.

Table 7. The panel regression results which classified by the unchanged of ΔTed_d and the change of ΔTed_{d-1} for ETF

Variables	$\Delta Ted_d = 0$ and $\Delta Ted_{d-1} > 0$	$\Delta Ted_d = 0$ and $\Delta Ted_{d-1} < 0$	$\Delta Ted_d = 0$ and $\Delta Ted_{d-1} = 0$
Ted	-0.001 (-1.07)	-0.002*** (-3.29)	
Vol	-0.001*** (-3.79)	0.001** (2.14)	
Volatility	4.218** (2.52)	13.165*** (14.06)	
$\mathrm{Ret_{d-1}}$	-0.090*** (-2.71)	-0.138*** (-5.34)	
GDP	-0.001 (-1.57)	-0.001 (-0.89)	
Intercept	0.001 (0.30)	0.001 (0.07)	

Notes: This table provides details of the funding liquidity effect on the daily returns by the unchanged of TED spread in d day and d-1 day (ΔTed_d), and the change of TED spread in d-1 day and d-2 day (ΔTed_{d-1}). We also classify ΔTed_{d-1}into three groups are bigger, smaller and equal to zero. The dependent variable is the daily returns. Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily return *(Ret_{d-1})*, and growth rate

($GDP_{i,m}$). The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Classified by Quantitative Easing

The financial crisis 2008 was considered as the most serious financial crisis since the Great Depression of the 1930s, caused by the U.S. subprime mortgage market and developed into the banking crisis, impacted to the world financial system. Financial market faced the insufficient liquidity issues. Therefore, U.S. government performed the quantitative easing methods, injecting a large amount of money into the market to rescue the confidence of investors. In this section, we deeply explore how the funding liquidity effect on daily returns when U.S. government performed the quantitative easing methods. We classified our sample into four difference groups based on three times of quantitative easing by the US government, consisting of QE1 (first quantitative easing), QE2 (second quantitative easing), QE3 (third quantitative easing), NonQE (other non-quantitative easing period).
[QE1 (first quantitative easing): QE1 is implemented from March 2009 to March 2010, and the scale is around 1.725 trillion US dollars, an average of more than 100 billion US dollars per month. The funding used to buy 1.25 trillion US dollars of mortgage-backed securities (MBS), $300 billion in US Treasury bonds, and $175 billion in institutional securities (Agency MBS).
QE2 (Second quantitative easing): The implementation time is from the end of November 2010 to June 2011, and the scale is about 600 billion US dollars. It is mainly used to purchase long-term bonds issued by the Ministry of Finance. The average monthly purchase amount is 750. Billion.
QE3 (Third Quantitative easing): Implementation time is from September 2012, with Agency MBS as the target, September purchase of $23 billion, and monthly purchase of $400 trillion from October. Different from the previous two rounds, the QE3 has no clear implementation deadline and will continue to be implemented until the US job market recovers. In January 2013, QE3 was further intensity, in addition to continuing to implement monthly purchases of $400 trillion in MBS, and increasing monthly acquisitions of $450 trillion in long-term public debt, that will increase the monthly bond purchases to $850 trillion. Although QE3 still has no clear deadline, it has joined the exit trigger, which is begin to adopt a gradual reduction in the scale of debt purchase when the expected inflation rate in the United States is higher than 2.5% and the unemployment rate is lower than 6.5%.].
Table 8 presents separately results using data from ETFs market in Panel A and Global Future market in Panel B. Our empirical result shows that the coefficients of *Ted* are significant negative at -0.001, with significant at the 1% level in both equity market and future market when US government do not perform the quantitative easing policies. Those results suggest that when the US government does not implement monetary and fiscal policies, U.S. investors face the shortage liquidity problem; they will withdraw their fund from the capital market, resulting in a decrease returns on securities market. In addition, the coefficients of *Ted* do not have a significantly negative effect on *Ret* for all three times of quantitative easing by the U.S. government (as shown in both panel A and panel B of Table 8). These results are line with our expectation. It indicates that during implementing three times easing methods, US government monetary policies not only directly impact the liquidity of funding market but also indirectly impact the capital market.

Table 8. The panel regression results by the change of quantitative easing methods effect

Variables	QE1	QE2	QE3	NonQE
Panel A: Global ETF data				
Ted	-0.001 (-0.31)	-0.004 (-1.42)	0.004 (0.95)	-0.001*** (-4.82)
Vol	0.001 (1.27)	-0.001* (-1.89)	0.001 (0.49)	-0.001*** (-4.27)
Volatility	5.716*** (4.86)	10.648*** (3.80)	1.308 (1.05)	0.601*** (5.16)
Ret_{d-1}	-0.081*** (-6.13)	0.0216 (1.34)	-0.031*** (-2.74)	-0.080*** (-22.97)
GDP	-0.00 (-0.62)	-0.001** (-2.44)	-0.001** (2.46)	0.001 (-0.86)
Intercept	-0.001 (-0.12)	0.002 (0.87)	-0.001 (-0.46)	0.001 (1.29)
Panel B: Global Future data				
Ted	0.003 (1.45)	-0.011 (-1.10)	0.009 (0.76)	-0.001*** (-3.01)
Vol	-0.002 (-1.08)	-0.008*** (-3.20)	-0.003 (-1.60)	-0.002*** (-2.62)
Volatility	1.791* (1.85)	3.152 (1.48)	1.498* (1.82)	-0.080** (-2.19)
Ret_{d-1}	-0.157*** (-11.67)	-0.028* (-1.66)	-0.066*** (-5.36)	-0.121*** (-28.95)
GDP	-0.003 (-0.90)	0.001 (0.27)	0.001 (0.54)	0.001** (2.09)
Intercept	0.005 (0.71)	0.005 (0.57)	-0.005 (-0.94)	-0.001 (-1.01)

Notes: This table provides details of the funding liquidity effect on the daily returns on three quantitative easing periods of the U.S. government. QE1 is implemented from March 2009 to March 2010, QE2 implementation time is from the end of November 2010 to June 2011, QE3 implementation time is from September 2012, and NonQE presents other non-quantitative easing periods. Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$).The controlvariable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use the panel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Effect of quantitative easing methods and the change of Ted spread

In this section, we classified our sample into groups based on the change of Ted spread and three times of quantitative easing by the US government. We explore whether exist the asymmetrical relationship between funding liquidity and returns in quantitative easing period. Table 9 presents separately results using data from ETFs market in Panel A and Global Future market in Panel B.

Table 9 shows that *Ted* have a significant negative effect on *Ret* at the 5% level in the ETFs market and the 1% level in the future market when U.S. government do not implement quantitative easing and market is loosening. The coefficient of *Ted* is -0.001, with significant at the 1% level when market unchanged (as shown in Panel A). These results are line with our expectation. It indicates that funding liquidity effects on returns when US government do not implement monetary policies.

In the first quantitative easing period, the coefficient of *Ted* is negative significant, ranging from -0.003 to -0.055, with significant at the 5% levels when monetary market is unchanged and loosen, respectively (as shown in Panel A). In Panel B of Table 9, the coefficients of *Ted* are significantly positive at 0.007, with significant at the 5% level, when the monetary market is loosening. This result clearly demonstrates that quantitative easing affects the relationship between the money market and the capital market.

In the second quantitative easing period, *Ted* has a significant negative significant effect on *Ret* at the 1% level when the monetary market is tightened (as results shown in Panel A). It implies that although the funding market is tightened, U.S. investors believe their government willing to inject capital into the market, thereby encouraging U.S. investors to invest to capital markets, leading to an increase price and returns.

Finally, in the third quantitative easing period, the coefficient of *Ted* does not have significant negative effect on *Ret* in both ETFs and future market. The main reason probably is the market has been injected a large amount of money, leading the excess liquidity in the capital market, therefore the impact of funding liquidity to returns not so significant. These results show strongly evidence that quantitative easing clearly affect the relationship between the money market and the capital market.

Table 9. The panel regression results which classified by the change of Ted spread

Variables	QE1 and ΔTed_d>0	QE1 and ΔTed_d<0	QE1 and ΔTed_d=0	QE2 and ΔTed_d>0	QE2 and ΔTed_d<0	QE2 and ΔTed_d=0	QE3 and ΔTed_d>0	QE3 and ΔTed_d<0	QE3 and ΔTed_d=0	Non QE and ΔTed_d>0	Non QE and ΔTed_d<0	Non QE andΔTed_d=0	Total
Panel A: Global ETF data													
Ted	0.001	-0.055**	-0.003*	-0.014***	-0.005	0.004	0.004	0.094	0.004	-0.001	-0.001**	-0.001***	-0.001***
	(0.44)	(-2.37)	(-1.95)	(-3.52)	(-1.63)	(0.78)	(0.54)	(1.60)	(0.73)	(-0.66)	(-2.21)	(-6.07)	(-7.09)
Vol	0.001	-0.001	0.001	-0.001	-0.001	-0.001	0.001	0.001**	-0.001	-0.001***	-0.001	-0.001***	-0.001***
	(1.49)	(-1.50)	(0.39)	(-0.40)	(-0.13)	(-0.91)	(0.59)	(2.20)	(-0.30)	(-2.91)	(-0.18)	(-3.17)	(-3.55)
Volatility	5.257**	-2.760	6.209***	20.237***	5.597	-8.429**	0.624	32.086**	1.998	0.936***	11.561***	0.016	0.763***
	(2.33)	(-0.29)	(-4.37)	(6.07)	(0.23)	(-2.23)	(0.33)	(2.37)	(1.21)	(4.92)	(14.56)	(0.11)	(6.91)
Ret_{d-1}	-0.004	0.113	-0.125***	0.139	0.180	0.005	-0.057***	0.051	0.001	-0.075***	-0.142***	-0.078***	-0.075***
	(-0.17)	(0.70)	(-7.51)	(0.71)	(0.74)	(0.20)	(-3.52)	(0.75)	(-0.09)	(-14.42)	(-6.48)	(16.02)	(-23.60)
GDP	-0.001**	0.001	0.001	0.009	-0.001	-0.001***	-0.001**	0.001	-0.001	0.001	-0.001**	-0.001	0.001
	(-2.45)	(0.94)	(0.60)	(1.23)	(-0.52)	(-3.70)	(-2.35)	(0.77)	(-1.28)	(0.13)	(-2.49)	(-1.16)	(-1.20)
Intercept	-0.004**	0.029**	0.004***	0.008**	0.028	-0.001	-0.001	-0.050*	-0.001	-0.001	-0.001	0.001***	0.001***
	(-2.44)	(2.17)	(3.40)	(2.48)	(1.10)	(-0.38)	(-0.15)	(-1.84)	(-0.44)	(-0.92)	(-0.48)	(2.85)	(4.87)
Panel B: Global Future data													
Ted	-0.005	0.007**	0.001	-0.017	-0.005	0.001	0.023	-0.001	0.001	-0.001	-0.001***	0.001	-0.001***
	(-1.28)	(2.45)	(0.01)	(-1.41)	(-0.33)	(0.01)	(1.38)	(-0.05)	(0.01)	(-0.12)	(-5.28)	(0.01)	(-3.03)
Vol	0.001	-0.003	0.001	-0.006	-0.006**	0.001	0.001	-0.007***	0.001	-0.002	-0.002**	0.001	-0.003***
	(0.05)	(-1.38)	(0.01)	(-1.50)	(-2.03)	(0.01)	(0.38)	(-2.66)	(0.01)	(-1.24)	(-2.21)	(0.01)	(-3.94)
Volatility	0.249	2.294**	0.001	9.430***	-8.621***	0.001	3.212**	-0.005	0.001	-0.306***	0.039	0.001	-0.071**
	(0.14)	(-1.97)	(0.01)	(3.55)	(-3.39)	(0.01)	(2.45)	(0.01)	(0.01)	(-4.58)	(0.92)	(0.01)	(2.02)
Ret_{d-1}	-0.101***	-0.174***	0.001	0.012	-0.081***	0.001	-0.050***	-0.084***	0.001	-0.114***	-0.137***	0.001	-0.117***
	(-4.18)	(-10.59)	(0.01)	(0.53)	(-3.17)	(0.01)	(-2.90)	(-4.67)	(0.01)	(-20.39)	(-20.90)	(0.01)	(-31.78)
GDP	-0.001	-0.001	0.001	0.001	-0.001	0.001	0.002*	-0.001	0.001	0.001	0.002***	0.001	0.001
	(-0.01)	(-0.25)	(0.01)	(0.04)	(-0.05)	(0.01)	(1.93)	(-0.72)	(0.01)	(0.42)	(3.52)	(0.01)	(0.88)
Intercept	0.004	-0.001	0.001	0.013	0.004	0.001	-0.014*	0.001	0.001	-0.001	-0.003*	0.001	0.001
	(0.38)	(-0.06)	(0.01)	(1.20)	(0.26)	(0.01)	(-1.77)	(0.17)	(0.01)	(-0.28)	(-1.94)	(0.001)	(0.37)

Notes: This table provides details of the funding liquidity effect on the daily returns on three quantitative easing periods (QE) of the U.S. government and the change of Ted spread in d day and d-1 day (ΔTed_d).We classify ΔTed_dinto three groups are bigger, smaller and equal to zero. QE1 is implemented from March 2009 to March 2010, QE2 implementation time is from the end of November 2010 to June 2011, QE3 implementation time is from September 2012, and NonQE presents other non-quantitative easing periods. Where the dependent variable is the daily returns for country i in day d. The independent variables include the TED spread (Ted_d) and control variables ($X_{i,d}$). The control variable ($X_{i,d}$) includes daily trading volume ($Vol_{i,d}$), daily trading volatility ($Volatility_{i,d}$), lagged daily returns (Ret_{d-1}) and growth rate ($GDP_{i,m}$). In Panel A, we use thepanel regression for 25 different countries on Financial Exchange Traded Fund (ETFs) market. In Panel B, we perform the panel regression for 22 different countries by using Global Future data. The t-values examine whether the regression coefficient is significantly different from zero. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

CONCLUSION

We explore the relation between the funding liquidity of US institutional investors and returns, using data from 25 financial ETFs and 22 financial future markets. We investigate the ways in which funding liquidity may have affected returns during the period from 1 January 2004 to 31 December 2015.

We observe that when the large US institutional investors face insufficient of liquidity, index returns of other countries will be negative effect. These results provide evidence to support for our hypothesis, a liquidity shock can trigger enormous redemption pressure for investors, causing the problems for the financial intermediaries funding. Our results show that the effects of funding liquidity on returns in the different countries are in different levels. Moreover, the impacts are in different level when the monetary market is tighter or looser. We also find evidence that US government monetary policies have significant impact to money market and funding market. In sum, our study provides a better overall understanding of the effect of the liquidity–supplier funding constraint on returns.

APPENDIX

Appendix A. ETF and Future countries list

Country	Country Code	Country	Country Code
Global EFT data		***Global Future data***	
Australia	AUS	Australia	AUS
Austria	AUT	Brazil	BRA
Belgium	BEL	Canada	CAN
Brazil	BRA	Denmark	DNK
Canada	CAN	France	FRA
China	CHN	Germany	DEU
France	FRA	Greece	GRC
Germany	DEU	Hong Kong	HKG
Hong Kong	HKG	Hungary	HUN
Ireland	IRL	India	IND
Italy	ITA	Italy	ITA
Japan	JPN	Japan	JPN
Malaysia	MYS	Malaysia	MYS
Mexico	MEX	Mexico	MEX
Netherlands	NLD	Netherlands	NLD
New Zealand	NZL	Singapore	SGN
Singapore	SGP	South Korea	KOR
South Korea	KOR	Spain	ESP
Spain	ESP	Sweden	SWE
Sweden	SWE	Switzerland	CHE
Switzerland	CHE	Taiwan	TWN
Taiwan	TWN	United Kingdom	GBR
Thailand	THA		
Turkey	TUR		
United Kingdom	GBR		

Appendix B. Classify countries by volume, freedom index, volatility

Group	Global EFT data					Global Future data				
Panel A: Classify countries by daily trading volume										
Lower	Ireland	Netherlands	New Zealand			Australia	Spain			
Medium	Australia	Austria	Belgium	Canada	France	Brazil	Canada	Denmark	Greek	Hong Kong
	Germany	Hong Kong	Italy	Malaysia	Mexico	Hungary	India	Italy	Japan	Malaysia
	Singapore	South Korea	Spain	Sweden	Switzerland	Mexico	Netherlands	Singapore	Sweden	Switzerland
	Taiwan	Thailand	Turkey	United Kingdom		Taiwan	United Kingdom			
Highest	China	Brazil	Japan			France	Germany	South Korea		
Panel B: Classify countries by freedom index										
Free (80-100)	Australia	Canada	Hong Kong	New Zealand	Singapore	Australia	Canada	Hong Kong	Singapore	Switzerland
	Switzerland									
Mostly free (70-79.9)	Austria	Germany	Ireland	Japan	Netherlands	Denmark	Germany	Japan	Netherlands	South Korea
	South Korea	Sweden	Taiwan	United Kingdom		Sweden	Taiwan	United Kingdom		
Moderately Free (60-69.9)	Belgium	France	Italy	Malaysia	Mexico	France	Hungary	Italy	Malaysia	Mexico
	Spain	Thailand	Turkey			Spain				
Mostly Unfree (50-59.9)	Brazil	China				Brazil	Greece	India		
Repressed (0-49.9)										
Panel C: Classify countries by daily trading Volatility										
Small Volatility	Italy	Japan	Malaysia			Australia	Malaysia	United Kingdom		
Medium Volatility	Australia	Austria	Belgium	Brazil	Canada	Canada	Denmark	France	Germany	Hong Kong
	China	France	Germany	Hong Kong	Netherlands	Hungary	India	Italy	Japan	Mexico
	New Zealand	Singapore	South Korea	Spain	Sweden	Netherlands	Singapore	South Korea	Spain	Switzerland
	Switzerland	Taiwan	Thailand	Turkey	United Kingdom	Taiwan				
Large Volatility	Ireland	Mexico				Brazil	Greece	Sweden		

REFERENCES

Admati, A. R., & Pfleiderer, P. (2015). A Theory of Intraday Patterns: Volume and Price Variability. *The Review of Financial Studies, 1*(1), 3-40.

Allen, L., Bali, T. G., & Tang, Y. (2012). Does systemic risk in the financial sector predict future economic downturns? *The Review of Financial Studies, 25*(10), 3000-3036.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of financial Economics, 17*(2), 223-249.

Binswanger, M. (2000). Stock market booms and real economic activity: Is this time different? *International Review of Economics & Finance, 9*(4), 387-415.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic perspectives, 23*(1), 77-100.

Brunnermeier, M. K., & Pedersen, L. H. (2008). Market liquidity and funding liquidity. *The Review of Financial Studies, 22*(6), 2201-2238.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics, 108*(4), 905-939.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics, 31*(3), 281-318.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *The journal of finance, 51*(5), 1681-1713.

Chen, S., & Chang, C.-P. (2015). Should bank loan portfolio be diversified under government capital injection and deposit insurance fund protection? *International Review of Economics & Finance, 38*, 131-141.

Chiu, J., Chung, H., Ho, K.-Y., & Wang, G. H. K. (2012). Funding liquidity and equity liquidity in the subprime crisis period: Evidence from the ETF market. *Journal of Banking & Finance, 36*(9), 2660-2671.

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of financial Economics, 56*(1), 3-28.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics, 10*(4), 407-432.

Comerton-Forde, C., Hendershott, T., Jones, C. M., Moulton, P. C., & Seasholes, M. S. (2010). Time variation in liquidity: The role of market-maker inventories and revenues. *The journal of finance, 65*(1), 295-331.

Duffee, G. R. (1992). *Trading volume and return reversals*. Retrieved from

Duffee, G. R. (1995). Stock returns and volatility a firm-level analysis. *Journal of financial Economics, 37*(3), 399-420.

Fair, Ray C. (2002). Events That Shook the Market. *The Journal of Business, 75*(4), 713-731.

Fama, E. F. (1990). Stock Returns, Expected Returns, and Real Activity. *The journal of finance, 45*(4), 1089-1108.

Fischer, B. (1976). *Studies of stock price volatility changes.* Paper presented at the Proceedings of the Business and Economic Statistics Section.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics, 19*(1), 3-29.

Gallant, A. R., Rossi, P. E., & Tauchen, G. (1992). Stock prices and volume. *The Review of Financial Studies, 5*(2), 199-242.

Goetzmann, W. N., & Massa, M. (2002). Daily Momentum and Contrarian Behavior of Index Fund Investors. *Journal of Financial and quantitative Analysis, 37*(3), 375-389.

Goldsmith-Pinkham, P., & Yorulmazer, T. (2010). Liquidity, Bank Runs, and Bailouts: Spillover Effects During the Northern Rock Episode. *Journal of Financial Services Research, 37*(2), 83-98.

Gromb, D., & Vayanos, D. (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of financial Economics, 66*(2-3), 361-407.

Gromb, D., & Vayanos, D. (2010). Limits of arbitrage. *Annu. Rev. Financ. Econ., 2*(1), 251-275.

GRULLON, G., LYANDRES, E., & ZHDANOV, A. (2012). Real Options, Volatility, and Stock Returns. *The journal of finance, 67*(4), 1499-1537.

Hameed, A., KANG, W., & VISWANATHAN, S. (2010). Stock Market Declines and Liquidity. *The journal of finance, 65*(1), 257-293.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *The journal of finance, 46*(1), 179-207.

Hasbrouck, J., & Seppi, D. J. (2001). Common factors in prices, order flows, and liquidity. *Journal of financial Economics, 59*(3), 383-411.

Huberman, G., & Halka, D. (2001). Systematic liquidity. *Journal of Financial Research, 24*(2), 161-178.

Hutson, E., Kearney, C., & Lynch, M. (2008). Volume and skewness in international equity markets. *Journal of Banking & Finance, 32*(7), 1255-1268.

J. Fleming, M., & Remolona, E. (1997). *What Moves the Bond Market* (Vol. 3).

Jacoby, G., Fowler, D. J., & Gottesman, A. A. (2000). The capital asset pricing model and the liquidity effect: A theoretical approach. *Journal of Financial Markets, 3*(1), 69-81.

Karpoff, J. M. (1987). The Relation Between Price Changes and Trading Volume: A Survey. *The Journal of Financial and Quantitative Analysis, 22*(1), 109-126.

Krishnamurthy, A. (2002). The bond/old-bond spread. *Journal of financial Economics, 66*(2-3), 463-506.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica, 53*(6), 1315-1335.

Kyle, A. S., & Xiong, W. (2001). Contagion as a wealth effect. *The journal of finance, 56*(4), 1401-1440.

Laeven, L., & Tong, H. (2012). US monetary shocks and global stock prices. *Journal of Financial Intermediation, 21*(3), 530-547.

Lamoureux, C. G., & Lastrapes, W. D. (1990a). Heteroskedasticity in stock return data: Volume versus GARCH effects. *The journal of finance, 45*(1), 221-229.

Lamoureux, C. G., & Lastrapes, W. D. (1990b). Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics, 8*(2), 225-234.

Lee, C. M., & Swaminathan, B. (2000). Price momentum and trading volume. *The journal of finance, 55*(5), 2017-2069.

Lin, J.-H., Tsai, J.-Y., & Hung, W.-M. (2014). Bank equity risk under bailout programs of loan guarantee and/or equity capital injection. *International Review of Economics & Finance, 31*, 263-274.

Nyborg, K. G., & Östberg, P. (2014). Money and liquidity in financial markets. *Journal of financial Economics, 112*(1), 30-52.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political economy, 111*(3), 642-685.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance, 44*(5), 1115-1153.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *The Review of Financial Studies, 12*(5), 975-1007.

DECISION SCIENCES INSTITUTE

Linking service innovation and financial performance in hospitals

David D. Dobrzykowski
University of Arkansas
ddobrzykowski@walton.uark.edu

Kathleen L. McFadden
Northern Illinois University
kmcfadden@niu.edu

ABSTRACT

Hospitals rely heavily on their workforce to develop and deliver innovative care. We examine key mediators that link service innovation and financial performance; namely, a committed workplace, experiential quality, and patient satisfaction. Data from three sources reveal important positive and negative relationships that are contingent on teaching status.

KEYWORDS: Healthcare Management

INTRODUCTION

Patients' perceptions of quality and satisfaction are important operational metrics for hospital leaders (Chandrasekaran et al. 2012; Nair et al. 2013). Reimbursement structures are shifting away from volume-based systems to programs that link hospital reimbursement payments to hospital outcomes, including experiential quality and patient satisfaction (Powell et al. 2012; Ding 2014). Concomitantly, as insurance plans increase patient out-of-pocket costs (Salzarulo et al. 2011), patients are becoming more engaged in selecting hospitals based on their level of satisfaction (Salzarulo et al. 2015). Patient Satisfaction is defined as the extent to which patients judge the overall hospital stay as favorable (Marley et al. 2004; Ancarani et al. 2011). Experiential Quality, on the other hand, has more to do with the service encounter and is defined as the extent to which patients judge the physical service setting and the multiple moments of contact with hospital employees as meeting their expectations (Pine and Gilmore 1999; LaSalle and Britton 2003; Gentile et al. 2007; Senot et al., 2016c). As such, experiential quality embodies the physical and social dimensions of a service encounter, sometimes referred to as servicescape in the service operations literature (Bitner 1992; Bitran et al. 2008; Rosenbaum and Massiah 2011). Patient satisfaction can be challenging and costly to achieve because it requires careful management of both the physical environment as well as the service encounter, which necessitates coordination among highly specialized service professionals to collectively deliver fairly customized care (Bitran et al. 2008; Lahiri and Seidmann 2012). Recent research has found that higher patient satisfaction scores are actually linked to increased healthcare costs (Fenton et al. 2012). As such, hospitals are increasingly interested in understanding how to improve patient satisfaction and translate those efforts into financial performance (Press Ganey 2013; Senot et al. 2016c).

Service innovations are frequently pursued for financial motivations, such as increasing net profits (Griffin and Page 1993). In addition, service firms develop innovations not only for financial reasons, but also to achieve other performance objectives like customer satisfaction (Ottenbacher 2007). In healthcare, researchers, government and industry leaders alike have suggested that a service innovation orientation is necessary to improve hospital performance across a multitude of dimensions (Dobrzykowski et al. 2015). This is due, at least in part, to the belief that operational performance measures, like experiential quality, patient satisfaction, and financial performance, are predicated on the service delivery system designed by healthcare operations managers (Cook et al. 2002; Zepeda and Sinha 2016; Zheng et al. 2017). Designing and managing the healthcare delivery system is challenging, given that hospitals operate as professional service organizations (PSOs) (Day et al. 2012; Heineke 1995). Specifically, as PSOs, a hospital's patient satisfaction scores rely on fostering a committed workplace that delivers new and innovative services that create positive patient experiences. This is because PSO's depend heavily on employees (e.g., individual healthcare providers) to possess and skillfully apply abstract expert knowledge during complex, high-customer contact interactions (Goodale et al. 2008). Furthermore, PSOs face a unique challenge of managing a flat management hierarchy with loose subordinate–supervisor relationships (Nembhard et al. 2009). These operational conditions can diminish organizational commitment, shifting loyalties toward self-interests, peers, professional societies and other constituents (Schneller and Smeltzer 2006; Nembhard et al. 2009; Boyer and Pronovost 2010). As such, the operational context is important to consider as hospitals attempt to deliver on service innovation because of the key role played by service providers in designing and delivering services that patients desire (Umashankar et al. 2011).

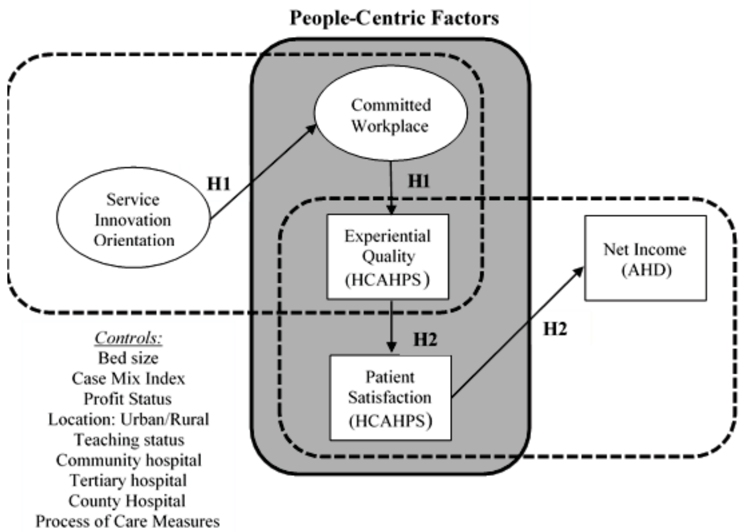
Our study investigates the roles of people-centric factors, namely a committed workplace, experiential quality, and patient satisfaction in linking a hospital's desire for innovation to improved financial performance. We address our research objective by conceptualizing a model based on the services operations (informed by the healthcare, PSO, and servicescape streams) and organizational commitment literature. The model is tested using a unique dataset collected from three sources: 1) primary data collected via survey measuring the extent to which hospitals report a service innovation orientation and committed workplace, 2) secondary archival data from the Centers for Medicare and Medicaid Services (CMS) measuring experiential quality and patient satisfaction, and 3) objective archival data from the American Hospital Directory (AHD) measuring net income.

RESEARCH MODEL

A Service Innovation Orientation is defined as the extent to which a hospital promotes new, pioneering services and is on the leading edge of technology (Dobrzykowski et al. 2015). It is created by hospital leaders and is one of many possible orientations (McFadden et al., 2009; Boyer et al., 2012; Dobrzykowski et al. 2015; Dobrzykowski et al., 2016). This study focuses on a service innovation orientation, given the recent mandates from the United States federal government requiring healthcare delivery innovations. For instance, the Centers for Medicare and Medicaid Services (CMS) is requiring hospitals to deliver experiential quality and patient satisfaction at a lower cost or suffer penalties as high as 2% of their CMS reimbursements (CMS, 2017). In order to achieve positive results from a service innovation orientation, the service organization must be able to understand the voice of the customer and translate their wants and needs into service design specifications that can be used in the delivery of quality service. In a service organization, the front-line employees (physicians, nurses, and other healthcare providers) are in the best position to learn from customer feedback (Dobrzykowski et al., 2015). Therefore, it is important to foster a motivated and committed organizational environment as a means of enhancing quality.

As a theoretical basis for this research, we draw from the service operations (informed by the healthcare, PSO, and servicescape streams) and organizational commitment literature streams to frame relationships among service innovation, committed workplace, experiential quality, patient satisfaction and net income. See Figure 1.

Figure 1: Hypothesized research model



The professional service operations literature motivates our study, and provides contextual support for our research model. A professional service organization (PSO) has been defined as "any organization which provides a service which is based on a professional diagnosis, i.e. on a thorough analysis by a qualified professional in a given field of facts or problems in order to gain understanding and guide future actions" (Harvey 1990, p. 6). The professional services literature highlights the distinct characteristics, environment and managerial challenges required for managing PSO processes (Goodale et al. 2008; Lewis and Brown 2012; Brandon-Jones et al. 2016; Zhang et al. 2016; Lawrence et al. 2016). PSOs' distinct characteristics make the organizational environment a key consideration as managers orchestrate operations with an eye toward organizational performance.

The PSO literature, with its inherently high focus on customer contact/customization, flexible processes with high labor intensity, and the need for "nudging" service personnel rather than implementing standard operating procedures, supports the idea that a committed workplace would enhance experiential quality. A Committed Workplace is defined as the extent to which the work environment enhances employees' attitudes towards work, encourages hard work and high-level job performance (Bargagliotti 2012; Jambulingam et al. 2005; Vera et al. 2016). Gowen et al. (2006) explains that the employee commitment approach has been conceptualized at both an individual and an organizational-level behavioral theory. This study follows the approaches of Bou and Beltan (2005) and Gowen et al. (2006) in conceptualizing workplace commitment at the organizational level.

Hypothesis 1. A Committed Workplace partially mediates the relationship between Service Innovation Orientation and Experiential Quality.

Patient experience is an important aim for hospital leaders as it is considered a cornerstone in the evaluation of health care quality (Danielsen et al. 2010). However, experiential quality may not improve a hospital's financial performance. Some firms have found that their investments to improve service offerings, with an aim toward 'delighting' customers, have proven unprofitable primarily because they fail to benefit from the economies of scale and learning curves realized in manufacturing settings (Rust et al. 1995; Sasser et al. 1997). In hospitals, economies of scale are difficult to achieve due to the substantial fixed costs associated with investments in cutting-edge infrastructure and amenities thought to improve patient experience (Senot and Chandrasekaran 2015). It may also be difficult to achieve consensus regarding infrastructure decisions among professional service providers given their autonomy and influence in the workplace (Harvey 1990; Schneller and Smeltzer 2006; Boyer and Pronovost 2010). Learning curves are hampered by the volume and variety of sources considered by patients in evaluating the quality of their experience (i.e., nurses, physicians, other staff members, and even "hotel-like" amenities such as the cleanliness of their hospital room) (LaSalle and Britton 2003). What is clear is that the customer plays an important role in the profitability of service firms (Bolton 1998). "Competition for healthcare dollars often rests on how satisfied patients are with their provider," (Russell, Johnson and White 2015: p. 1159).

Hypothesis 2. Patient Satisfaction partially mediates the relationship between Experiential Quality and Hospital Net Income.

METHODS

Survey data were collected from 166 hospitals in 46 states in the USA, and paired with data from two other sources – 1) archival data from the Centers for Medicare and Medicaid Services (CMS) collected using individual patient surveys to measure experiential quality and patient satisfaction, and 2) objective data on net income from the American Hospital Directory (AHD).

Instrument Development

We collected survey data to measure service innovation orientation and committed workplace as multi-dimensional psychometric measures. These are hospital phenomena which are behavioral in nature and thus it is appropriate to capture these variables using psychometric approaches (Gowen et al. 2006; Dobrzykowski et al. 2015). We developed measurement instruments using a three step process (Churchill, 1979; Bardhan, et al. 2007).

Measures

Service innovation orientation and committed workplace were collected from primary data with the scales developed using the procedures described above. Experiential quality and patient satisfaction were collected from the Center for Medicare and Medicaid Services (CMS) Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey. HCAHPS measures patients' perspectives on their care following inpatient stays at acute care facilities in the USA (see also Chandrasekaran et al. 2012; Ding, 2014; Senot et al. 2016b). The HCAHPS survey is administered to a random sample of patients on an ongoing basis and reported publically via the Hospital Compare website. This process produces a dyadic sample – paired sets featuring hospitals reporting on operational practices and patients reporting on their experience and satisfaction. The mean of the seven items was calculated to produce an aggregate measure of experiential quality. This methodological approach is common in the OM literature (See Boyer et al. 2012; Goldstein and Iossifova 2012; Queenan et al. 2011). Patient satisfaction was also collected from the HCAHPS survey and deals with the overall hospital experience (Ancarani et al. 2011; Marley et al. 2004).

Our dependent variable is net income, which is reported as a percentage by the American Hospital Directory (AHD). AHD is a publically available data source providing financial performance data for more than 6,000 hospitals in the USA. AHD data is derived from both public and private sources, including the Federal Centers for Medicare and Medicaid Services. Smith et al. (2013) is a related study that has analyzed AHD net income data.

Data Collection

The dataset was populated by a cross-sectional self-administered internet-based survey. The sample frame was created from a random list of acute care facilities that are members of the American Hospital Association. A telephone solicitation method was employed to populate the sample frame with email addresses of prospective respondents (see McFadden et al. 2009). Our response rate is 46.8% (302 responses / 644 prospective acute care facilities). Two statistical tests were conducted to examine non-response bias (Armstrong and Overton 1977): the chi-square test and t-tests (Bardhan et al. 2007). Chi-square tests performed on dichotomous binary variables for hospital type (tertiary, community, or critical access) as well as for membership in a hospital system. A t-test examined mean differences for bed size, similar to Meyer and Collier (2001). A total of 124 hospitals declined participation in the survey and were designated as non-respondents. We collected non-respondent data for hospital type using the internet, and data for bed size and system affiliation membership was provided from the AHA. These tests produced no statistically significant differences between the respondents and non-respondents, providing evidence of an absence of non-response bias (Armstrong and Overton 1977).

Analysis and Results

The analysis and results will be discussed at the conference.

Discussion and Conclusion

Our study makes several important contributions to the literature that ought to influence the way hospital operations leaders design healthcare services with an aim toward satisfying patients and improving hospital net income. First, we show that service innovation plays a dual role in positively influencing workplace commitment, but negatively affecting experiential quality and patient satisfaction. Second, we found that service innovation has a positive indirect effect on experiential quality, partially mediated by a committed workplace. Third, experiential quality and patient satisfaction do not generally drive net income. Fourth, our results show that the pathway from service innovation orientation to hospital financial performance is contingent upon teaching status, or a hospital’s strategic positioning toward differentiation. We find that teaching status mutes the negative effect of service innovation on patient satisfaction, and amplifies the effects of experiential quality and patient satisfaction on net income. Thus, it is important to understand how their hospital’s strategic position influences the patient outcomes they achieve from an innovation orientation, and when (under what conditions) to expect positive financial results.

References available upon request.

DECISION SCIENCES INSTITUTE

Blockchain Applications in Engineering and Manufacturing: A Systematic Review

Joseph Kasten
The Pennsylvania State University – York Campus
1031 Edgecomb Ave.
York, PA 17403
jzk99@psu.edu
1-717-718-6775

ABSTRACT

The paper presents the results of a systematic review of the literature pertaining to the use of blockchain technology in the engineering and manufacturing activities. The uses of blockchain have been growing quickly in the literature, and this fast pace can make it difficult for researchers to keep abreast of the state of the art. Using a systematic approach to search the primary electronic indices, the compiled list of papers is then analyzed and a number of themes are identified. The paper concludes with a set of conclusions about the current state of the literature and suggestions for improvement.

INTRODUCTION

Interest in blockchain technology, the data management method underlying cryptocurrencies, has gotten the attention of researchers and developers in a widening number of industries. After finding a following in the financial industry (Nakamoto, 2008), it has become a focus of interest in the supply chain (Treblmaier, 2018), insurance (Sehgal, 2017), and agriculture (Lin et al, 2017) industries. In many of these industries, the interest has only begun to take shape within the past few years as an understanding of blockchain has become more widespread. In some fields, such as engineering and manufacturing, the evaluation of blockchain as an appropriate tool did not begin until a few years ago, but has continued with an increasing rate and scope. Because of this phenomenon, it becomes difficult for researchers and practitioners to keep abreast of the current approaches to the use of the tool. This situation is exacerbated by the wide variety of current research streams that include blockchain as well as the sheer number of possible outlets for this research. While journals are relatively easy to identify with the use of one or more widely available electronic databases, conference proceedings can be much more challenging to identify and obtain. To assist researchers and practitioners in gaining an understanding of the current state of the research, a systematic review of the literature that surrounds the application of blockchain to engineering and manufacturing systems is undertaken in this paper.

Engineering and manufacturing systems are targeted in this review because of their importance to the creation of value and wealth in any society as well as the very nature of these activities. Both engineering and manufacturing involve the creation and use of high-precision information that must often be shared between users and departments within a firm, or between organizations within a supply chain or joint venture. The distributed nature of these operations, both geographically and organizationally, creates vulnerabilities as these information items circulate within the system. Each change of possession or transformation event must be tracked within a record that cannot be changed or deleted as they are of vital importance to both the success of the enterprise and its ability to analyze its internal processes toward the goal of continuous improvement. Blockchain's ability to provide data immutability in a widely distributed environment is one of its most important features and makes it a likely candidate for use in these fields.

The paper is organized as follows: The next section provides an overview of blockchain technology and extends the discussion as to its application to the engineering and manufacturing processes, the following section describes the methodology followed in the review, the findings of the review are detailed next, and that is followed by a discussion and conclusion including the limitations of the study and directions for additional research.

BLOCKCHAIN OVERVIEW

Blockchain is a distributed ledger data storage mechanism that allows data to be stored in a manner that is available to all entities having the proper level of permission but does not allow for the changing of those data without the notice of all other entities in the ecosystem (Nakamota, 2008). In order to accomplish these two objectives, blockchain utilizes a cryptographic hashing mechanism to determine how data are stored on the chain and to prevent it from being changed. Data to be stored on the chain is combined into blocks of a predetermined size. These blocks are presented to the entities entrusted with the control of the block additions, which might be one entity on a private blockchain architecture or many on a public chain. In order for the block to be added to the chain, a hash puzzle must first be solved. The hash puzzle requires that a random number (called the “nonce”) be found that, when hashed together with data from the new block and the last block added to the chain, results in a hash reference value that meets some predetermined criterion (e.g. three leading zeros). The properties of a cryptographic hash function are such that knowledge of the hash reference requirement does not assist in the identification of the nonce, which means that the entities wishing to solve the puzzle must guess at a nonce and then execute the hashing algorithm repeatedly until the nonce is found. Only when the nonce is finally identified and verified can thee new block be added. After that occurs, work begins on the next block to be added.

The data immutability inherent in blockchain derives directly from this hashing process. Once data are successfully added to the chain, another property of cryptographic hashing requires that any change to the underlying data in a given block necessarily changes its hash value, rendering the solution to the hash puzzle using those data invalid, which in turn renders all subsequent hash solutions invalid (Drescher, 2017). The longer the data are on the chain, the further “buried” it becomes in subsequently added data, and the more difficult it is to change without all other hashes being affected. Therefore, data added to the blockchain remains unchanged and can act as an “official” or immutable version of the data, which is very useful for data that is being used in important and/or expensive operations across multiple organizational or national boundaries, such as exists in a typical engineering or manufacturing environment.

Blockchains have additional characteristics or options that make them a compelling choice for the applications being considered in this paper. The option exists to store data not on the blockchain, but rather in an off-chain storage facility with only the hash value of the data stored on the chain. This has the advantage of increasing the speed of adding blocks to the chain as well as removing the ability of entities with access to the chain to view the data of other entities in the chain environment. This allows all entities to validate the data by comparing hash values with those held in off-chain storage without divulging the actual data. The speed by which blocks are added can also be increased by reducing the complexity of the hash puzzle. An easier puzzle will enable the puzzle to be solved more quickly and still provide data immutability, but might only be applicable to a private chain with only trusted parties involved. Lastly, access to the blockchain can be left open to all (permissionless) or to only those with the need to access it (permissioned). These permissions can be customized to permit only those entities that own the chain to write to it but others that depend on it for verification to have read only access.

With these and other beneficial characteristics available to blockchain designers and users, there would seem to be several aspects of engineering and manufacturing that would benefit from a blockchain-based architecture. The methodological approach to reviewing this literature is detailed in the following section.

METHODOLOGY

In order to perform a systematic literature review, a repeatable and verifiable process must be in place (Briner & Denyer, 2012). For the present research, the process used consists of the following steps:

- Identify the research question(s).
- Locate and select relevant studies
- Critically appraise the studies
- Analyze and synthesize the findings
- Disseminate the findings

The research question for this review is:

- For what aspects of the engineering and manufacturing processes have researchers or developers proposed or applied blockchain?

Within this research question, two research objectives are also investigated:

- Within the engineering/manufacturing blockchain literature, what types of methodologies are being used?
- Within the engineering/manufacturing blockchain literature, how long has blockchain been investigated as a possible solution technology?

To locate relevant studies, a set of “ground rules” are developed to frame the literature to be evaluated. For the current study, the focus will remain on peer-reviewed articles in journals, conference proceedings, or book chapters in edited volumes. Future expansion of the study will involve other sources such as trade journals, unpublished studies, and other so-called “gray” literature in order to both identify any applications not covered in the present study and to understand any differences in emphasis between the traditional academic literature and other sources of information.

Searches are performed in major electronic databases including ABI Inform Global, Emerald, IEEE Explore, JStor, ScienceDirect, Scopus, and Springer, using the following keywords in all possible combinations: blockchain (including digital ledger, distributed ledger, and shared ledger), engineer(ing), design, Manufacture(ing), and Industry 4.0. Searches were accomplished in the article title, abstract, and keywords. All articles having one or more of these search terms were selected for initial inclusion in the study. This initial search resulted in sixty-nine studies.

The selected studies were examined to ensure that they actually were focused on the topics at hand. In most cases, the abstract was sufficient to determine the article’s suitability for the study, however some abstracts were not sufficiently descriptive and the full text of the article was consulted. After this step, six studies were removed as they were not specifically focused on either engineering or manufacturing but were rather more general descriptions of blockchain that only mentioned its possible application to these processes. The results of the analysis of the included studies are presented in the following section.

FINDINGS

The results of the literature analysis are divided into two sections. The first section presents the descriptive statistics of the literature and describes the characteristics of the body of work derived based on the previously described inclusion criteria. The second section describes the themes identified in the literature that describe the research surrounding the application of blockchain to engineering and manufacturing.

Descriptive Statistics

The sixty-two articles included in the study have the following characteristics:

- 50 manufacturing applications, 12 engineering applications
- Only nine articles include any empirical analysis, the rest are descriptive or propositional
  - 6 studies utilized some form of experiment to validate the proposed system
  - 3 studies used a formal case study to demonstrate the proposed system
- Sixteen studies are focused on a particular industry, the rest are more general in nature.
  - construction (9)
  - automotive (2)
  - pharmaceuticals (2)
  - chemical (1)
  - appliances (1)
  - mining (1)
- The articles are published in one of the following three forms:
  - conference proceedings (31)
  - journal articles (26)
  - book chapters (5)

The search process described in the article identification protocol does not use any date specification. Nevertheless, there were no articles retrieved prior to 2016. Figure 1 shows the rapid rise in the number of studies applying blockchain to engineering and manufacturing. The data reflect articles retrieved on May 1, 2019, which suggests that the number of applicable studies published during 2019 will likely be even higher if the pace of publication continues.

Research Themes

After examining each of the articles included in the study, six research themes or emphases emerged, and these will be detailed in this section. There is necessarily a certain amount of overlap between these categories, largely because of the level of integration of the applications being described or proposed. However, the papers are categorized based on the most significant contribution put forth by the authors.

Blockchain and Data Validity

The first, and most popular (21 papers), research emphasis centers on blockchain's ability to provide assurance regarding the validity of a value, information entity, or even an individual. This ability is applied in order to prove the identity of a person or part, determine the level of access of a user, and more broadly, to create trust between two parties that might not have much of a basis to trust each other. This is in keeping with blockchain's reputation as a trust-

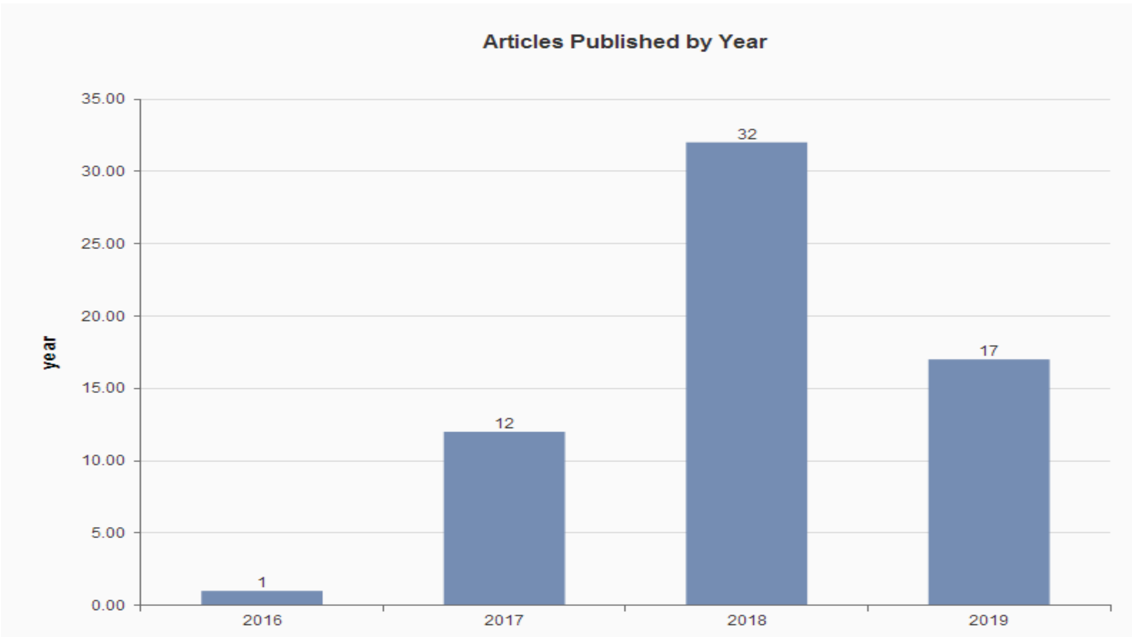


Figure 1: Number of Articles Published Applying Blockchain to Engineering or Manufacturing as of May1, 2019

building tool (Kundu, 2019). Most of the papers in this category discuss blockchain's potential contribution to trust-building from a theoretical perspective, but some of them offer system designs that propose in more detail how blockchain can contribute to inter-organizational or interpersonal trust. For example, Mohamed and Al-Jaroodi (2019) describe a middleware approach to provide more trustable and reliable communications between autonomous manufacturing machines in a distributed manufacturing environment. In the same vein, Oak, Jhala, & Khare (2018) proposes a solution that ensures that none of the multiple parties in a design process dominates the process and makes changes to the design without notification of the other contributors. This is written for a software design environment but is equally applicable to other design and manufacturing processes. Lastly, Kasten (2019) describes a blockchain-based system to ensure that no changes can be made to a CAD-based product design without all parties in the process, including both internal and external partners, knowing about it. This allows the parties to have trust in the data they are using, if not in the parties they are working with. Table 1 lists the papers in this research theme, its research paradigm (description of possible use or concrete system proposal) and their key findings.

Each of the papers in this section leverage the data immutability properties of blockchain to provide the services described. With data embedded on the chain, or at least their hash references, there is very little chance that the data can be changed without the notice and approval of the other entities in the ecosystem. This characteristic takes the place of trust in providing confidence that the data is as it was supposed to be at the time of its creation, regardless of the intention of some to tamper with it. Thus, data maintains its validity and processes that depend on it can continue safely without the overhead of a third party to maintain its safety.

Author(s)	Descriptive/Proposal	Key Findings
Abdullah & Faizal (2018	descriptive	BC applications to security
Anto & Nugraha (2018)	proposal	BC system for calibration of digital multimeters
Belle (2017)	descriptive	BC to increase trust in construction industry
Doumas & Lombardi (2018)	proposal	Validating CAD models w/BC
Fernández-Caramés & Fraga-Lamas (2019)	descriptive	Use BC to add trust in distributed mfg system
Fraga-Lamas & Fernández-Caramés (2019)	proposal	Use of BC to increase security, privacy, and anonymity
Holler, Barth & Fuchs (2019)	descriptive	BC use in managing product lifecycle in automotive ind.
Holtkemper & Wieninger (2018)	descriptive	BC use in making data tamper proof and secure
Kobzan et al (2018)	descriptive	BC use in making data tamper proof
Kasten (2019)	proposal	Use of BC to ensure CAD model validity
Kremenova & Gajdos (2019)	descriptive	Use of BC to improve security and verifiability of data
Lemes & Lemes (2019)	descriptive	Use of BC to integrity of CAD data
Li, Greenwood & Kassem (2019)	descriptive	BC use in political and social aspects of engineering/mfg
Mohamed & Al-Jaroodi (2019)	descriptive	BC to create more trustable and traceable smart mfg.
Nuss, Puchta & Kunz (2018)	descriptive	BC to create tamper-proof DB
Oak, Jhala & Khare (2018)	proposal	BC middleware to coordinate collaborators in dist. mfg system
Plotnikov & Kuznetsova (2018)	descriptive	BC's role in creating trust in pharmaceutical mfg.
Towne (2019)	proposal	BC to keep track of part authenticity certificates
Turk & Kline (2017)	descriptive	BC for use in creating trustworthiness of construction docs and processes
Skwarek (2017)	proposal	Use of BC to achieve data reliability in communications between sensors and machines
Yin et al (2017)	descriptive	BC to address security issues in M2M communications

Table 1: Papers focusing on use of blockchain to increase trust and validity

Blockchain and Communication

Blockchain is identified by these authors as a tool to enhance the ability to communicate and to foster interaction among the entities in a system, be they machines, organizations, or humans. The twelve papers in this section include a number of proposals for tools that use blockchain to enhance the communication abilities of various types of systems to achieve various goals that support the overarching goals of the organization. Some examples include an Internet-of-Things (IoT) approach for gathering data from the structure supporting the walls and ceilings of underground mines to provide more meaningful warnings to miners (Singh, Kumar & Hötzel, 2018), a system that will assist in the management of the product design process (Papakostas, Newell & Hargaden, 2019), and helping to track the specific parts applicable to a specific version of an automobile (Heber & Groll, 2018). Interestingly, there are a higher percentage of proposal-based papers in this grouping, possibly due to the relative ease of conceiving of the role of blockchain in the communications process. Table 2 provides an overview of the papers discussing the role of blockchain in enhancing communications and integration.

Author(s)	Descriptive/Proposal	Key Findings
Ablyazov & Petrov (2019)	descriptive	Use BC to improve communication among construction participants
Afansez et al (2018)	descriptive	BC to create network of industrial components
Chung et al (2019)	proposal	Use BC to help interpret sensor data from equipment and human body
Downey, Bauchot & Röling (2018)	descriptive	BC for contract and workflow
Goldstein & Imbault (2018)	descriptive	Use of BC in fractal manufacturing
Heber & Groll (2018)	proposal	Use BC to trace products throughout product lifecycle management
Papakostas, Newell & Hargaden (2019)	proposal	BC for managing product design process
Sikorski, Haughton & Kraft (2017)	proposal	BC use to facilitate machine-to-machine interactions
Singh, Kumar & Hötzel (2018)	proposal	BC and IoT to enhance mine safety
Yan et al (2017)	proposal	BC to facilitate the cooperative dev of power electronic devices
Zhang, Liu & Shen (2017)	proposal	BC to coordinate collaborative manufacturing
Li et al (2018)	proposal	BC to promote knowledge sharing within a manufacturing ecosystem

Table 2: Papers focused on using blockchain to enhance communications

The distributed nature of the blockchain technology can take the place of more complicated communication processes, thus simplifying the inter- or intra-organizational communications that need to take place. With the peer-to-peer communications process inherent in a distributed ledger architecture, the requirements of communications, local access, data immutability and validity can all occur simultaneously without the need for external monitors to moderate the process. Depending on the particular architecture, certain rules might be needed to control the order in which data are added to the chain or permissions might be adjusted to ensure that only certain entities can add data, but once in place these rules will maintain order on the chain without user or organizational intervention.

Blockchain and Cloud Manufacturing

Cloud manufacturing can be defined as an outgrowth of cloud computing in that manufacturing resources are located outside of the manufacturing firm (Zissis & Lekkas, 2012). These resources can be strictly manufacturing process (machining, assembling, etc.) or can be broadened to include ancillary processes such as engineering and design, testing, supply chain, and other activities necessary for the creation of physical products. Like cloud computing, these resources are interacted with via the Internet, usually including the provision of engineering and manufacturing data, contracts, and other information necessary for the creation of products. The relationships can be short or long term, and involve entities spread around a town or around the world. With that as background, the third research theme is the use of blockchain in the cloud manufacturing environment (11 papers).

The papers that are placed in this category utilize blockchain to address certain issues found specifically, but not exclusively, in cloud manufacturing, such as trustworthiness and distributed manufacturing networks. Bahga and Madisetti (2016) combine blockchain with the Internet of Things (IoT) concept to improve the level of interaction between peers without the need for an intermediary. A cloud manufacturing system with thirty-two manufacturing services is evaluated as to its ability to provide increased scalability and security using blockchain over more traditional, centrally controlled data storage approaches (Li, Barenji & Huang, 2018). Lou et al (2018) simulate the use of a blockchain-based approach in the coordination of distributed manufacturing resources using a punishment mechanism to promote cooperation. The papers comprising the cloud manufacturing theme are presented in Table 3.

As mentioned above, the papers in this section utilize the characteristics of blockchain to enhance the ease, accuracy, and efficiency of the cloud manufacturing process. By creating an agreed-upon version of the data required for this distributed system to operate and leverage the inherent advantages of the cloud platform, blockchain acts as a lubricant of sorts to ease the flow of data (and therefore trust) between machines (Barenji, Li, & Wang, 2018), organizations (Li et al, 2018), and credit-granting organizations (Liu, Jiang, & Leng, 2017).

Author(s)	Descriptive/Proposal	Key Findings
Angrish et al (2018)	proposal	Verification of BC data on distributed network
Bahga & Madisetti (2016)	proposal	Enable peers to interact without trusted intermediary
Barenji et al (2018)	descriptive	BC for decentralized peer to peer network
Barenji, Li & Wang (2018)	proposal	Use BC to promote machine-to-machine comm.
Innerbichler & Damjanovic-Behrendt (2018)	proposal	Use Federated Byzantine Agreement to ensure trust between federated nodes
Isaja & Soldatos (2018)	descriptive	Uses of BC in decentralized manufacturing
Li, Barenji & Huang (2018)	proposal	Use BC to improve security and scalability of dist. mfg
Li et al (2018)	proposal	Use BC for knowledge sharing for injection mold redesign
Liu, Jiang & Leng (2017)	proposal	Use NC to coordinate credit assurance within social manufacturing context
Lou et al (2018)	proposal	Use NC to encourage cooperation among nodes in cloud manufacturing
Yu et al (2019)	proposal	Use BC to enhance information transparency and decentralization

Table 3: Papers focused on using blockchain to support cloud manufacturing

<u>Blockchain and Efficiency</u>

The fourth theme of research is the use of blockchain to improve the efficiency of a manufacturing process or system and to increase the value it brings to the product and, thus, to the firm (10 papers). If the literature is a guide, there are a number of areas in manufacturing and engineering that can be made less costly and bring more value using blockchain in some manner. Sorting of materials and goods can be an expensive, manual process, but Petrali, Isaja, and Soldatos (2018) provide guidance in the use of blockchain and smart contracts to model and optimize the sorting process. Carreno et al (2018) use blockchain to control data entry into an installation design expert system. The costs of data storage and manipulation using blockchain and traditional database technologies are compared by Rimba et al (2018). The papers that utilize blockchain to bring about efficiency and increase value are listed in Table 4.

Author(s)	Descriptive/Proposal	Key Findings
Backman et al (2017)	proposal	Use BC to enhance autonomous equipment leasing
Berger, Penzanstadler & Drögehorn (2018)	descriptive	Identify use of BC in safety-critical systems to reduce liability
Carreno et al (2019)	proposal	BC as a means of controlling data intro to designer ES
Isaja & Soldatos (2018)	proposal	Industrial autonomous systems using BC and edge computing
Miller (2018)	proposal	BC enabled sharing of IoT data to all participants in business network
Kapitonov et al (2018)	descriptive	BC supported autonomous agent network
Liu & Cai (2018)	proposal	automatic decision-making value system for pharma
Mohamed, Al-Jaroodi & Lazarova-Molnar (2019)	descriptive	Describes the role of Industry 4.0 technology, inc BC, to reduce greenhouse emissions and costs
Petrali, Isaja & Soldatos (2018)	proposal	Use of BC to support optimal sorting mechanism
Rimba et al (2018)	descriptive	compares cost of BC and traditional RDBMS

Table 4: Papers focused on using blockchain to increase efficiency and value

Blockchain and Building Information Model

A building information model (BIM) is used in the construction industry to plan, design, construct, and maintain buildings (Belle, 2017). They are essentially 3D Computer Aided Design models that display the foundation, structure, electro-mechanical lines, HVAC, and other aspects of a building’s construction. It is used as a design and engineering tool, a construction guide, and a repository of data that can be used to efficiently run the building as well as plan for improvements. This understanding of BIM should relate easily to the use of blockchain. As a single source of information, a technology that provides for data immutability while facilitating access would seem to fit well. Zheng et al (2019) find that blockchain is useful in providing for security of the BIM as well as guarantees that the current data has not been tampered with. Zheng et al (2019) provide for a blockchain-based BIM that provides for context-aware access of data, thereby creating a much more responsive construction tool. Table 5 presents the five articles in this theme.

Author(s)	Descriptive/Proposal	Key Findings
Li, Greenwood & Kassem (2019)	descriptive	Suggest BC not yet mature enough for BIM use
Mathews, Robles & Bowe (2017)	descriptive	Explores role of BC in the evolving construction industry
Zheng et al (2019)	proposal	BC used to facilitate BIM data security and audit
Zheng et al (2019)	proposal	BC to provide context-aware BIM access
Ye et al (2018)	descriptive	BC to increase security and transparency of BIM

Table 5: Papers focused on using blockchain for BIM

Blockchain and Additive Manufacturing

Additive manufacturing, sometimes known as 3D printing or direct digital manufacturing, is the process of creating a part from a CAD model by applying layers of material (plastic, metal, etc.) until the final part shape is reached. This technology can be used to build prototypes and production parts. Because it is always created from a CAD model, the validity of the model is of paramount importance. The three papers in this last group focus on anti-counterfeiting (Kennedy et al, 2017) and intellectual property protection (Engelmann et al, 2018; Holland et al, 2017). These papers are listed in Table 6.

Author(s)	Descriptive/Proposal	Key Findings
Engelmann et al (2018)	descriptive	BC as a tool to protect intellectual property
Holland et al (2017)	descriptive	BC as a tool to protect intellectual property
Kennedy et al (2017)	proposal	Use BC to trace lanthanide nanomaterial to prevent counterfitting

Table 6: Papers focused on using blockchain in additive manufacturing

DISCUSSION

There are a number of themes that emerge from these results beyond those of the research areas already identified. The first is the lop-sided nature of the results in terms of the ratio of engineering to manufacturing research emphases. This review was planned to include both engineering and manufacturing because of the close relationship of the two disciplines. In essence, they are highly interconnected segments of the same value chain (Porter, 1985). Therefore, a simultaneous search of both literatures was a logical approach to the project. In reviewing the papers included in this review, the reason for the emphasis on manufacturing is that most of them concentrate on the use of blockchain to improve the logistical aspect of manufacturing (production line management, part authentication, contract and workflow management) or the management of manufacturing information (intellectual property of additive manufacturing, machine-to-machine communication, credit assurance validation), both of which have already been shown to be useful applications of blockchain in other contexts (Wang, Han, & Beynon-Davies, 2019). The use of blockchain in engineering seems to be limited, at the present, to verification and validity, which is another application of blockchain that has been shown to be very useful (Kasten, 2019).

The second theme that is apparent from this collection of papers is that there has been a significant amount of prescriptive activity in this area of inquiry, but very little in the way of empirical analysis of the applicability of blockchain to the engineering or manufacturing processes. Some of the papers discuss the evaluation of blockchain in a well-controlled experimental or simulation environment, and others apply a limited case study approach, but there has yet to be a substantial empirical investigation performed that would provide evidence of the applicability, or lack thereof, of blockchain to these activities. This is similar to the application of blockchain to other activities such as supply chain (Treiblmaier, 2018) and healthcare (Casino, Dasaklis, & Patsakis, 2019). The cause of this imbalance has a few roots, not the least of which is the relative newness of this technology beyond its initial introduction to the financial sector. The level of activity displayed over the previous two years suggests that there will be an increasing focus on empirical research in the near future as long as these proposed systems become reality. However, there are significant barriers to the implementation of blockchain-based systems, as there are with any new technology.

Despite the level of interest in both the academic and popular literature, there are pervasive questions surrounding the performance and complexity of blockchain-based solutions, as well as questions of whether this is a ground-breaking technology or a solution looking for a problem. These misgivings are being dealt with in a theoretical basis in the literature, as shown by this review, but they will act as significant barriers to implementation until they are investigated by empirical research. Blockchain, and the literature surrounding it, need to mature so that the normal resistance to change often brought about by technology change can be addressed. This expansion of the literature will also serve to reduce the uncertainty felt by decision-makers about how it works, what benefits it brings, and the expenses involved in its implementation.

There was very little mention of the changes to existing organizational processes, if any exist, required to introduce blockchain-based systems. Users' perception of a system's usability as well as their perception of how it will help them do their job and bring value to the organization has a great deal to do with how well a newly implemented system is accepted by the user community (Davis, 1989). Many of the systems described will operate in the background and as such will have little impact on user processes or the user interface, but in some instances a new way to look at manufacturing data and its management will require a change in how processes, and even the data itself, are conceived of on the part of both the user and management. So, like any other technological innovation, organizational change processes must accompany these implementations. Moving up the decision-making hierarchy, management must first determine if the evolution of their technological architecture requires the use of blockchain in order to support their strategic arc. As always, the alignment of technology and strategic direction is essential to the success of the organization's attempt to use technology to gain a competitive advantage.

The last theme is also one that becomes apparent with its absence. Blockchain has difficulty accepting changes to data already existing on the chain. This difficulty is due to the requirement of solving the hash puzzle to add the next block to the chain. Changing data is difficult, or possibly impossible, because it would disrupt the hash reference of the block being changed as well as the hash references of all subsequent blocks, requiring all subsequent hash puzzles to be solved again. In cases of chains with millions of blocks and fairly difficult hash puzzle requirements this could be an infeasible process. However, engineering and manufacturing processes require the ability to revise data easily and quickly, while retaining the validity of its original form. In some blockchain systems, revisions are handled by adding a block with the revised data to the end of the chain and revising any connected pointers accordingly. However, this might require significant programming effort and might not be feasible in all environments. Further research is needed to fully understand this shortcoming (some will consider it a feature that provides data immutability) in the various applications outlined in the papers that make up this review.

CONCLUSION

This paper provides a systematic review of the extant academic literature that describes the use of blockchain technology in the engineering and manufacturing activities. The literature surrounding these two activities were reviewed together because of the close relationship between these two pursuits and, therefore, the presumed close relationship between their literatures. The review of the literature finds that while there is a quickly growing body of work that provides descriptive and proposed uses of blockchain in these fields, there is scant empirical analysis that accurately describes how well this technology serves these activities. Thus, this article serves as a call for not only an effort to examine existing blockchain installations in these and other fields, but also for a broadening of the exploration of blockchain applicability to these and other undertakings.

There are a few limitations of this study. First, while this is a systematic review of the literature, it cannot claim to be exhaustive. Most of the major indices have been examined, but not all, though the level of duplication of publications experienced leads to the conclusion that the research is very near the point of exhaustion. The second limitation is that the study does not include non-academic literature such as trade journals, unpublished studies, and the like. Third, it does not include literature found by searching terms such as Internet of Things, autonomously guided vehicles (AGV), or other terms which have applicability in many industries beyond engineering and manufacturing. The reason for their omission is to decrease the complexity of this stage of the study. They will be included in subsequent stages.

Future directions of enquiry include the inclusion of other forms of literature and search terms, as noted above, to provide a fuller picture of the state of the knowledge base surrounding this topic. It is hoped that a review article such as this, which is by definition only a snapshot of the field under examination, provides a starting point for future researchers and practitioners to address a particular research direction, and that it can form a foundation for research that will add to the work already accomplished.

REFERENCES

Abdullah, R. S. & Faizal, M. A., (2018), Block chain: Cryptographic method in fourth industrial revolution. *International Journal of Computer Network and Information Security, 11,* 9-17.

Ablyazov, T., & Petrov, I., (2018), Influence of blockchain on development of interaction system of investment and construction activity participants. *IOP Conference Series: Materials Science and Engineering, 497,* 1.

Afanasev, M. Y., Fedosov, Y. V., Krylova, A. A., & Shorokhov, S. A., (2018), An application of blockchain and smart contracts or machine-to-machine communications in cyber-physical production systems. *2018 IEEE Industrial Cyber-Physical Systems (ICPS),* 13.

Angrish, A., Craver, B., Hasan, M., & Starly, B., (2018), A case study for blockchain in manufacturing: "FabRec": A prototype for peer-to-peer network of manufacturing nodes. *Procedia Manufacturing, 26,* 1180-1192.

Anto, I. A. F., & Nugraha, I. G. B. B., (2018), Blockchain-based for calibration of digital multimeter system design. *International Conference on ICT for Smart Society (ICISS),* 1-6.

Backman, J., Yrjölä, S., Valtanen, K., & Mämmelä, O., (2017), Blockchain network slice broker in 5G. *IoT Business Models, Users, and Networks,* 1-8.

Bahga, A. & Madisetti, V. K., (2016), Blockchain platform for industrial Internet of Things. *Journal of Software Engineering and Applications, 9*, 533-546.

Barenji, A. V., Guo, H., Tian, Z., Li, Z., Wang, W. M., & Huang, G. Q., (2018), Blockchain-based cloud manufacturing: decentralization. In Peruzzini et al (Eds.), *Transdisciplinary Engineering Methods for Social Innovation of Industry 4.0.* IOS Press, 1003.

Barenji, A. V., Li. Z., & Wang, W. M., (2018), Blockchain cloud manufacturing: Shop floor and machine level. *Proceedings of Smart SysTech 2018.* 1-6.

Belle, I., (2017), The architecture, engineering and construction industry and blockchain technology. In Ji, G. & Tong, Z. (eds.), *Proceedings of the 2017 National Conference on Digital Technologies in Architectural Education,* 279-284.

Berger, C., Penzenstadler, B., & Drögehorn, O., (2018), On using blockchains for safety-critical systems. *ACM/IEEE 4th International Workshop on Software Engineering for Smart Cyber-Physical Systems,* 30.

Briner, R. B. & Denyer, D., (2012), Systematic review and evidence synthesis as a practice and scholarship tool. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management,* Oxford University Press.

Carreno, R., Aguilar, V., Pacheco, D., Acevedo, M. A., Yu, W., & Acevedo, M. E., (2018), An IoT expert system shell in block-chain technology with ELM as inference engine. *International Journal of Information Technology & Decision Making, 18*(1), 87-104.

Casino, F., Dasaklis, T. H., & Patsakis, C., (2019), A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics, 36,* 55-91.

Chung, K., Yoo, H., Choe, D., & Jung, H., (2019), Blockchain network based mining process for cognitive manufacturing. *Wireless Personal Communications, 105,* 583-597.

Davis, F. D., (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-339.

Dounas, T & Lombardi, D., (2019), A CAD-blockchain integration strategy for distributed validated digital design. *BIM Applications, 1,* 223-230.

Downey, L. X., Bauchot, F., & Röling, J., (2018), Blockchain for business value: A contract and work flow management to reduce disputes pilot project. *IEEE Engineering Management Review, 48*(4), 86.

Drescher, D. (2017). Blockchain Basics. New York: Apress.

Fernández-Caramés, T., & Fraga-Lamas, P., (2019), A review on the application of blockchain to the next generation of cybersecure Industry 4.0 smart factories. *IEEE Access, 7,* 45201-45218.

Fraga-Lamas, P. & Fernández-Caramés, T., (2019), A review on blockchain technologies for an advanced and cyber-resilient automotive industry. *IEEE Access, 7*, 17578-17598.

Goldstein, R., Imbault, F., Usländer, T., & de la Gastine, H., (2018), Fractal production: Reprogramming "Industrie 4.0" around resource and energy efficiency? *2018 IEEE International Conference on Environmental and Electrical Engineering (EEEIC),* 1-5.

Heber, D. & Groll, M. W., (2018), How the blockchain fosters E/E traceability for MBSE and PLM in distributed engineering collaboration. *2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD),* 125-130.

Holler, M., Barth, L., & Fuchs, R., (2019), Trustworthy product lifecycle management using blockchain technology – experience from the automotive ecosystem. In J. Stark (Ed.), *Product Lifecycle Management (Volume 4): The Case Studies, Decision Engineering,* 13.

Holtkemper, D. & Wieninger, S., (2018), Company data on blockchain: A juxtaposition of technological drivers and potential applications. *Proceedings of PICMET '18: Technology Management for Interconnected World,* 1-7.

Isaja, M. & Soldatos, J. K., (2018), Distributed ledger architecture for automation, analytics and simulation in industrial environments. *IFAC PapersOnLine 51-11,* 370-375.

Isaja, M. & Soldatos, J., (2018), Distributed ledger technology for decentralization of manufacturing processes. *2018 IEEE Industrial Cyber-Physical Systems (ICPS),* 696-701.

Innerbichler, J. & Damjanovic-Behrendt, V., (2018), Federated Byzantine Agreement to ensure trustworthiness of digital manufacturing platforms. *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchain for Distributed Systems,* 111.

Kapitonov, A., Berman, I., Bulatov, V., Lonshakov, S., & Krupenkin, A., (2018), Robonomics based on blockchain as a principle of creating smart factories. *Fifth International Conference on Internet of Things: Systems, Management and Security (IoTSMS),* 78-85.

Kasten, J. (2019), Blockchain application to engineering model management. *Proceedings of the Northeast Decision Sciences Annual Conference,* 362-388.

Kennedy, Z. C., Stephenson, D. E., Christ, J. F., Pope, T. R., Arey, B. W., Barrett, C. A., & Warner, M. G., (2017), Enhanced anti-counterfeiting measures for additive manufacturing: coupling lanthanide nanomaterial chemical signatures with blockchain technology. *Journal of Materials Chemistry C, 5,* 9570-9578.

Kobzan, T., Bienderra, A., Schriegel, S., Herbst, T., Müller, T., & Jasperneite, J., (2018), Utilizing blockchain technology in industrial manufacturing with the help of network simulation. *IEEE 16th International Conference on Industrial Informatics (INDIN),* 152-159.
Kremenova, I. & Gajdos, M., (2019), Decentralized networks: The future Internet. *mobile Networks and Applications,* https://doi.org/10.1007/s11036-018-01211-5.

Kundu, D., (2019), Blockchain and trust in a smart city. *Environment and Urbanization ASIA, 10*(1), 31-43.

Lemes, S. & Lemes, L., (2019), Blockchain in distributed CAD environments. In I. Karabegović (Ed.), *NT 2019: New Technologies, Development and Application II,* 25-32.

Li, J., Greenwood, D., & Kassem, M., (2019), Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. *Automation in Construction, 102,* 288-307.

Li, J., Greenwood, D., & Kassem, M., (2019), Blockchain in the construction sector: A socio-technical systems framework for the construction industry. In I. Mutis and T. Hartmann (Eds.), *Advances in Informatics and Computing in Civil and Construction Engineering,* 51.

Li, Z., Barenji, A. V., & Huang, G. Q., (2018), Toward a blockchain cloud manufacturing system as a peer to peer distributed network platform. *Robotics and Computer Integrated Manufacturing, 54,* 133-144.

Li, Z., Liu, L., Barenji, A. V., & Wang, W., (2018), Cloud-based manufacturing blockchain: Secure knowledge sharing for injection mould redesign. *Procedia CIRP 72,* 961-966.

Li, Z., Wang, W. M., Liu, G., Liu, L., He, J., & Huang, G. Q., (2018), Toward open manufacturing: A cross-enterprises knowledge and services exchange framework based on blockchain and edge computing. *Industrial Management & Data Systems, 118*(1), 303-320.

Lin, Y., Petway, J. R., Anthony, J., Mukhtar, H., Liao, S., Chou, C., & Ho, Y., (2017), Blockchain: The evolutionary next step for ICT E-Agriculture. *Environments, 4*, 50.

Liu, J., Jiang, P., & Leng, J., (2017), A framework of credit assurance mechanism for manufacturing services under social contract manufacturing context. *13th IEEE Conference on Automation Science and Engineering (CASE),* 36.

Liu, X. & Cai, S., (2018), An automatic decision-making value system with pharmaceutical manufacturing industry based on blockchain. *11th International Symposium on Computational Intelligence and Design,* 56-60.

Lou, P., Zhu, C., Zhang, X., Jiang, X., & Li, Z., (2018), Research on the cooperative behavior in cloud manufacturing. *AsiaSim 2018,* 241-254.

Mathews, M., Robles, D., & Bowe, B., (2017), BIM + blockchain: A solution to the trust problem in collaboration? *CITA BIM Gathering 2017.*

Miller, D., (2018), Blockchain and the Internet of Things in the industrial sector. *IT Professional, 20*(3), 15-18.

Mohamed, N. & Al-Jaroodi, J., (Applying blockchain in Industry 4.0 applications. *IEEE 9th Annual Computing and Communications Workshop and Conference (CCWC),* 852-858.

Mohamed, N., Al-Jaroodi, J., & Lazarova-Molnar, S., (2019), Leveraging the capabilities of Industry 4.0 for improving energy efficiency in smart factories. *IEEE Access, 7,* 18008-18020.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. www.bitcoin.org/bitcoin.pdf. Accessed 12 July, 2018.

Nuss, M., Puchta, A., & Kunz, M., (2018), Towards blockchain-based identity and access management for Internet of Things in enterprises. In S. Furnell et al (Eds.), *TrustBus 2018,* 167-181.

Oak, R., Jhala, K., & Khare, M., (2018), Smart collaboration mechanism using blockchain technology. *2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud),* 117-121.

Plotnikov, V. & Kuznetsova, V., (2018), The prospects for the use of digital technology "blockchain" in the pharmaceutical market. *MATEC Web of Conferences, 193.*

Papakostas, N., Newell, A., & Hargaden, V., (2019), A novel paradigm for managing the product development process utilizing blockchain technology principles. *CIRP Annals – Manufacturing Technology,* in press.

Petrali, P., Isaja, M., & Soldatos, J. K., (2018), Edge computing and distributed ledger technologies for flexible production lines: A white-appliances industry case. *IFAC PapersOnLine 51-11,* 388-392.

Porter, M. E., (1985). *Competitive Advantage: Creating and Sustaining Superior Performance.* New York: Free Press.

Rimba, P., Tran, A. B., Weber, I., Staples, M., Ponomarev, A., & Xu, X., (2018), Quantifying the cost of distrust: Comparing blockchain and cloud services for business process execution. *Information Systems Frontiers,* 1-19, doi: 10.1007/s10796.018.0876-1.

Sehgal, A., (2017), Blockchain's insurance business implementation. *International Journal of Computer Applications, 174*(3), 32.

Sikorski, J. J., Haughton, J., & Kraft, M., (2017), Blockchain technology in the chemical industry: Machine-to-machine electricity market. *Applied Energy, 195,* 234-246.

Singh, A., Kumar, D., & Hötzel, J., (IoT based information and communication system for enhancing underground mines safety and productivity: Genesis, taxonomy and open issues. *Ad Hoc Networks, 78,* 115-129.

Skwarek, V., (2017), Blockchains as security-enabler for industrial IoT – applications. *Asia Pacific Journal of Innovation and Entrepreneurship, 11*(3), 301-311.

Towne, W. B., (2019), A blockchain-based database for qualified parts. *SAE Technical Paper 2019-01-1343.*

Treblmaier, H., (2018), The impact of the blockchain on the supply chain: A theory-based research framework and a call for action. *Supply Chain Management: An International Journal, 23*(6), 545-559.
Turk, Z. & Klinc, R., (Potentials of blockchain technology for construction management. *Procedia Engineering, 196,* 638-645.

Wang, Y., Han, J. H., & Beynon-Davies, P., (2019), Understanding blockchain technology for future supply chains: A systematic literature review and research agenda. *Supply Chain Management: An International Journal, 24*(1), 62-84.

Yan, Y., Duan, B., Zhong, Y., & Qu, Z., (2017), Blockchain technology in the Internet Plus: The collaborative development of power electronic devices. *IECON 2017 – 43rd Annual Conference of the IEEE Industrial* Electronics *Society,* 922.

Ye, Z., Yin, M., Tang, L., & Jiang, H., (2018), Cup-of-water theory: A review on the interaction of BIM, IoT and blockchain during the whole building lifecycle. *35th International Symposium on Automation and Robotics in Construction (ISARC 2018).*

Yin, S., Bao, J., Zhang, Y., & Huang, X., (2017), M2M security technology of CPS based on blockchains. *Symmetry, 9,* 193.

Yu, C., Zhang, L., Zhao, W., & Zheng, S., (2019), A blockchain-based service composition architecture in cloud manufacturing. *International Journal of Computer Integrated Manufacturing,* DOI: 10.1080/0951192X.2019.1571234

Zhang, F., Liu, M., & Shen, W., (2017), Operation modes of smart factory for high-end equipment manufacturing in the Internet and big data era. *2017 IEEE Interntional Conference on Systems, Man, and Cybernetics,* 152-157.

Zheng, R., Jiang, J., Hao, X., Ren, W., Xiong, F., & Ren, Y., (2019), bcBIM: A blockchain-based big data model for BIM modification audit and provenance in mobile cloud. *Mathematical Problems in Engineering, 2019,* Article ID 5349538.

Zheng, R., Jiang, J., Hao, X., Ren, W., Xiong, F., & Zhu, T., (2019), CaACBIM: A context-aware access control model for BIM. *Information, 10,* 47.

Zissis, D., & Lekkas, D., (2012), Addressing cloud computing security issues. *Future Generation Computer Systems, 28,* 583-592.

DECISION SCIENCES INSTITUTE

Startups as Suppliers and Sources of Innovation: Archetypes of Selection Approaches

Stephan M. Wagner
Department of Management, Technology, and Economics
Swiss Federal Institute of Technology Zurich
Weinbergstrasse 56/58, 8092 Zürich, Switzerland
Email: stwagner@ethz.ch

Stefan Kurpjuweit
Department of Management, Technology, and Economics
Swiss Federal Institute of Technology Zurich
Weinbergstrasse 56/58, 8092 Zürich, Switzerland
Email: skurpjuweit@ethz.ch

Thomas Y. Choi
Department of Supply Chain Management
W. P. Carey School of Business
Arizona State University
Email: thomas.choi@asu.edu

ABSTRACT

Startups are increasingly becoming suppliers to established firms. However, prior research has only studied how buying firms select established firms as suppliers. It remains unclear which processes, tools or organizational approaches firms use when selecting startup firms as suppliers. This inductive, qualitative case study research collects data from 20 established buying firms and investigates how they select startups as suppliers. The results suggest that firms that are willing and able to adapt their selection approach to startups are expected to show a higher selection performance, meaning that they are more likely to select suitable startups as suppliers.

KEYWORDS: New venture supplier, startup, supplier selection, supplier evaluation, open innovation, case study

INTRODUCTION

Under the umbrella of the open innovation paradigm, established firms are tapping into external sources of innovation to increase their innovation output (Chesbrough, 2006). Suppliers and startups are considered particularly attractive innovation partners for established organizations (Azadegan & Dooley, 2010; Weiblen & Chesbrough, 2015). Traditionally, both sources have been studied from different perspectives. While the selection of suitable startups has been studied primarily from the perspective of strategic alliances, venture capital, or corporate incubators and accelerators (e.g., Baum & Silverman, 2004; Aerts, Matthyssens & Vandenbempt, 2007; Moschner et al., 2019), selecting innovative suppliers has been an important stream within the supply chain management literature (e.g., Koufteros, Cheng & Lai, 2007; Oke, Prajogo & Jayaram, 2013). However, scholars and business practice have only recently recognized the potential of merging the two perspectives.

Following several calls for research (Kickul et al., 2011; Shepherd & Patzelt, 2013), scholars have started integrating the supply chain with the entrepreneurship literature by investigating buyer-supplier relationships between established buying firms and new venture suppliers [Following Zaremba et al. (2016), we use the term new venture supplier for startups that became suppliers to established firms.] (Zaremba, Bode & Wagner, 2016; Rottenburger & Kaufmann, 2019). At the same time, leading corporations such as the premium automaker BMW announced that it would be increasing its supplier relationships with startups (Gimmy et al., 2017). We define *new venture suppliers* as young and innovative startups that can bring radical benefits to buying firms through sourcing or co-developing new technologies, products or services.

The first step in building such asymmetric buyer-supplier relationships is the selection of suitable new venture suppliers. Along with buying firms' increasing outsourcing activities in the last decades (Spina et al., 2013), scholars have studied various aspects of the supplier selection process such as its link to overall firm performance (Kannan & Tan, 2002), product quality (Gonzalez, Quesada & Mora Monge, 2004) or the use of evaluation criteria for determining suitable suppliers (Ho, Xu & Dey, 2010). However, these studies do not distinguish established from startup firms as suppliers, even if startups are struggling with legitimacy and are thus less likely to be selected (Stinchcombe, 1965; Brachtendorf, 2016). Furthermore, because of startups' lack of resources, established processes, and their high failure rate, buying firms face substantial challenges evaluating the true capabilities of NVS (Stuart, Hoang & Hybels, 1999; Reuber & Fischer, 2005). In addition, buying firms prefer suppliers with well-recognized certifications and strong track records of high quality and timely delivery (Choi & Hartely, 1996; Kannan & Tan, 2002).

Recent empirical findings indicate that buying firms need to organize the selection process of new venture suppliers differently. For instance, Zaremba et al. (2016) investigated how established buying firms can work with new venture suppliers to achieve desired relationship outcomes. They show that buying firms that demonstrate an innovation orientation during supplier selection tend to buy more from startups and to be more innovative with them. By the same token, a cost orientation during supplier selection reduces the number of successful innovations with startups. In a later study, Zaremba et al. developed the concept of *new venture partnering capability* (*NVPC*), which captures a buying firm's effectiveness in leveraging the potential of new venture suppliers (Zaremba, Bode & Wagner, 2017). They show how buying firms evaluate, develop, and communicate, as well as how they govern the buyer-supplier relationship with startups. Their findings with regard to the selection of suitable new venture suppliers suggest that startups' evaluation requires more time and effort. At the same time, the evaluation criteria must be applied more flexibly to give these young firms a chance against their more mature competitors.

These two studies are a valuable first attempt at exploring this new phenomenon. However, we still know remarkably little about how buying firms actually identify and evaluate this new supplier type, what makes a startup a suitable new venture supplier and which organizational functions are involved in the selection process. A better understanding of how established firms can identify, evaluate, and select suitable new venture suppliers would help to improve the success rates of subsequent collaborations with positive effects on established firms' innovation output and startups' survival rates. This study therefore answers the following research questions: What are the dominant profiles (or archetypes) of the buying firms as they select new venture suppliers? How are these profiles linked to the chance of selecting suitable new venture suppliers?

To answer these questions, we adopt a multiple case study design, which is particularly suitable for generating theory (Eisenhardt, 1989). The findings of this study extend the sparse research in the emerging field of new venture suppliers and contribute to the supplier selection literature. Our analysis led to the development of a typology of three distinct buying firm archetypes which differ (1) in their strategic focus, (2) in the type of startups they are engaging with, (3) the organizational responsibility of the selection process, (4) in their identification strategy, and (5) in their evaluation approach. Based on these findings, we derive a set of propositions that describes how the likelihood of selecting a suitable new venture supplier is connected with specific adaptations that the case firms made regarding their selection approaches. Moreover, we provide an initial description of a suitable new venture supplier. The archetypes offer insights into how and why buying firms select startups as suppliers and help firms classify and benchmark their selection processes.

LITERATURE REVIEW

New Venture Suppliers as a Source of Competitive Advantage

From a resource-based perspective, firms can achieve competitive advantage by combining their valuable, rare, imperfectly imitable, and non-substitutable resources with the complementary assets of other organizations (Barney, 1991; Dyer & Singh, 1998). Some scholars even consider resource complementarity as the most important driver for interorganizational collaboration (Dyer, Singh & Hesterly, 2018). New venture suppliers offer vast opportunities for established firms to complement their existing resources bases. Perhaps, most importantly, startups are a source of radical innovation and a driver of technological change (Ireland, Hitt & Sirmon, 2003; Song & Di Benedetto, 2008). In contrast, a buying firm's incumbent supply base might be a source of incremental and routinized innovation since buying firms initiate numerous innovations on their own, which rarely results in truly radical improvements (Wagner & Bode, 2014).

Startups also differ from established firms in terms of worse access to financial resources and a lack of management and manufacturing capabilities (Hitt et al., 2001; Terjesen, Patel & Covin, 2011). However, they are often recognized as excellent problem solvers and product developers for several reasons. First, their organizational prerequisites such as their ability to make quick decisions without lengthy approval procedures create an advantage regarding speed and flexibility compared to mature organizations (Hogenhuis, van den Hende & Hultink, 2016). Second, their way to communicate is more direct, informal, and more frequent than that of established firms facilitating collaborative activities (Das & He, 2006). Finally, young firms must develop their innovations quickly to reach market maturity as they risk exhausting their funding. This risk puts their focus on fast market introduction in order to generate first cash flows (Poutziouris, 2003). In sum, these characteristics make them ideal partners for large firms, particularly for accelerating product development projects.

For startups, strong and partnership-like supply chain relationships can yield substantial advantages. As Arend and Wisner (2005, p. 405) point out: "To the entrepreneur, SCM has many potential benefits […] (e.g.) the ability to leverage its scalable competences in a cooperative network through fast and feasible access to complementary partner assets." For example, established firms can contribute tangible and intangible resources that startups can access through participating in their customers' supplier development programs (Zaremba et al., 2016). Moreover, relationships with buying firms that offer specifications, feedback, and finally purchase orders help startups generate cash flows and gain market legitimacy (Bhalla & Terjesen, 2013; Gimmy et al., 2017). In fact, some studies support these theoretical arguments. For instance, Dutta and Hora (2017) find that startups' downstream alliances (to industry incumbents) increase their invention and commercialization success and Terjesen et al. (2011) show that the value of a startup's operational capabilities depends on its operating environment and the type of alliances it forms. Alliances with a diverse set of customers and suppliers, which are geographically dispersed, seem to have a positive effect on the value of startups' operational capabilities.

New Venture Supplier Identification

The traditional supplier selection literature often describes supplier selection as an optimization problem, in which buying firms select the most suitable supplier out of a given set of available firms without the necessity of identifying them (e.g., Kim & Wagner, 2012). The identification process typically relates to the ability to distinguish innovative from less-innovative suppliers. For example, Pulles, Veldman and Schiele (2014) emphasize that buying firms should search for suppliers with a collaborative attitude, high levels of professionalism (i.e., skills and expertise), and deep technical knowledge. Buying firms use a variety of strategies to engage with potentially innovative suppliers, such as supplier days, supplier workshops, or supplier competitions (Langner & Seidel, 2009; Schiele, 2010). Nonetheless, also for these approaches, buying firms need to be aware of a set of potential suppliers they can invite. Startups, however, are probably not part in this pool of firms, as many established firms lack experience in collaborating with them (Schaettgen & Mur, 2016) and thus miss ties to startup ecosystems. Moreover, startups often do not belong to a specific industry since neither the potential applications for their technology, nor their potential customers are well-defined (Chesbrough, 2006). Thus, buying firms must search for attractive startups in different supply markets.

The open innovation literature describes the ways in which firms can search for external innovations. Companies might hire technology scouts, use intermediaries (e.g., consultants), or search the internet for attractive startups (Ili, Albers & Miller, 2010). Using the internet to find potential innovation partners has decreased search costs, as companies now can find attractive startups in communities, blogs, or via crowdsourcing platforms (West & Bogers, 2014). While screening has become more affordable, many companies face the challenge of having to process large numbers of identified startups. Siemens, for instance, identifies about 1,200 startups every year (Weiblen & Chesbrough, 2015). As processing and evaluating such large numbers is time-consuming, West and Bogers (2014) conclude that firms must find an optimal level of search.

In contrast, there are initiatives where buying firms passively identify startups through web-based innovation platforms where suppliers can submit their ideas or products (Kurpjuweit, Reinerth & Wagner, 2018). For startups, these web-based platforms are usually the interface to outside-in programs with the objective to make the "existing startups' technology accessible and useful for the sponsoring corporation" (Weiblen & Chesbrough, 2015, p. 72). According to

Weiblen and Chesbrough (2015), these new engagement models differ from traditional models such as corporate venture capital by the absence of the corporation's equity involvement. For building buyer-supplier relationships with startups, these traditional models have had problems with the innovation transfer in the past and might be unsuitable as their equity involvement may create difficulties for startups to engage with customers that are competing in the same industry as their equity provider (Gimmy et al., 2017). Therefore, the identification of potential new venture suppliers might rely on non-equity based models such as scouting and screening or web-based innovation platforms. However, both mechanisms require the buying firm to view the identification process as an integral part of the supplier selection problem which distinguishes the selection of new venture suppliers from the selection of mature suppliers.

New Venture Supplier Evaluation

Evaluating suppliers against predefined criteria is one of the major tasks in supplier selection contests. Based on a supplier's recent history, buying firms apply evaluation criteria such as price, quality, flexibility or consistent delivery to determine the supplier's qualifications (Choi & Hartely, 1996; Ho et al., 2010). However, startups lack a proven track record of high performance which pose challenges for buying firms to apply these criteria properly. This lack of history is often referred to as *legitimacy*, defined as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p. 574). Missing legitimacy constitutes a major disadvantage for startups in supplier selection (Brachtendorf, 2016). In addition, startups' organizational limitations and their lack of supplier resources and capabilities (Das & He, 2006) might pose substantial challenges in later collaboration stages that require project management and manufacturing capabilities (Hogenhuis et al., 2016).

From the buying firm's perspective, this liability of newness creates considerable uncertainty about a startup's true supplier performance potential, decreasing its chance of being selected as a supplier, especially when competing against an established firm (Zimmerman & Zeitz, 2002). Zaremba et al. (2016, p. 162) even point out that "new ventures should be regarded as a distinct subset of suppliers in the marketplace." Therefore, buying firms must adjust their supplier selection process in response to the specific requirements of startups. According to Zaremba et al. (2017), buying firms that have integrated the capabilities of new venture suppliers, examined these startups more closely by performing additional evaluations and by allocating more time and human resources to the selection process. However, once a startup has shown that it possesses specific base capabilities, the standard evaluation criteria were relaxed and applied more flexibly. For example, buying firms required the startups to present fewer references or certifications.

In general, the literature indicates that buying firms must adapt their supplier selection process when considering new venture suppliers. This study, therefore, aims to shed light on how different archetypes of buying firms tailor their selection process to startups, and how these adaptations affect the likelihood that a suitable new venture supplier will be chosen.

METHODOLOGY

The unexplored phenomenon of buying firms selecting startups as suppliers demands an inductive case study approach (Barratt, Choi & Li, 2011). This research design fits best with the nascent stage of the research in the field of new venture suppliers (Edmondson & McManus, 2007) and is well suited for generating new theory (Glaser & Strauss, 1967; Ketokivi & Choi, 2014) – especially for developing typologies and archetypal classifications (e.g., Bode, Hübner & Wagner, 2014; Pauwels et al., 2016).

Research Setting and Case Selection

The empirical setting of this research is established buying firms which have recently selected at least one startup as their supplier. We refined our new venture supplier definition from the introduction by setting an age limit of six years. This is consistent with the entrepreneurship literature (Song et al., 2008) and has been used in other investigations of new venture suppliers (Zaremba et al., 2017). We restricted the sampling process to buying firms which have selected startups that offered a tangible product that was either integrated into the buying firms' products or intended to improve the value creation process of a buying firm's core business (e.g., new production machines). The technology had have been developed by the startup and required at least some form of customization (not solely distribution). We focused on the selection of manufacturing ventures as we expected from this setting to be able to observe selection decisions that are particularly relevant for the buying organizations. Furthermore, we analyzed supplier selection decisions that cannot be easily reversed. Relationships with component suppliers are often difficult to terminate, as they can involve relationship-specific investments and highly specific product interfaces which require domain knowledge so that buying firms hope to build long-term relationships with those suppliers to safeguard their investments (Wagner & Bode, 2014).

Both in academic research and in business practice, sourcing from manufacturing ventures is relatively new. In order to identify buying firms that have already selected new venture suppliers, we screened German-speaking startup blogs and magazines. We then contacted established firms that formed asymmetric collaborations that fit our criteria. In the contact email, we described the purpose of the study and requested the name of the person within the organization who was in the best position to talk to us about an example of the firm's selection process. Then, we had a brief conversation over phone or email over whether or not the example fit our study setting. Informants who provided useful examples were then asked for an expert interview. At the end of each interview, we asked each interviewee whether he or she knew of another case from his or her professional network that would be appropriate for our research (i.e., snowball sampling).

The firms that we considered for our sample generate annual sales exceeding USD 100 million and have been in business for several decades. In fact, most firms were founded more than 80 years ago; only one firm is substantially younger (*3D Printing*). However, this firm develops, manufactures and distributes selective laser sintering (SLS) machines, which is still a fledgling but increasingly important business (Durach, Kurpjuweit & Wagner, 2017). In addition, with more than 1000 employees and sales exceeding USD 300 million, *3D Printing* is one of the leading organizations in additive manufacturing today.

By setting the USD 100 million threshold for sales, firm size within the sample naturally varies. Since firm size seems to affect a firm's ability to partner effectively with startups (Zaremba et al., 2017), by including established firms with different sizes, we expected to bring more variation to the phenomenon, which usually generates deeper insights (Flick, 2014). Table 1 summarizes the sample demographics. Notably, the informants' organizational functions vary considerably among the case firms, indicating that procurement is not the dominant function of the selection process.

Table 1: Overview of Case Firms

#	Company Code	Industry Sector	Employees	Revenues (in EUR)	Firm age (in years)	Country	Informant(s)	Archival data (pages)
1	Adhesive	Consumer & Industrial Goods	10k-50k	>20bn	>100	GER	Dir. new business development	54
2	Automation	Automation, Machine Building	10k-50k	1-5bn	>80	GER	Head of R&D strategy	36

3	Automotive	Automotive OEM	>100k	>50bn	>100	GER	Head of startup relations	92
4	Biopharma	Biopharma & Lab technology	1k-10k	1-5bn	>100	GER	Process development consultant	-
5	Car	Automotive OEM	10k-50k	5-20bn	>100	CZ	Innovation specialist	61
6	Electronics	Industrial Goods	>100k	>50bn	>100	GER	Head of project procurement	129
7	Excavator	Construction Machinery	1k -10k	0.1-1bn	>50	GER	Head of divisional R&D	48
8	Filtration	Industrial Filtration	1k-10k	1-5bn	>80	GER	Head of divisional SCM	-
9	Food	Retail, Food Production	>100k	>50bn	>100	GER	Head of IT	-
10	Lineartech	Machine Building	1k-10k	1-5bn	>80	CH	CEO	
11	Logistics	Logistics & Transportation	>100k	>20bn	>80	GER	Corporate strategy	141
12	Piping	Piping systems	10k-50k	1-5bn	>100	CH	Head of strategy & M&A	61
13	Medtech	Medical Technology	100-1k	0.1-1bn	>40	CH	Head of SCM	-
14	Pharma	Pharmaceutical industry	>100k	5-20bn	>100	GER	Business development manager	28
15	Robotics	Robotics, Energy Technology	>100k	>20bn	>100	CH	(1) Corporate R&D scientist (2) Business unit head of SCM (3) Global SCM manager	14
16	Rollercoaster	Transport Systems	100-1k	0.1-1bn	>50	CH	CEO	43
17	Security	Building technologies	10k-50k	1-5bn	>100	CH	CPO	-
18	Smarthome	Building Technology	100-1k	0.1-1bn	>50	A	CEO	-
19	Transportation	Transportation	10k-50k	5-20bn	>100	CH	(1) Former head of startup relations (2) Head of startup relations (3) Startup specialist (4) Purchasing manager (5) Purchasing innovation manager	88
20	3D Printing	Additive Manufacturing	1k-10k	0.1-1bn	>25	GER	Business development manager	51

Data Collection

For each case firm, we conducted at least one semi-structured interview with the manager, who was responsible for the selection process. We used a semi-structured interview guide with open-ended questions that let the interviewees reflect on a specific example (see Appendix). We began by asking the informant to refer to the example they had mentioned in the phone call or email prior to the interview. In some cases, we changed the example if the informant came up with a more suitable one at the beginning of the interview. During the interviews, we tried to determine how the case firms identified, evaluated and chose this specific startup, and whether the selection process differed in relation to established or other new venture suppliers. We also asked how the case firms managed the relationship after the selection, and how successful the collaboration has been so far, to understand the impact of the selection process on the buyer-supplier relationship.

In the second part of the interview, we asked more general questions about the case firms' motivation to engage with new venture suppliers, how risky the selection process was, and how they would characterize a suitable new venture supplier. While the core topics of the interview guide remained consistent, we adapted individual questions during the course of data collection and added firm-specific questions for each interview (Patton, 2015). After each interview, we took notes to capture first impressions and thoughts (Miles & Huberman, 1994).

The interviews took place between February and November 2016; each lasted an average of 60 minutes. All informants agreed to be audio recorded as long as we treated the collected data confidentially and guaranteed their anonymity. We transcribed the recorded interviews, resulting in 280 pages of text stored in a database to which all researchers had access (NVivo 10). One of the authors translated the quotes presented in this manuscript from German to English while a research assistant checked the translated quotes for semantic correctness. A professional copy editor then checked for grammatical correctness and for comprehensibility. We concluded the data collection when new interviews generated only few additional insights (Glaser & Strauss, 1967; Flick 2014). In total, we conducted 26 interviews with 20 case firms. We carried them out in person whenever the proximity and availability of the informants permitted (18 out of 26).

Since especially in large firms, many people participate in the selection process, we triangulated the information provided by our informants by conducting additional interviews within the buying firms' organizations. Two large case firms (*Robotics* and *Transportation*) granted us such access. The remaining informants of the other large companies in our sample were confident that they had provided sufficient details about the selection process and were therefore unwilling to provide access to their colleagues.

To complement and triangulate our interview data, we collected additional information about the case companies' collaborative activities with startups. Specifically, we collected press releases from the firms' websites and newspaper articles using the WISO database, one of the largest databases for magazine and newspaper articles published in German. We limited our search to the last five years and collected only press releases and newspaper articles in which the word "start-up" appeared. This procedure resulted in 189 pages of press releases and 616 pages of articles for analysis. In this article, quotes from these archival sources are marked "AD."

Data Analysis

We analyzed and coded the interviews simultaneously with the data collection (Glaser & Strauss, 1967). We began the data analysis by writing individual case descriptions of each firm's typical selection approach and how it differs from selecting established firms. We also prepared a short version of the case description including the key learnings of each case which we sent to our interviewees. We discussed their feedback and included it in our analysis.

After analyzing the cases individually, we moved on to a cross-case analysis in order to investigate the variations among cases. To this end, we used various case displays, but we also compared cases pairwise to identify similarities and differences (Eisenhardt, 1989). In a first step, we inductively generated first-order concepts (codes) without any predefined coding structure. Thereby, we tried to stay as close as possible to the terminology of our informants, in order to generate truly informant-centric codes (Gioia, Corley & Hamilton, 2013). For example, we generated codes such as "close future technological gaps" something that had been explicitly stated by the interviewees. Two researchers participated in this coding process. One of the authors and a research assistant coded the transcripts independently to avoid biased coding and ensure a free interpretation of the data. A code could be only one sentence or a whole paragraph, and the same sentence could belong to different codes. This process resulted in dozens of codes generated by each researcher. During data collection and analysis, we constantly compared the codes generated by each researcher. We discussed differences, and merged codes when necessary. For example, several codes had the same meaning but slightly different names (e.g., startup stage vs. startup maturity) or were based on different levels of analysis (e.g., passive identification vs. internet-based application form). We drastically reduced the total number of codes to ensure alignment with our research questions, which resulted in a final list of 13 codes.

After coding all interviews with this joint coding structure, we calculated the inter-rater reliability. With a Cohen's κ of .67, the two researchers arrived a substantial agreement (Landis & Koch, 1977). Then, both researchers discussed all nodes which were zero or below and adjusted them accordingly. This procedure increased of the inter-rater reliability to .81, reflecting an "almost perfect" agreement (Landis & Koch, 1977).

In a second step, we were able to aggregate the generated 13 codes to five design themes (second-order themes): *strategic focus*, *startup type*, *organization*, *identification*, and *evaluation*. In contrast to our first-order codes, they are more abstract and theory-centric (Gioia et al., 2013). In a final step, we aggregated the five design themes under three overarching theoretical dimensions (i.e., archetypes). We provide a detailed explanation of this procedure in our analysis and results section. Table 3 provides an overview of the resulting data structure.

ANALYSIS AND RESULTS

This section presents the findings of our cross-case analysis. We first provide evidence of how a suitable new venture supplier can be characterized. Then, we inductively develop five design themes by comparing the selection approaches of the case firms. Finally, we use these themes to build a typology of buying firms selecting new venture suppliers.

What is a Suitable New Venture Supplier?

Before we can answer the question of how the case firms select new venture suppliers, we must clarify the objective of the selection process. In general, buying firms want to select suitable suppliers based on pre-defined evaluation criteria, such as cost, delivery, or quality performance (e.g., Choi & Hartely, 1996; Ho et al., 2010). In contrast to the selection of established suppliers, it is, however, less clear which characteristics a suitable new venture supplier should and should not exhibit. Since no single case firm could give us an explicit definition of a suitable new venture supplier, we analyzed the features our informants were looking at when making their selection decisions. By asking them why they were selecting a specific new venture supplier, we could identify four categories that distinguish suitable from unsuitable new venture suppliers. We provide representative quotes for each category in Table 2 and describe each category.

Table 2: Representative Quotes Supporting the Definition of Suitable New Venture Suppliers

	Related Questions	Suitable New Venture Supplier	Inadequate New Venture Supplier
Quality Fit	• Does a startup possess above-average resources and capabilities? • Do startups' resources and capabilities fit to a buying firm's requirements for a specific use case?	"[A high-quality startup] has already passed a well-known accelerator or is already funded by venture capital. Before that, we do not take the startups because we do not want to do this due diligence. We do not want to worry about whether the team is good enough, whether they have their investing agreements […] the whole legal basis of the company. I do not want to do that. Others do this better." (Automotive)	"If a startup has not yet raised any money with its idea, then there will be a reason for that […]. Does the team have strategically relevant competencies? For example, an IT startup without software engineers [...]. This is not a K.O. criterion, but it would become very difficult." (Logistics) "In the end, it was a very negative experience for us, because the startup did not survive [because it ran out of funding]." (Robotics)
Strategic Fit	• Does the technology fall into a buying firm's strategic innovation roadmap?	"Is the startup's technology applicable for measuring the tire pressure or is it interesting for autonomous driving? Autonomous driving would be better in this case." (Automotive) "I strongly believe that there needs to be a strategic link. […] The strategic link is extremely important […] and it could be that a startup has a technology that complements our own technology	"If a startup supplies us with a critical component, then the question is whether we would collaborate at all, because we should produce such components by ourselves." (Automation) "[The startup] only delivers a single part, because we aim to keep the core know-how with us." (Automation)

		portfolio, which was the case with [startup name].” (Adhesive)	
Technological Fit	• Is the technology better than everything comparable in the market?	“There was a technical due diligence, where different benchmark parts were produced to compare the technologies, and it was quickly clear that [startup name] was already leading in terms of quality.” (3D Printing) “Startups are always relevant when it comes to specific, new know-how. New know-how where there are no resources on the market and no established players yet.” (Piping)	“There were cases where startups questioned the existence of the second law of thermodynamics. If such statements come, you can sort them out. For me, that would be such a technical KO criterion. In the meantime, they have also gone bankrupt.” (Robotics)
Market Fit	• Is there someone willing to pay for a startup’s technology inside the buying firm?	“The product must bring benefits, which means you have to recognize, where you gain something. [The product] needs to address someone’s current or long-term pain point.” (Automation) “It is the best for us when a startup already has concrete visions for the application of their technologies. We have just received an application for freight transport that is very detailed regarding the problem of train composition. Of course, this is great for us because we can immediately see the USP very clearly.” (Logistics)	If the technology is not interesting for anyone within our organization, then we do not have to talk about the other things.” (Robotics)

Quality fit. This attribute stands for the quality of startups’ resources and capabilities compared to other startups in the market. According to our interviewees, a startup’s quality can be assessed independently from its product since some startups are better managed, have better processes, possess more know-how, or have acquired more funding than others. Consequently, these startups are expected to have higher survival rates and are easier to manage as suppliers. Besides seeking startups with above-average qualities, suitable new venture suppliers should possess resources and capabilities that fit the potential use case, on which they will be collaborating, as the quote of *Logistics* in Table 2 illustrates.

Strategic fit. The second attribute refers to the fit between a startup’s technology and the buying firm’s innovation strategy. As the quote of *Automotive* in Table 2 shows, autonomous driving is a key technological area for this case firm. *Automotive*, thus, aims to engage with startups that help to strengthen its competitive position in this area which seems to be particularly important for the entire automotive industry today. Our data further suggest that although some technologies lie within the scope of buying firms’ innovation roadmaps, firms such as *Automation* refuse to collaborate in what they consider as their core technological areas. Both firms would always prefer an acquisition to a buyer-supplier relationship as they both prefer engaging with new venture suppliers in more distant technological fields. Nevertheless, both examples show that established firms align their selection decision with their innovation strategy in order to achieve a strategic fit.

Technological fit. The third category suggests that a startup’s product must fit the technological requirements of the buying firm. Often this implies that the offered technology must be substantially better than comparable products in the market, or that a startup offers something so new that buying firms cannot compare it with technologies provided by established suppliers. According to our interviewees, startups’ selection chances become very low, if comparable products from established suppliers are available.

Market fit. According to our data, high-quality startups which offer technologically sophisticated and strategically important products to buying firms do not always find a paying customer within the buying firm's business units for their technologies. Since neither startup relationship managers nor the innovation management department have the budget and the decision making power to purchase startups' technologies, startups need to convince business unit leaders, purchasing, or product managers of the advantages of their technology. Hence, a suitable new venture supplier is one whose technology finds a paying internal customer. Clarifying whether a real demand for a technology in a business unit exists is, therefore, an integral part of the selection process.

New venture suppliers exhibiting all four of these attributes are expected to have better chances of survival, as their resource bases and their value propositions are highly attractive for potential customers. Accordingly, we define a suitable new venture supplier as a high-quality startup that offers technologies, products or services that fit the buying firm's strategic, technological and market needs.

Comparing the Selection Approaches of Buying Firms

In this section, we compare the selection approaches of our 20 case firms. We inductively develop five design themes that describe how buying firms selecting new venture suppliers. Specifically, the selection processes of the case firms differ with regard to a buying firm's (1) strategic focus, (2) type of the selected startups, (3) organizational approach, (4) identification efforts, and (5) its evaluation approach. These five elements emerged from our qualitative data during the coding process (Table 3). To estimate how they differ across the case firms, we assign intensities (+, ++, +++) for each element and each case firm. In the following, we explain the five themes and the way in which we assigned the intensities.

Table 3: Representative Quotes Supporting the Data Structure

Second-Order Themes	First-Order Concepts	Representative Quotes
Strategic Focus	Solve a current technical problem	"In our industry, you only cooperate with a startup, if there is a customer request that is not easy to solve." (Excavator)
	Close future technological gaps (Innovation focus)	"There are many startups with which we have only loose collaborations in order to evaluate the technology. These collaborations are not based on a specific product or a decision that we need this technology now, but rather we try to find out how the technological landscape looks like." (Robotics)
	Profit from startups' innovations and capabilities	"Depending on where [the startups] come from, they have different intellectual approaches and opportunities which we do not have. For example, at some universities, there are highly specialized labs and they can try things out and accept risks which we cannot." (Adhesive)
Startup Type	Startups' stage	**Early stage** "We had the idea to do something with early-stage startups. We have collaborated with many later stage startups but what about all these startups with maybe five employees? These early-stage startups are even more interesting for a company like ours in the sense of innovation impact." (Automotive) **Later stage** A startup that contacts us and says it has an entirely new product has only a 0.01 percent chance to become selected." (Food)
	Distance to buying firm's knowledge base	**Distant to own knowledge base** "I am open to all good ideas. We do not restrict ourselves, because if we do, we miss the opportunity to find something exciting." (Logistics) **Close to own knowledge base** „We are only interested in material science startups." (Adhesive)

Organization	No clear responsibility	"Either strategic procurement or the research group that is interested in the technology." (Excavator)
	Responsibility assigned to existing department	"We have alerted our sales departments to keep us informed of new startups in the market. The intention is that all roads lead to our department." (Biopharma)
	Dedicated startup department	"In 2016 we established an independent unit designed to partner with […] start-ups with innovative ideas for shaping the future of electrification, automation and digitalization, and thereby turn those ideas into viable businesses" (Electronics; AD).
Identification	Active Identification	No active identification "No, we do not screen the market." (Smarthome) Active identification "There is this search and evaluation team […] This means the colleagues go to all the fairs out there and look out who has something to offer." (Pharma)
	Passive Identification	**No passive identification process** "Depending on whom they [the startups] are contacting, they can be lucky or unlucky. Sometimes they contact developers who know only their area, but if they contact people who have an overview, these people would then contact our corporate research department, and then it would reach me. This is based on an internal network without a formal process behind it. However, it works best if you contact the right people directly." (Robotics) **Passive identification process available** "The application form is a way to efficiently scan through startup applications." (Automotive)
Evaluation	Process	**No established evaluation process** "This startup contacted our HR manager, so we invited the founders to present their ideas. They were very convincing, so we decided just to try it out." (Piping) **Established evaluation process available** "We have developed a catalog of 20 to 30 criteria. This is like a blind tasting. Each of us evaluates the startups independently, then I merge the results, and we make an analysis. […] If a startup gets a poor evaluation, against our initial expectations, we ask why and we discuss it." (Logistics)
	Criteria	**No startup specific criteria** "We assess which quality management system they have, if the product documentation is available if the registrations are all correct […] We evaluate the typical things." (Medtech) **Startup specific criteria** "We evaluate a startup's team, its technology, innovation, the strategic relevance for us, the financial conditions of the startup […]" (Logistics)
	Functions involved	**No expert involvement from other functions** They [the startup and the potential user] typically meet somewhere and figure out whether the product would be attractive. Then this manager checks out whether it makes sense for him, so that he could profit from it and then it goes its way." (Piping) **Experts from other functions involved** "Of course we are working closely with people from R&D and the business unit for the technical and commercial evaluation. We will never collaborate with a startup if there is not one business unit which says: good idea! We are going to support you." (Adhesive)

Strategic focus. The motivation for firms' selection of startups as suppliers differs across the case firms. Some firms started buyer-supplier relationships with startups because they had a specific (technical) problem for which they could neither find any established supplier nor had the expertise in-house. We assigned buying firms with this narrow strategic focus with

"+." Other firms adopt a more forward-looking approach and develop technology roadmaps from which they derive the areas in which they aim to engage with startups. Such companies clearly seek to profit from startups' innovations and consider them as a potential source of external innovation (++). The third type of firm emphasizes the particular strengths of startups. Besides innovativeness, they aim to tap into startups' special capabilities such as their speed, and flexibility, their ability to take the risks that established firms will not, or they hope to be able to bring some of the younger firm's entrepreneurial spirit to their own corporate cultures (+++).

New venture type. The second design element refers to the profile of the selected startups. Our analysis reveals two distinctive features of the selected startups: maturity and distance to the buying firm's knowledge base. Both factors determine a new venture supplier's risk-return profile. Early-stage ventures from a distant knowledge base can be expected to have a stronger innovation impact than later-stage ventures from a buying firm's own knowledge base. The reason is that when engaging with the latter type, it is more likely that competitors have already adopted the technology and that such innovations tend to be more incremental than innovations from completely different areas. We assign buying firms that put more emphasis on searching for early-stage startups from distant knowledge fields with "+++." The other extreme comprises buying companies that prefer later-stage ventures from the own knowledge base (+). Finally, firms in an intermediate position are assigned "++."

Organization. This category relates to the organizational structures with which buying firms approach the selection of new venture suppliers. These approaches range from organizations without any defined responsibilities for startups (+), to organizations with defined responsibilities, embedded in existing departments (++), to firms that have established dedicated startup departments sitting between the corporates and the startups, thus facilitating the engagement process (+++).

Identification. The most renowned firms constantly receive so many applications from startups that processing them poses a challenge. We describe this way of identifying startups as *passive identification* because the buying firm becomes identified by startups. However, a sole passive approach is unsatisfactory for most buying firms. Therefore, many firms also search for new venture suppliers, which we term *active identification*. In contrast to selecting established firms as suppliers, there is no pool of former suppliers to which buying firms can send a request for proposal, which makes this search process difficult. Overall, the intensity and procedure of the identification activities vary across the case firms. While some firms do not identify any startups at all (+), other firms follow either an active or a passive identification strategy (++). The third group of firms identifies startups both actively and passively (+++).

Evaluation. This category contains all evaluation efforts a buying firm makes to assess a startup before the relationship begins. The findings indicate that firms differ primarily in the procedures they follow and the evaluation criteria they apply. The two extremes range from having no processes or any predefined criteria to structured multi-stage evaluation approaches involving different departments and experts. Consequently, companies that do not evaluate any startups at all or that follow a completely unstructured evaluation process are assigned "+." We assign companies "+++," if they have implemented a structured evaluation process, assess startups according to predefined startup-specific evaluation criteria and regularly involve experts in the evaluation process. If at least one of the three evaluation constructs is available (established evaluation process, startup-specific criteria, or involvement of different functions), we assign these firms "++."

Selection Archetypes of Buying Firms

Based on these observed differences and similarities in the five design themes, we were able to assign all case firms to one of three overarching theoretical dimensions (i.e., archetypes): (1)

Skeptical Buyers, (2) *Realistic Adapters*, and (3) *Systematic Selectors*. The assigned archetype is based upon the average score of the five elements and a discussion among the authors (Table 4). In the following, we will outline the key characteristics of the archetypes, which are summarized in Table 5.

Table 4: Assigned Intensities per Case Firm

Company Code	Strategic Focus	Startup Type	Organization	Identification	Evaluation	Resulting Archetype[a]
Adhesive	+++	++	+++	++	++	RA
Automation	+++	++	++	++	++	RA
Automotive	+++	+++	+++	+++	+++	SS
Biopharma	++	+	++	++	++	RA
Car	++	++	++	++	++	RA
Electronics	++	++	+++	+++	++	RA
Excavator	+	++	+	+	++	SB
Filtration	++	++	+	+	+	SB
Food	+	+	+	+	+	SB
Lineartech	+	++	+	++	+	SB
Logistics	+++	+++	+++	+++	+++	SS
Piping	+	++	+	++	+	SB
Medtech	++	++	++	++	++	RA
Pharma	+++	++	+++	+++	++	SS
Robotics	++	+++	++	++	++	RA
Rollercoaster	+	+	+	+	+	SB
Security	+	++	+	++	+	SB
Smarthome	+	++	+	+	+	SB
Transportation	+++	++	+++	+++	+++	SS
3D Printing	+++	+++	+++	++	++	SS

Note: [a] “SB” = Skeptical Buyer, “RA”= Realistic Adapter, “SS” = Systematic Selector.

Table 5: Overview of Different Archetypes of Buying Firms Selecting New Venture Suppliers

	Skeptical Buyers	Realistic Adapters	Systematic Selectors
Theme	*Do not adjust their supplier selection process to startups.*	*Do partially adjust their supplier selection process to startups.*	*Do fully adjust their supplier selection process to startups.*
Strategic Focus	• Select startups to solve specific business problems • Startups are often the only option for a specific problem	• Aim to close future technological gaps in their roadmaps • Startups are one potential source of external innovation	• Aim to get access to innovations, which could not have come from own R&D or other external sources • Aim to get access to startups’ unique capabilities
New Venture Type	• Later stage startups close to own knowledge base • Rather incremental innovations	• Medium position between Skeptical Buyers and Systematic Selectors	• Early stage startups from more distant knowledge bases • Rather radical innovations
Organization	• No clear responsibilities • Responsibility highly depends on the first contact person	• Responsibility is assigned to an existing department or manager	• Dedicated startup department that manages the selection process
Identification	• No identification efforts	• Focus on active identification, processes for passive identification are missing	• Focus on active and passive identification
Evaluation	• Completely unstructured evaluation process • Startups are evaluated based on the intuition of managers who established the contact	• Somewhat structured evaluation process • Dedicated team/manager evaluates startups according to relaxed standard criteria • Involve experts from other departments in startup evaluation	• Structured evaluation process with several stages and predefined startup-specific evaluation criteria

Skeptical Buyers. Firms of the first archetype use a highly unstructured approach when selecting new venture suppliers. *Skeptical Buyers* do not pursue the strategic objective to engage with startups in the first place, they engage with them by accident or because no other established supplier was available. Hence, *Skeptical Buyers* neither actively nor passively identify startups. For *Skeptical Buyers*, the risks of asymmetric collaborations often outweigh the potential benefits. They are particularly risk-averse and prefer sourcing only noncritical technologies from new venture suppliers. One reason for this is that these firms do not trust these startups to be reliable suppliers. As one informant states:

> We had serious internal discussions about this topic [supply risk]. Especially that [the startup] has no experience in managing production processes. (Excavator)

The lack of trust can also result from the lack of experience in selecting new venture suppliers; firms of this archetype do not assign clear responsibilities for startups. However, *Skeptical Buyers* do not use their traditional supplier selection process either. Only *Excavator* once invited a startup to make an offer for a new steering unit for its construction machines. This startup had to compete against several established suppliers, all of which offered less innovative products but met the remaining evaluation criteria (e.g., financial stability, certifications). In the end, the buying firm selected the startup because a senior R&D manager convinced his CEO of the advantages of the startup's unique technology. Other case firms do not use their traditional supplier selection processes as they know that they are not suitable for selecting new venture suppliers:

> […] as a startup, being able to participate in tenders is almost impossible. That is why we are trying to keep them away from this topic as long as possible. (Transportation)

Since *Skeptical Buyers* neither use their traditional selection processes, nor establish any new processes, they typically assess startups without utilizing any predefined criteria. Instead, senior managers decide whether a startup is a suitable supplier or not. These managers frequently rely on their intuition, so they often start collaborations simple because they like an idea or the founders of the venture. Hence, it is more an ad-hoc decision than a well-considered choice.

Realistic Adapters. The innovation strategy of the firms we assigned to the second archetype emphasizes open innovation, in which collaborating with startups is considered a potential source of external innovation. From their innovation strategy, these companies derive technological areas that may become important in the future. Hence, these firms base their engagement with new venture suppliers primarily on the motivation to close technological gaps and to access startups' innovativeness. As one of our informants explained:

> Although we offer a lot of different surface coatings, we are not able to produce this specific one. However, this coating lies within a relevant technological area for us which is why the collaboration is a useful completion for our portfolio. (Adhesive)

From an organizational perspective, *Realistic Adapters* define clear responsibilities for startups within their organizational structures. For example, *Medtech* has appointed a business developer to be responsible for startups, while *Robotics* assigned the startup competence to its strategic R&D department. These dedicated managers or teams carry out this new responsibility in addition to their daily business. They are keen to be recognized as the principal startup contact within their organizations so that all relevant information is gathered in one place. The designated managers orchestrate all identification efforts, whereby firms of this archetype emphasize identification activities such as going to fairs or conferences, searching the internet and most importantly, using their industry network to find attractive new venture suppliers. These companies also realized that they can screen their industry by leveraging internal departments with external contacts such as sales, procurement or local R&D:

> We have alerted our sales departments to keep us informed of new startups in the market. The intention is that all roads lead to our department. (Biopharma)

For *Realistic Adapters*, evaluating new venture suppliers is a matter of teamwork. The dedicated team evaluates all proposals. Nonetheless, this evaluation team is not always capable of evaluating all relevant aspects. Then, they enlist experts from other departments as needed. For instance, R&D engineers often conduct the technical assessment of the startup's product. Furthermore, *Realistic Adapters* know that they cannot apply their standard evaluation criteria to new venture suppliers. Most firms of this archetype, therefore, relax these criteria or conduct only a light version of the supplier assessment. As the informant of *Electronics* told us:

> I cannot expect a startup to have all the certifications upfront. Instead, we use a special approach that is simpler and less complicated in order to be able to begin a collaboration with a startup. […] We do not carry out full audits but we assess the most important things. For example, who owns the IP rights? (Electronics)

> We need to make compromises with such small suppliers. The reason we want to cooperate is because we want access to the innovation. But one cannot expect this firm to think and act like a large supplier that has been on the market for a long time. We are definitely aware of that. (Electronics)

Systematic Selectors. The third archetype has the most advanced supplier selection approach. For these companies, engaging with startups is a strategic objective since they believe that such collaborations offer radical benefits that other external partners do not. In addition, they not only consider startups as attractive partners due to their innovativeness but they also emphasize their special capabilities. *Systematic Selectors* generally do not restrict themselves to startups from their own knowledge base. As the following quote shows:

> I am open to all good ideas. We do not restrict ourselves, because if we do, we miss the opportunity to find something exciting. (Logistics)

Systematic Selectors have modified their organization by creating dedicated startup departments in which relationship managers work full-time positions on issues related to startups. These departments are more visible within the organization than, for example, an R&D department that has been tasked with this responsibility. Even though *Systematic Selectors* actively search for startups, they are also comfortable with the passive way of identification, because they know how costly and labor-intensive it is to screen the market themselves. By encouraging startups to apply via the firm's website, they can focus on evaluating the incoming proposals instead of screening the whole market. These firms can even convince their actively identified startups to submit an application in order to evaluate all proposals by the same process. In addition, this procedure provides an overview of all startup contacts, facilitating comparisons and enabling the management to analyze their startup activities statistically. The firms of this archetype provide information about technologies they are interested in on their web pages so that applicants can know immediately whether or not their products meet the needs of the buying firm. These "pain points" are coordinated with top management and the related business units.

The evaluation process consists of several stages with a clear objective to sort out applicants who do not fit the buying firm's requirements. Hence, contingent on the stage, experts from different departments are involved alongside the relationship managers who are present throughout the process. *Systematic Selectors* are the only firms whose evaluation criteria are akin to those that venture capital firms typically apply before investing. One informant describes their standard process for applying these criteria as follows:

> We have developed a catalog of 20 to 30 criteria. This is like a blind tasting. Each of us evaluates the startups independently, then I merge the results, and we make an analysis. […] If a startup gets a poor evaluation, against our initial expectations, we ask why and we discuss it. (Logistics)

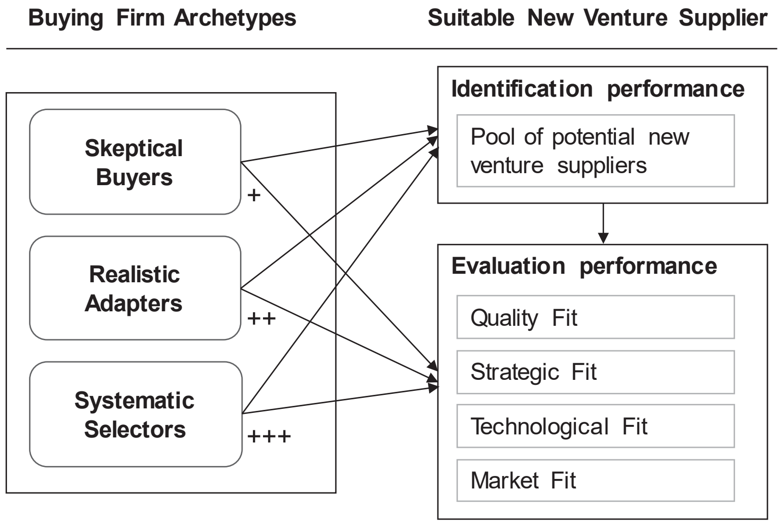
As these criteria differ from traditional supplier evaluation criteria, our case firms highlight that the evaluation team possesses the skills to apply these criteria properly. Therefore, many of the employees from the startup department have previously worked in the entrepreneurial sector. A final peculiarity of the evaluation process of *Systematic Selectors* is that the potential adopter of the technology is involved in the last evaluation stage:

> Before startups are ever accepted, the internal clients are identified within [Automotive]. They need to have a strong interest in using a startup's technology in the future. These internal clients then collaborate with the startup in a real innovation project, like they would work with any other technology partner, like Bosch or Conti. (Automotive, AD)

A MODEL OF SELECTING NEW VENTURE SUPPLIERS

In this section, we compare the characteristics of each archetype with the dimensions of suitable new venture suppliers (quality, strategic, technological, and market fit), and develop propositions regarding the performance implications of major adaptations of firms' selection approaches. We summarize our findings in a theoretical model in Figure 1.

Figure 1: Theoretical Model that Connects Buying Firm Archetypes with their Likelihood to Select Suitable New Venture Suppliers



Note: “+” = limited chances to select suitable new venture suppliers, “++” = medium chances to select suitable new venture suppliers, “+++” = high chances to select suitable new venture suppliers.

The analysis of the selection process of firms of the first archetype reveals that most *Skeptical Buyers* have only selected a small number of new venture suppliers. They would engage with more; however, they do not actively search for potential new venture suppliers in the market but wait for accidental encounters with them. For instance, *Piping’s* informant explained that the startup they selected was located next to one of their facilities and that other contacts to startups were always established through personal networks; sometimes startups simply contacted senior managers directly:

> This startup contacted our HR manager, so we invited the founders to present their ideas. They were very convincing, so we decided just to try it out. (Piping)

This statement reveals another problem of the selection process. Although *Skeptical Buyers* are skeptical of startups’ capabilities, they do not carefully evaluate them. They invite startups for pitches or they discuss their products in personal meetings but there is no formal evaluation, which raises doubts whether these firms properly assess the quality of new venture suppliers. In particular, the lack of predefined evaluation criteria and the missing startup specific experience of the selecting managers increase the risk of selecting low-quality startups. We know from *Realistic Adapters* and from *Systematic Selectors* that they adjusted their selection processes because they found them unsuitable for selecting new venture suppliers. For instance, the head of *Automotive’s* startup department explained how the selection process had worked before they professionalized it:

> [Before we adjusted the evaluation process] an engineer went to a conference at the MIT and got approached by 20 startups who all told him that they have invented the Holy Grail. Of course, he is thrilled. Because he compared them with his team which is still using Windows Explorer 6. […] of course he is enthusiastic, but he cannot tell the difference between a good startup and a bad one. He has never heard of the Y Combinator and does not know what the difference between seed stage and later stage

> is, and he probably thinks, if a startup has not yet raised venture capital, then that is even better. (Automotive)

This quote shows that selecting new venture suppliers requires special processes and skills that *Skeptical Buyers* do not yet possess. Therefore, it is unlikely that they will make the best choice of new venture supplier. Accordingly, we propose:

> **Proposition 1**: Buying firms that do not adjust their supplier selection process for startups are unlikely to select suitable new venture suppliers.

Identification performance. The analysis of the selection process of *Realistic Adapters* and *Systematic Selectors* in terms of their identification efforts reveals several important findings. Unlike *Skeptical Buyers*, *Realistic Adapters* use a variety of channels to establish contacts with startups. They are open to external innovation and strongly emphasize active identification strategies which are consistent with previously defined strategic search areas. These search areas are coordinated within the organization. *Realistic Adapters* keep the whole organization apprised of the technologies that interest them, so that employees can evaluate whether or not a startup could be a worthwhile supplier. Especially managers from departments that maintain close contacts with universities (R&D), customers (sales), or suppliers (procurement) are instructed to notify the dedicated team whenever a potentially attractive startup appears. This seems to be an efficient way to screen close markets without having to employ large screening units. For instance, *Medtech's* dedicated startup manager receives dozens of proposals of potential partners from other departments every year; this has already resulted in a small number of collaborations.

Although the centralization of the identification process facilitates engagement, and *Realistic Adapters* get in contact with substantially more startups than *Skeptical Buyers*, the identification process of *Realistic Adapters* has also some weaknesses:

> Depending on whom they [the startups] are contacting, they can be lucky or unlucky. Sometimes they contact developers who know only their area, but if they contact people who have an overview, these people would then contact our corporate research department, and then it would reach me. This is based on an internal network without a formal process behind it. However, it works best if you contact the right people directly. (Robotics)

In contrast to *Medtech*, *Robotics* is a multinational company with facilities all over the world. The informal forwarding of startup proposals is considerably more difficult than for a smaller company like *Medtech*. It is also hard to keep the whole organization informed about strategic search areas. The *Robotics* informant admitted that only a very small fraction of business units forward proposals to him, which shows that the responsible managers of *Realistic Adapters* often lack internal visibility. However, this missing external visibility might be even a more serious challenge as *Realistic Adapters* tend not to advertise their startup activities on special web pages, making it difficult for startups to find the right contact.

Systematic Selectors have found a solution for this problem. Their dedicated startup departments are much more visible internally and externally as they announce that they want to engage with startups. At the same time, through the online application forms, applications always receive the right contacts while keeping identification costs low:

> We communicate through startup media that we are interested in startup collaborations. The startups then come to us. Screening the whole market by yourself is enormously expensive. (Transportation)

In conclusion, the firms of the third archetype present the most sophisticated means of identifying startups. Although they use the same kinds of active identification and they leverage other departments to identify more startups, they can be distinguished from *Realistic Adapters* because their central focus is on passive identification strategies supported by a highly visible startup department. The following propositions summarize our findings regarding the identification process:

> **Proposition 2a**: Buying firms that effectively orchestrate departments with external contacts for active identification purposes identify more potential new venture suppliers and are thus more likely to select suitable new venture suppliers.

> **Proposition 2b**: Buying firms that establish a passive identification strategy receive more incoming startup proposals and identify startups more efficiently than firms that only pursue an active identification strategy. Thus, they are more likely to select suitable new venture suppliers.

> **Proposition 2c**: Buying firms that establish a dedicated startup function enhance their visibility which increases the number of incoming startup proposals. Thus, they are more likely to select suitable new venture suppliers.

Evaluation performance. The supplier selection process of *Realistic Adapters* seems to be better suited for achieving the objective of engaging with suitable new venture suppliers than is the process of the *Skeptical Buyers*. They have adjusted their selection process in two important ways. First, there is a clear organizational responsibility for selecting new venture suppliers. The team that evaluates these new suppliers gains experience and accumulates startup-specific knowledge, which might lead to a more accurate assessment of the startup's quality. Second, this team does not work completely unassisted. Instead, it involves potential adopters of the startups' innovations or technological experts from the R&D function to evaluate whether the technology fulfills technological and market requirements. As one informant explains:

> Of course we are working closely with people from R&D and the business unit for the technical and commercial evaluation. We will never collaborate with a startup if there is not one business unit which says: good idea! We are going to support you. (Adhesive)

Although these two measures make the evaluation process less permeable for unsuitable startups, firms of the second archetype have not adopted a fully standardized evaluation approach. Every startup is evaluated individually without any predefined startup-specific criteria. They only use the standard criteria that have been developed for their established suppliers. This is a major difference from *Systematic Selectors* who send all applicants through the same funnel-shaped evaluation process. Each stage of this evaluation process has predefined criteria and experts who evaluate them. The following two quotes show the differences between the evaluation criteria used by *Realistic Adapters* (Medtech) and *Systematic Selection* (Logistics):

> We assess which quality management system they have, if the product documentation is available if the registrations are all correct […] We evaluate the typical things. (Medtech)

> We evaluate a startup's team, its technology, innovation, the strategic relevance for us, the financial conditions of the startup […] (Logistics)

Finally, after ensuring that only high-quality startups with technologically attractive products that meet the strategic objectives of the buying firm have gone through the process so far, the startup department seeks a business unit or a department to act as an "internal client" interested in sourcing the product. This last stage ensures a market fit, because only if a business unit wants to purchase the innovation and provides the budget for a joint innovation project can the startup be selected. Hence, *Systematic Selectors* offer startups a clear path after selection. Among them, *Automotive* put a great deal of effort into achieving market fit through this "pull process." As its informant explains:

> This is a pull from the business unit and not a push. We do not say work together with the startup, instead, we ask them, whether they want to work with the startup. Then, they must prove that they would benefit from the collaboration and they have to offer a real project. […] Thus, there is no "Not-Invented-Here syndrome" because each department is convinced that they get access to a technology that they cannot find somewhere else. (Automotive)

In sum, *Systematic Selectors* appear to be more likely to select suitable new venture suppliers than *Realistic Adapters* or *Skeptical Buyers* as all startups go through the same evaluation process and put more effort and qualification into the assessment of the startups' quality. Based on our analysis, we formulate the following propositions about the evaluation process:

> **Proposition 3a**: Buying firms with clear organizational responsibility for startups' selection accumulate startup specific knowledge which improves their ability to evaluate startups' quality and are thus more likely to select suitable new venture suppliers.

> **Proposition 3b**: Buying firms that apply startup specific evaluation criteria can better evaluate a startup's quality and are thus more likely to select suitable new venture suppliers.

> **Proposition 3c**: Buying firms that involve experts from different functions in the evaluation process can better evaluate the strategic, technological, and market fit and are thus more likely to select suitable new venture suppliers.

> **Proposition 3d**: Buying firms involving potential internal clients (i.e., potential users of the technology) in the evaluation process can better evaluate the market fit and are thus more likely to select suitable new venture suppliers.

DISCUSSION

The main purpose of this study was to examine how buying firms select startups as their suppliers. From our observations, we developed a typology of three archetypes based on five elements of differentiation. By analyzing the commonalities and differences among them, we derived propositions explaining performance differences of the analyzed selection approaches. The findings of this study have the following implications.

Theoretical Implications

Despite requests for research on the supply chain management/entrepreneurship interface (Kickul et al., 2011; Shepherd & Patzelt, 2013) and increasing evidence from business practices that established firms regularly select startups as suppliers (Gimmy et al., 2017), the selection of these non-traditional suppliers has scarcely been studied. The development of three selection archetypes in this study is thus one of the first attempts to generate insights into this new phenomenon. Many other scholars have also developed archetypes, especially to explore a new topic (e.g., Bode et al., 2014; Pauwels et al., 2016). Our typology underlines the importance of distinguishing established from new venture suppliers in supplier selection (Zaremba et al., 2016). The results indicate that buying firms that hesitate to tailor their supplier selection process to the peculiarities of startups are less likely to select suitable new venture suppliers. In other words, business practices and the supplier selection literature should consider startups as a distinct supplier type for which different sets of tools, evaluation criteria, and processes are necessary.

We also found that determining the suitability of a new venture supplier is quite challenging as standard evaluation criteria are not applicable. Through this exploratory multiple case study research, we discovered that in addition to a new startup's *quality*, its product needs to fit the *technological*, *strategic* and *market* requirements of the buying firm. These findings extend previous studies suggesting that startups must fulfill a certain base capability to become selected (Zaremba et al., 2017). This base capability is akin to our concept of new venture quality fit, which *Systematic Selectors* in particular emphasize during the selection process. Zaremba et al. (2017) found that buying firms that have developed a *new venture partnering capability* (*NVPC*) adapt their supplier management and evaluation process to the peculiarities of new venture suppliers and evaluate their new venture suppliers more thoroughly than their established suppliers. *Systematic Selectors* therefore seem to be the only group of firms from our sample that exhibit a strong NVPC. Consequently, one can expect *Systematic Selectors* not only to select more suitable new venture suppliers but that also to collaborate more effectively with them after the selection.

The exploration of *Systematic Selectors'* evaluation processes reveals that buying firms not only need to apply standard evaluation criteria more flexibly, as found by Zaremba et al. (2017) but that they also need to develop new evaluation criteria designed to distinguish *high-quality* from *low-quality* startups. The venture capital literature has explored the evaluation criteria for a long time (e.g., Baum & Silverman, 2004). This research stream could thus provide useful insights for the selection of new venture suppliers. Buying firms and venture capital companies share the objective to select startups with above-average survival rates. Hence, criteria that determine whether to invest in a startup might overlap with criteria that determine whether a startup has the potential to become a reliable supplier. Nevertheless, despite adding more concrete evaluation criteria, the selection process remains less formalized, less deterministic and perhaps also less rational than the selection process of established suppliers which many studies describe as an optimization problem where the best-performing companies according to several evaluation criteria will be selected (e.g., Wetzstein et al., 2016). When new venture suppliers are being selected, there will probably never be an objectively best solution.

The proposed archetypes illustrate differences in the selection processes of the case companies. In response to their strategic objective to establish asymmetric buyer-supplier relationships, *Systematic Selectors* have made the most modifications to the supplier selection process for startups. They are the only ones with a systematic approach to selecting startups, whereby they put substantial emphasis on creating and filling a funnel of potential new venture suppliers from which they actively select the most suitable. By doing so, they remove as much randomness as possible from the process, which makes them the archetype most likely to select suitable new venture suppliers.

Realistic Adapters recognized that they need to actively identify potential new venture suppliers as they cannot select suppliers from a given set of available firms (Kim & Wagner, 2012). They have somewhat adapted their selection process for these suppliers, but they still hesitate to fully adapt and allocate resources for this purpose. For *Realistic Adapters*, startups are means to an end. They have long-term innovation roadmaps and want to source the technologies externally. They consider startups as one of many potential innovation sources. Consequently, many of these companies adopted a “wait-and-see” attitude, trying to gain experience while avoiding significant investments.

Finally, *Skeptical Buyers* do not create special processes, criteria or organizational structures. They are particularly skeptical whether the potential advantages of such asymmetric partnerships outweigh the risks. Therefore, these companies do not actively screen for startups, which gives these relationships a random character. They are also less interested in close collaborations after the selection has taken place. Instead, they focus more on directly purchasing the product. Moreover, their selection process greatly depends on principal decision makers. In most cases, top managers follow their intuition in choosing startups and then promoting the collaboration, often against internal obstacles. These findings support prior studies’ findings regarding the key role of principal decision makers in asymmetric collaborations (Minshall et al., 2010; Zaremba et al., 2017). Generally, firms of this archetype are expected to have only limited chances to select suitable new venture suppliers.

Managerial Implications

The proposed typology provides managers with a framework to classify their new venture supplier selection process, enabling them to identify potential areas of improvement. In particular, *Skeptical Buyers* should reconsider their current supplier selection practice and adjust their structures and processes accordingly. In this way, they can orient themselves on the five elements of differentiation (i.e., design themes). For identifying startups, buying firms should leverage their departments with external contacts. This approach would broaden the reach of the search without incurring additional costs. Moreover, buying firms should consider collaborating with external screening firms to extend the search to other industries which would help to overcome any lack of identification capabilities (Zaremba et al., 2016). On top of this, buying firms need to communicate their needs to startups. To do this effectively, they also need to offer a standardized way of gaining access to the buying firm (e.g., via an application process). Then, buying firms should concentrate their supplier evaluation efforts on determining the inherent quality of a startup. Therefore, buying firms need to adapt their evaluation criteria to the startup environment. To acquire the necessary know-how, the selecting managers need to be constantly involved in the selection process. Buying firms should internally assign the responsibility for new venture suppliers. Often, the responsibility is not assigned to procurement but R&D or a newly created startup department. Then, it is important to clarify the role of procurement to avoid conflicts over competences. Finally, becoming a *Systematic Selector* might not be advisable for all buying firms, as the establishment of a startup department could add more bureaucracy and tie up resources. Thus, many firms might be better off assigning the startup topic to an already existing department.

Limitations and Future Research Opportunities

We are aware of the following limitations which should be considered when interpreting the results. We used an exploratory qualitative research design that can lead to idiosyncratic conclusions (Eisenhardt, 1989). We tried to compensate for this potential shortcoming by selecting our case firms carefully and by interviewing informants from a variety of manufacturing firms and industries. In addition, we focused only on manufacturing ventures offering tangible high-tech products. Nonetheless, we are confident that our findings are also relevant for selecting startups with intangible products (e.g., service or software ventures) since most of our case firms have also selected service ventures with the same approach.

Naturally, distilling firms into archetypes is a simplification of reality, but we are convinced that this archival classification is a good starting point for future research. Specifically, future quantitative research could test the presented propositions with a larger sample. Furthermore, many established firms lack a general understanding of startups and are not familiar with how to identify them. Therefore, scholars could investigate how established firms can effectively outsource their identification activities to external intermediaries. Also, there is still a great deal of uncertainty about how established firms can manage new venture suppliers systematically after they became selected. Finally, from the startup perspective, it would be worthwhile to explore how startups perceive the supplier selection process and how they can successfully maneuver through it.

APPENDIX

Interview guide

A) Profile
- What is the main business of your company? What products does it cover?
- Please describe your role/competencies in the organization.
- What strategy does your company generally follow when selecting suppliers?
- What experience has your company already gained with startups?
- Which startups has your firm selected as suppliers within the last 12 months?

B) Specific selection example
Let us talk about the selection of [name of specific startup discussed prior to the interview]
- How did you find this startup? (Or did it find you?)
  - Is this the usual way you identify startup suppliers?
  - How is the identification process organized in your firm?
- How did you evaluate this startup?
  - Which evaluation criteria did you apply?
  - Had this startup fulfill certain base/minimum criteria to become selected?
  - Who did evaluate this startup?
  - Did the evaluation process differ from other startups you or your firm selected?
- If you compare the selection process of this startup with the selection process of a normal supplier, what is different?
- How did you proceed after the selection?
  - What is the current state of the collaboration?
  - How do you manage this collaboration? (Is this the usual approach?)
  - Who manages the collaboration?
  - How would you assess the collaboration so far? Has it been successful? If so, why?

C) General questions
- If you think about the other startups your firm has selected in the last 12 months, why did you selected them? What makes these startups attractive to you as suppliers?
- Which products/technologies are these startups delivering to you?
- What is a suitable startup supplier for your firm in your opinion?
  - Is [name of the startup from the exemplary case] a suitable startup supplier? Why? What are inadequate startup suppliers?
- Do startups also approach you?
  - What are you doing then? Are there some processes in place? How do you decide whether the product/technology is of interest to you?
- Do startups also apply for traditional tenders?
  - How do you proceed then?
- Do you see particular risks when selecting a startup as your supplier?

D) Conclusion and outlook
- Please draw a brief conclusion. How would you assess the current state of your firm's startup supplier selection process?
- What are the main insights and learnings you have gained in selecting startup suppliers?
- How can your firm become better in selecting startups as suppliers?
- Would you like to add something to the answers?

REFERENCES

Aerts, K., Matthyssens, P., & Vandenbempt, K. (2007). Critical role and screening practices of European business incubators. *Technovation, 27* (5), 254-267.

Arend, R. J., & Wisner, J. D. (2005). Small business and supply chain management: Is there a fit? *Journal of Business Venturing*, *20* (3), 403-436.

Azadegan, A., & Dooley, K. J. (2010). Supplier innovativeness, organizational learning styles, and manufacturer performance: An empirical assessment. *Journal of Operations Management, 28* (6), 488-505.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17* (1), 99-120.

Bhalla, A., & Terjesen, S. (2013). Cannot make do without you: Outsourcing by knowledge-intensive new firms in supplier networks. *Industrial Marketing Management*, *42* (2), 166-179.

Barratt, M., Choi, T. Y., & Li, M. (2011). Qualitative case studies in operations management: Trends, research outcomes, and future research implications. *Journal of Operations Management*, *29* (4), 329-342.

Baum, J. A., & Silverman, B. S. (2004). Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *Journal of Business Venturing, 19* (3), 411-436.

Bode, C., Hübner, D., & Wagner, S. M. (2014). Managing financially distressed suppliers: An exploratory study. *Journal of Supply Chain Management*, *50* (4), 24-43.

Brachtendorf, C. (2016). *Selecting Innovative and Start-Up Suppliers*. Dissertation No. 23302, ETH Zurich.

Chesbrough, H. W. (2006). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Boston, MA: Harvard Business Press.

Choi T. Y., & Hartley J. L. (1996). An exploration of supplier selection practices across the supply chain, *Journal of Operations Management*, *14* (4), 333-343.

Das, T., & He, I. Y. (2006). Entrepreneurial firms in search of established partners: review and recommendations. *International Journal of Entrepreneurial Behavior & Research, 12* (3), 114-143.

Durach, C. F., Kurpjuweit, S., & Wagner, S. M. (2017). The impact of additive manufacturing on supply chains. *International Journal of Physical Distribution & Logistics Management*, *47* (10), 954-971.

Dutta, D. K., & Hora, M. (2017). From invention success to commercialization success: Technology ventures and the benefits of upstream and downstream supply-chain alliances. *Journal of Small Business Management*, *55* (2), 216-235.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review, 23* (4), 660-679.

Dyer, J. H., Singh, H., & Hesterly, W. S. (2018). The relational view revisited: A dynamic perspective on value creation and value capture. *Strategic Management Journal*, *39* (12), 3140-3162.

Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review, 32* (4), 1246-1264.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14* (4); 532-550.

Flick, U. (2014). *An Introduction to Qualitative Research*. London: Sage.

Gimmy, G., Kanbach, D., Stubner, S., Konig, A., & Enders, A. (2017). What BMW's corporate VC offers that regular investors can't. *Harvard Business Review Digital Articles*. 2017, (July 27), 2-6.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, *16* (1), 15-31.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine.

Gonzalez, M. E., Quesada, G., & Mora Monge, C. A. (2004). Determining the importance of the supplier selection process in manufacturing: A case study. *International Journal of Physical Distribution & Logistics Management, 34* (6), 492-504.

Hitt, M. A., Ireland, R. D., Camp, S. M., & Sexton, D. L. (2001). Strategic entrepreneurship: Entrepreneurial strategies for wealth creation. *Strategic Management Journal, 22* (6-7), 479-491.

Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research, 202* (1), 16-24.

Hogenhuis, B. N., van den Hende, E. A., & Hultink, E. J. (2016). When should large firms collaborate with young ventures? Understanding young firms' strengths can help firms make the right decisions around asymmetric collaborations. *Research-Technology Management, 59* (1), 39-47.

Ili, S., Albers, A., & Miller, S. (2010). Open innovation in the automotive industry. *R&D Management*, *40* (3), 246-255.

Ireland, R. D., Hitt, M. A., & Sirmon, D. G. (2003). A model of strategic entrepreneurship: The construct and its dimensions. *Journal of Management, 29* (6), 963-989.

Kannan, V. R., & Tan, K. C. (2002). Supplier selection and assessment: Their impact on business performance. *Journal of Supply Chain Management, 38* (3), 11-21.

Ketokivi, M., & Choi, T. Y. (2014). Renaissance of case research as a scientific method. *Journal of Operations Management*, *32* (5), 232-240.

Kickul, J. R., Griffiths, M. D., Jayaram, J., & Wagner, S. M. (2011). Operations management, entrepreneurship, and value creation: Emerging opportunities in a cross-disciplinary context. *Journal of Operations Management, 29* (1–2), 78-85.

Kim, D. Y., & Wagner, S. M. (2012). Supplier selection problem revisited from the perspective of product configuration. *International Journal of Production Research, 50* (11), 2864-2876.

Koufteros, X. A., Cheng, T. E., & Lai, K. H. (2007). "Black-box" and "gray-box" supplier integration in product development: Antecedents, consequences and the moderating role of firm size. *Journal of Operations Management*, *25* (4), 847-870.

Kurpjuweit, S., Reinerth, D., & Wagner, S. M. (2018). Supplier innovation push: Timing strategies and best practices. *Research-Technology Management*, *61* (2), 47-55.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159-174.

Langner, B., & Seidel V.P. (2009). Collaborative concept development using supplier competitions: Insights from the automotive industry. *Journal of Engineering and Technology Management*, 26 (1-2), 1-14.

Miles, M.B. & Huberman, A.M. (1994). *Qualitative Data Analysis.* 2nd ed., Newbury Park, CA: Sage.

Minshall, T., Mortara, L., Valli, R., & Probert, D. (2010). Making "asymmetric" partnerships work. *Research-Technology Management*, *53* (3), 53-63.

Moschner, S.-L., Fink, A., Kurpjuweit, S., Wagner, S. M., & Herstatt, C. (2019). Towards a better understanding of corporate accelerator models. *Business Horizons*, *62* (4), in press.

Oke, A., Prajogo, D. I., & Jayaram, J. (2013). Strengthening the innovation chain: The role of internal innovation climate and strategic relationships with supply chain partners. *Journal of Supply Chain Management*, *49* (4), 43-58.

Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods: Integrating Theory and Practice*. 4th ed., Thousand Oaks, CA: Sage Publications.

Pauwels, C., Clarysse, B., Wright, M., & Van Hove, J. (2016). Understanding a new generation incubation model: The accelerator. *Technovation*, *50–51*, 13-24.

Poutziouris, P. (2003). The strategic orientation of owner-managers of small ventures: Evidence from the UK small business economy. *International Journal of Entrepreneurial Behavior & Research*, *9* (5), 185-214.

Pulles, N. J., Veldman, J., & Schiele, H. (2014). Identifying innovative suppliers in business networks: An empirical study. *Industrial Marketing Management, 43* (3), 409-418.

Reuber, A. R., & Fischer, E. (2005). The company you keep: How young firms in different competitive contexts signal reputation through their customers. *Entrepreneurship Theory & Practice, 29* (1), 57-78.

Rottenburger, J. R., & Kaufmann, L. (2019). Picking on the new kid: Firm newness and deception in buyer-supplier negotiations. *Journal of Purchasing and Supply Management*, in press.

Schaettgen, N., & Mur, S. (2016). *The Age of Collaboration: Startups and Corporates Need Each Other.* Vienna: MM Ventures and Arthur D. Little.

Schiele, H (2010). Early supplier integration: The dual role of purchasing in new product development. *R&D Management*, *40* (2), 138-153.

Shepherd, D. A., & Patzelt, H. (2013). Operational entrepreneurship: How operations management research can advance entrepreneurship. *Production and Operations Management, 22* (6), 1416-1422.

Song, M., & Di Benedetto, C. A. (2008). Supplier's involvement and success of radical new product development in new ventures. *Journal of Operations Management*, *26* (1), 1-22.

Song, M., Podoynitsyna, K., Van Der Bij, H., & Halman, J. I. (2008). Success factors in new ventures: A meta-analysis. *Journal of Product Innovation Management*, *25* (1), 7-27.

Spina, G., Caniato, F., Luzzini, D., & Ronchi, S. (2013). Past, present and future trends of purchasing and supply management: An extensive literature review. *Industrial Marketing Management, 42* (8), 1202-1212.

Stinchcombe, A. (1965). Social structure and organizations. In: March, J.G. (ed.), *Handbook of Organizations*. Chicago, IL: Rand McNally, 142-193.

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly, 44* (2), 315-349.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20* (3), 571-610.

Terjesen, S., Patel, P. C., & Covin, J. G. (2011). Alliance diversity, environmental context and the value of manufacturing capabilities among new high technology ventures. *Journal of Operations Management, 29* (1), 105-115.

Wagner, S. M., & Bode, C. (2014). Supplier relationship-specific investments and the role of safeguards for supplier innovation sharing. *Journal of Operations Management*, *32* (3), 65-78.

Weiblen, T., & Chesbrough, H. W. (2015). Engaging with startups to enhance corporate innovation. *California Management Review,* 57 (2), 66-90.

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: A review of research on open innovation. *Journal of Product Innovation Management*, *31* (4), 814-831.

Wetzstein, A., Hartmann, E., Benton, W. C. Jr, & Hohenstein, N.-O. (2016). A systematic assessment of supplier selection literature – State of the art and future scope. *International Journal of Production Economics, 182*, 304-323.

Zaremba, B. W., Bode, C., & Wagner, S. M. (2016). Strategic and operational determinants of relationship outcomes with new venture suppliers. *Journal of Business Logistics*, *37* (2), 152-167.

Zaremba, B. W., Bode, C., & Wagner, S. M. (2017). New venture partnering capability: An empirical investigation into how buying firms effectively leverage the potential of innovative new ventures. *Journal of Supply Chain Management, 53* (1), 41-64.

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. Academy of Management Review, 27 (3), 414-431.

DECISION SCIENCES INSTITUTE

Understanding Value Co-Creation Experiences in Innovation Communities: The Case of Makerspaces

Yi-Tzu Chang
Royal College of Art, UK
Email: yi-tzu.chang@network.rca.ac.uk

Cheng-Chieh Hsiao
Shih Hsin University, Taiwan
Email: cchsiao@mail.shu.edu.tw

ABSTRACT

Innovation has become a value co-creation process where people make new creation together with emerging technologies (e.g., 3D printing and laser cutting) in innovation communities, which often provide their members with diverse resources. This study draws upon motivation-opportunity-ability (MOA) theory to understand members' value co-creation experiences in innovation communities. Makerspaces are selected as the research context due to their practical significance. After interviewing 16 participants from different Makerspaces, the results indicate the types of value co-creation and MOA-relevant antecedents in innovation communities. Finally, implications for innovation research and practice are discussed.

KEYWORDS: Value Co-creation, Emerging Technology, Innovation Community, Makerspace, Motivation-Opportunity-Ability Theory

INTRODUCTION

With the advancement of emerging technologies, people often make innovation and creation in different innovation communities, such as crowdfunding platforms and Makerspaces. For example, the Maker movement has rapidly become a global phenomenon (MacMillan, 2012). Makers, the creators and innovators in Makerspaces, can now fabricate new items with open source software, hardware and the emerging technology like 3D printing and laser cutting (Wang, 2013). Free, accessible and editable open source resources have also empowered individuals, and various communities to participate value co-creation across socioeconomic classes, geographies, and cultures.

According to a report from the *Times*, the output value of the Maker industry and 3D printer reached 2.2 billion USD in 2012, and an expected value of 8.41 billion USD by 2020 (Bajarin, 2014). *USA Today* also reported that the annual funding given to Maker-related industries is about 29 billion USD (Stewart, 2013).

Makers tend to join innovation communities like Makerspace, which give them the opportunity to find new co-workers with mutual interests and to build a team or found a company. Makerspaces also convene talented people in various types of activities such as a speech or a startup competition. Compared to a traditional office unit, the cross-disciplinary environment found in Makerspaces draws more attention to the creative workers (Cavalcanti, 2013). This is why Makerspaces have become more and more important in new innovation and product development.

Makers and creators have transform the creating experience into the value co-creation process with others in communities of innovation. Value co-creation relies on Makers' and co-creators' enthusiasm in exploring, experimenting, and fabricating, while also requiring knowledge, expertise, and a degree of community's support. However, little extant research has investigated makers' value co-creation or innovation experiences in Makerspaces.

As a result, this study attempts to explore members' value co-creation experiences in innovative communities (e.g., Makerspaces) and analyze how value co-creation occurs from the perspective of MOA theory (MacInnis & Jaworski, 1989; MacInnis et al., 1991). Drawing on empirical data from in-depth interviews, the analysis reveals an interactive relationship between Makers and their communities on value co-creation. Our findings are signification contributive to innovation research and technology management.

RESEARCH CONTEXT

This study takes Makerspace as an example of innovation communities. The term “Maker” originates from the “DIYer (Do it yourself)” culture, which focused on empowering individuals to learn simple handcraft techniques on their own in one’s garage in the 70s and 80s in America. One quintessential example is Steve Jobs, the founder of Apple, who successfully built the prototype of the Apple I with his friends in his garage (Gray, 2014). This kind of environment has now known as the community of innovation, which indicate the place like Makerspace and Hackerspace that allows creators can discuss their work and collaborate with the community. The idea of the Makerspace has its root in MIT (Massachusettes Institute of Technology)’s Fab Labs (fabrication laboratory), which has had a significant influence on Makerspaces. Fab Labs are defined as the design to fabricate things, and “consist of digital equipment for designing products and the digitally driven tools to create them” (Burke, 2014).

Typical equipments in Makerspaces include 3D printers, and laser cutters, which may provide coaching and consulting services to help to hone and implement startup collaborations. Several developed countries such as the United States, the United Kingdom, and Singapore have established industrial parks for talented people to exchange their ideas, which give Makers the opportunities to pursue and share further development of their products and ideas to the worldwide audience. It also aims to support members by creating an open platform and a space for cross-disciplinary idea exchange, conducting brainstorming and creative workshops. These spaces also encourage Makers, self-learners and micro-startup teams to identify and learn about the resources and technologies they will need for their work. There are more and more design schools in Taiwan starting to operate their own Makerspaces to help students to strengthen collaborative skills. To facilitate group cohesion and creativity, they encourage forming teams that interdisciplinary members to facilitate spontaneous learning and integrative discussions (Sawyer, 2007).

Innovation communities like Makerspaces build friendly, well-resourced environments for value co-creation among members; they give creators multiple channels to learn new knowledge and skills; and they organize regular events for sustained idea exchange in the innovation community. These help to motivate creators to actively participate in these spaces. With resources and opportunities from these communities, Makers can convert their imagination and ideas into real, tangible works. They can also find ways to give back to the community after learning new skills. Makers strive to transcend time and resource limitations of traditional products by exploring new methods and technologies together.

LITERATURE REVIEW

Innovation as Value Co-Creation

In this study, innovation is treated as a process of value co-creation. From a perspective of marketing, value co-creation is defined as consumers assume an active role and create value together with the firm (Koleret al. 2011; Prahalad and Ramaswamy 2004a). And process-based value co-creation framework consist three main typess: (1) customer value-creating process—in a business-to-customer relationship, (2) supplier value-creating process (3) encounter process. These types have share the different indications in innovation communities. First, the *customer value-creating process* is similar to the creators who tend to explore themselves and build their own product. This type of the maker start their creation with their problems in daily life, which inspire them to create a better solution to help themselves and others. Second, the *encounter process* is more about team-based co-creation. New ideas and products that are co-created by multiple groups in school labs or startup competition, share a mutual collaborative essence and match the culture of value co-creation in the Maker community. There are less samples about the *supplier value-creating process* for now, which is the part that community could work on in the future.

For Makerspaces, the more makers engage in community events like meetups and workshops, the more they have the chance to learn from each other. These participatory experiences help them to build things in teams and to bring products to market. This cyclical system helps build an accumulation of creative energy in the community, which gives creators a better chance to achieve breakthroughs in their work. The recent trend of micro-enterprise (Cheng, 2014) has led increasing numbers of people to start their own businesses. The process of value co-creation with different teams makes it possible for a product's creator to get an actual response from the market. The innovation community can be built to support the pattern and the environment of the creative industry (Anderson, 2012). There are industrial parks that support local Makers in other countries, such as Tech-city in England (*The Economist*, 2013), D camp in Korea (Wu, 2014) and BLOCK71 in Singapore (Lee, 2014). These industrial parks create an ecosystem for Makers by providing resources for startup groups, including consulting services, incubators, and accelerators.

Motivation-Opportunity-Ability (MOA) Theory

MOA theory is derived from the Motivation-Opportunity-Ability model. This theory was first applied in the field of industrial psychology (Lawshe, 1945). MacInnis and Jaworski (1989) extended this initial work by examining the influence of the three elements of MOA theory on marketing communication from the viewpoint of social psychology. Recently, MOA theory has been used to investigate value creation in consumer communities (e.g., Gruen et al., 2007). MacInnis, Moorman, and Jaworski (1991) argued that all three factors are essential and irreplaceable in accounting for human work processes. If one of the elements is missing, it is likely to lead to a major negative effect on productivity and value creation. In the MOA framework, Motivation represents the factors affecting how an individual pursues purposive action; Opportunity accounts for the environment and mechanisms constraining individual behavior; and Ability refers to personal assets, skills, and knowledge (Rothschild, 1999).

Hallaban (2000) views Motivation as the primary factor in the MOA model. This driving force represents an abstract process, which is an important component in initiating action. Macinnis, Moorman, and Jaworski (1991) conceive of Motivation as a combination of will, pleasure, interest, and desire to engage in meaningful conversation. Schiffman and Kanuk (2000) defined Motivation as a private inner force, which leads people to take action. In their model, the existence of this force is triggered by unsatisfied needs. The rise of crowdfunding platforms has helped many Makers turn their creations into businesses, shifting their activities from simply making fun and useful works to producing full, commercial product lines. Typical communities exchange ideas through workshops or one-off classes using simple tools. Makerspaces have provided the community with an alternative that incorporates digital tools such as 3D printers, CNC and laser cutters. Maker activities not only attract experienced Makers, but they also extend the variation of activities and facilitate new challenges in the production process. Fab cafe, a combination of a Makerspace and a coffee shop, is a new type of Makerspace. It offers a comfortable place for people to make new connections and create new works together.

The second factor of MOA theory is Opportunity, which reflects a situation that is conducive to achieving a desired outcome. An opportunity can be evaluated based on its achievement or failure (MacInnis et al., 1991). Rothschild (1999) explained that opportunity is the external environment plus a relational mechanism, which can bring about practical action in an individual. For Makers, opportunities include high-speed delivery in information, networks of human resources, and consulting services. Creators can use online public information such as open source tutorials to help their creation work (Wang, 2013). Different from Makers, the community at large is often managed by the operator of the Makerspace. These operators tend to add value by offering traditional handcraft classes. Creating good experiences in the space can encourage people to organize other events in the community, which makes interdisciplinary idea exchange possible. These opportunities are appealing not only to Makers and investors with common interests but also to other creative members of the community who wish to connect with each other.

Ability is the third factor of the MOA model. Rothschild (1999) indicated that ability comprises many aspects, including the relevant skills that a person is equipped with as well as the individual's ability to implement practical solutions using theoretical knowledge. Hoyer and MacInnis (1997) described ability as any resources an individual possesses that can contribute to creative achievements, including funds, social capital, skills, and knowledge. Gruen et al. (2006) defined ability as the crucial resources that can make things happen, such as intelligence and knowledge. Makers today often have backgrounds in design, electronics, and architecture, which give them the advantage in learning and using digital tools. Cross-disciplinary collaboration expands the range of possibilities and tends to maximize the interests of all individuals involved. Online platforms make the communication widely accessible, and help Makers promote and share their ideas more broadly. A Maker community can promote its vision by convening people from different industries and giving them a chance to learn collaboratively. Shared knowledge and tools in the space can allow new types of works to be imagined, prototyped, and completed. The community can also invite international speakers to share their experience, thereby building international bridges in both academic and industrial fields.

METHOD

Participants

This study conducted in-depth interviews with 16 Makers and Makerspaces operators in Taipei. Over 90% interviewees were males, which shows that female creators remain minority in the community. The average age of the interviewees were 32 year-old, whom are from design, creative, education and medical sectors. Sample innovation communities included Makerbar, Open Lab, Taipei Hackerspace, and Fablab Taipei, where the interviewees have their membership joined at least five to seven years, and one of the most experienced operator have been in the community for over 11 years.

Interviewees shared their experiences openly, and exhibited their works in the Maker Faire exhibition and participated in value co-creation in communities of innovation with digital tools like 3D printer and laser cutter. Interviews took place in Makerspaces or community gathering spaces, which allowed the researcher to observe interactions between Makers and their communities. The factors considered in gathering these data were chosen to test application of MOA theory, which emphasizes the influence of space, environment, and resources on human behavior.

Data Collection

This study used qualitative research methodology and gathered data through semi-structured, in-depth individual interviews. Semi-structured interviews are also known as semi-standardized interviews or guided interviews. For this approach, the researcher designs an outline before the interview based on the questions and the aims of the research. However, the interviewer does not always follow the order of the outline in the process. Normally, the interviewer can adjust the order and scope of questions based on the context and flow of the interview. Semi-structured interviews have the following advantages: (1) They allow the interviewer to collect data with a relatively open-mind on specific issues. The researcher may receive an unexpected result by taking a semi-standardized interviews approach. (2) The interviewee tends to reflect their experience more openly with fewer limitations in the interview. (3) Semi-standardized interviews tend to be an adequate approach for researchers who wish to learn about personal experiences of the interviewee or to compare the results with other data. (Guan, 2010)

The participants in this study were divided into two groups for in-depth interviews: (1) Makers who build their works in Makerspace, which includes MakerBar Taipei, FutureWard and Fablab Taipei; and (2) Operators who run Makerspaces, including Open Lab and Taipei Hackerspace. The follow-up interview questions are listed below: (1) How do you know about the Makerspace? (2) What is the reason you started operating this new type of space/ community? (3) What do you wish to change by running this space/community? (4) Have your experiences and the collaborations in Makerspace/Community influenced your work? If so, how?

After conducting in-depth interviews with Makers and Operators, the researcher analyzed the data using the motivation, opportunity, and ability framework. By reviewing the goals and achievements of both groups, the research aims to interactively identify weaknesses and gaps in these Maker programs in order to advance the development of this new industry.

Data Analysis

After conducting in-depth interviews, they were transcribed and analyzed. Two authors extracted and categorized the elements of value co-creation and its MOA antecedents. After the initial result was completed, it was validated by academics who specialize in this area of theory and checked by our interviewees.

FINDINGS

After qualitative analysis, this study illustrates the value co-creation experience and relevant factors within the MOA framework that foster member creation in innovation communities. These factors are related to the time when each individual began participating in communities of innovation (i.e., Makerspaces) and their reason for choosing to join this particular environment, which also helps to understand value co-creation experiences in innovation communities.

Types of Value Co-Creation in Innovation Communities

The results showed that two genres of value co-creation and innovation process in innovation communities (i.e., Makerspaces). The first type of value co-creation is innovation by yourself but inspired by others. In this type, innovation is created by an individual maker whose idea is inspired by his or her interaction with other members in communities of innovation. Interviewee G described that "*Maker is about inventing and modifying what you need. You will need to search information online and ask helps in the community to accomplish your project. Instead of offering you all the resource, you gather interdisciplinary people here at the same time, and inspire their imagination*."

The second is team-based innovation. In this type, innovation is an outcome of group dynamics among several members who work together for creation. Makers collaborate with cross-disciplinary talents to expend their project into a bigger scale. As Interviewee H described that "*For me, I care about work division. If it's a hard core project, I tend to collaborate with someone who already have knowhow. Like people in the community or people who know the program, and I will work on the design part to help finishing the project*."

Motivations for Value Co-Creation in Innovation Communities

Motivation is the first factor in MOA theory, which represent the driving force that leads an individual to take action. Our analysis revealed three categories that are related to motivation in this context, which included the following: (1) a delight in the making process, (2) a sense of achievement resulting from creation, (3) an interest in innovative experimentation. These three categories of motivation sparked Makers' interest in the Maker movement and captured the reasons for their continued involvement. These are discussed in detail below.

1. A Delight in the making process

Most interviewees first discovered the Maker movement through their interest in fabricating items. Different experiences in educational backgrounds also led to various results. Once a Maker had an initial experience of experimental creation, they typically discovered many more possibilities for continuing in the creative process. Different operation methods may also trigger Makers motivation and interest in creating new items. There were also some people whose motivations were closely linked with their own personal interests. Interviewee F stated: "*That was also the first time for us to do a workshop with another organization. One just felt that the process was so fun and interesting, even if it didn't balance your cost. You wouldn't think about the cost when making it. It's more about having this amazing opportunity for everyone. We used a 3D printer to make that piece at that time.*"

2. Sense of achievement resulting from creation

After experiencing the fun of fabrication, Makers often try to extend the development and value of their creations. It is easier for people to feel a sense of achievement when making items that match the Maker's original idea. Being motivated by a sense of achievement also leads to a pattern in the creative production process. In the community events and the co-creation process, most Makers are proud of the works they build from scratch. Aside from seeking a sense of achievement, community operators are dedicated to creating platforms that allow members to participate in more activities. Interviewee B explained: "*For me, it's about gathering people in my life, and we solve problems together. It is not about showing off how good I am in fixing things but about how we live. I have a sense of satisfaction when making my creation and helping other people. The more that people help others, the more they will get the sense of achievement. The reason I didn't emphasize [the importance of helping others] in the first place is that people will just show up if they are interested*."

3. An interest in innovative experimentation

There are also people who think their interests and development are limited by current environments and systems of production. They wish to break out of the status quo by changing fields and make improvements to other industries using digital fabrication. Interviewee O said: "*My system supports medical design with design. I hope to use digital tools to update the complicated medical systems in a modern way. My approach is to make products that are easier for people to operate*."

Opportunities for Value Co-Creation in Innovation Communities

The presence of an Opportunity indicates that a situation is conducive to achieving a desired outcome, and its presence can help to predict achievement or failure. Opportunity represents the external environment that brings about practical action from an individual. People not only join a Maker community to start creating; they also gain various opportunities in the process, which allow them to complete their project. There are three categories of opportunities, which include the following: (1) new opportunities from the community to extend an individual's work; (2) diverse modes of learning; and (3) interdisciplinary knowledge exchange.

1. New opportunities from the community to extend an individual's work

Normally, Makers will not buy their own 3D printer, so they go to a nearby makerspace to use specialized tools. It is important for makers to meet and collaborate with the community in the space. With technical assistance, creators can speed up production and execution on their project. Interviewee C described a Makerspace as: "A place with machines where one can also make things and conduct research with other people."

2. Diverse learning

Based on the type of Makerspace, a community may organize various events such as workshops and speeches. Workshops are chances for makers to develop new learning patterns, which let them engage people from different background and interests. Adjustments and improvements in their work may also lead to creative collaborations and finding new business partners. Aside from participating in workshops, members in the Maker community also have the opportunity to offer courses on their own, which can be a memorable experience. Interviewee I commented: "*I got the opportunity to start my own workshop. This resulted from a discussion with OO who helped to make it really happen. I think this was an important experience, which affected me a lot. It was not just the project but also the space that provides you with this kind of environment and people, which allows you to have chances to learn. It also helped to build a starting point for the future. I think this was the best part, and I also learned a lot*."

3. Interdisciplinary knowledge exchange

Unlike creators who are used to working on their own, Makers in this generation tend to share their creations online, which has led to an interdisciplinary knowledge exchange between online and offline members in the Maker community. Even foreign creators who have never heard of the Maker movement or have no experience in creating can connect with other members in the Taiwanese community. Interviewee N remarked: "*It's all about sharing knowledge and getting knowledge from each other about different subjects*." Those who cannot participate in an actual community may still be aware of information about Makerspaces online. Interviewee A described: "*What Fablab is doing is actually extending the spirit of things. You have to do things, learn things, and share them. You can make your creations and their influence extends through the Internet*."

Other than online information exchange, events like seminars are equally important. Seminars and forums often allow interdisciplinary talents to reach large audiences. Discussion is also a way to increase knowledge exchange.

Abilities for Value Co-Creation in Innovation Communities

Ability is the third factor of the MOA model, which comprises resources that an individual possesses that contribute to creative achievements, including funding, social capital, skills, and knowledge. The types of creations that people can complete are determined their ability. There are three categories of ability, which include the following: (1) integration of new and old cultural practices and repurposing outdated technologies, (2) empowerment resulting from creative modes of education, and (3) fabrication of novel, customized products in startup businesses.

1. Integration of new and old cultural practices and repurposing outdated technologies

Makers today not only have the ability to learn about digital tools, but they also apply what they know in their life, such as repairing objects. This incorporation of lived experience helps to renovating the production process using cultural practices. It shows that makers can solve actual problems that benefit people while extending the spirit of their traditional culture. With the access to new tools and spaces increases public

familiarity with fabrication techniques. Introducing easy, accessible techniques can engage more people to join the movement. Interviewee A commented: "*Fablab was originally built for facilitating all kinds of knowledge collaboration with digital fabrication. What digital fabrication does is lower the requirements [of using this technology] by equipping normal people with machines they would normally use only in the lab*."

2. Empowerment resulting from creative modes of education

Innovation communities represent social structures that link creators together. These concepts help contribute to the goal of empowering individuals. Many Taiwanese feel that their actions and imaginations are limited by aspects of their formal education. However, new tools adopted recently in new educational programs have shaped new perspectives. Interviewee L, a teacher, said: "*I'm actually organizing [Maker activities] in our school club, which gives me the opportunity to create, as well as the time, the space, and the ability to give back to the kids, mainly to provide resources to them. I think it's important for teachers to provide resources and direction to students, which is more akin to guiding than directing*." Makerspaces are equipped with all kinds of tools to support creators and encourage people to develop new abilities. More participation from a community also inspires more creation.

3. Customization and creative, entrepreneurial ability

To adapt to a quickly-changing entrepreneurial environment, makers not only must find a way to develop customized products, but they also must find suitable Makerspaces for themselves. This can help them connect to new product lines and online platforms, ultimately enabling them to build strong business models. Products made by the new generation of designers thus focus more on personalization and customization. It is natural for producers and creators to incorporate digital fabrication into the customization process. In addition, digital fabrication also lowers requirements for digital fabrication and encourages more people to create at home. This has already been applied to several fields. Interviewee G described his experience: "*I was making customized products. With digital tools and design, this often leads to customization for the majority [of customers]. It means everyone can make their own product*."

Discussion

Recently, communities of innovation have become more and more popular. Makerspaces is one of such communities where emerging technologies are provided for value co-creation among members or makers. This study applied the Motivation-Opportunity-Ability (MOA) theory to explain the interactive relationship between Makers and their communities, considering each actor's creative assets in order to extend the applicability of the MOA framework and provide pragmatic recommendations to the Maker movement in Taiwan.

This study relies on evidence from 16 interviews with key figures in Taiwan's Maker movement, finding three principal motivations for joining this movement: (1) a delight in the making process, (2) a sense of achievement resulting from creation, (3) an interest in innovative experimentation. Through pursuing these motives, the Makers typically discovered new opportunities in their communities to extend their work, such as marketable applications, diverse learning, and interdisciplinary knowledge exchange. Finally, through these opportunities, Makers sustained further creative collaborations, whose scope included (1) integration of new and old cultural practices and repurposing outdated technologies, (2) empowerment resulting from creative modes of education, and (3) fabrication of novel, customized products. According to our findings, this study is helpful for practitioners in innovation communities, such as crowdfunding platforms, Makerspaces, and R&D departments. Our findings can be applied to (1) innovation education, (2) new product development, (3) innovation management through group dynamics, (4) crowdfunding, and (5) micro-entrepreneurship.

Even if there are successful promotions in some communities government, Current promotions in communities have yet to reach the mainstream. There is also a lack of public discussion between communities. Taiwanese design education could also better support students in incorporating handicrafts into their work. Most students come to Makerspaces for school assignments, which may inhibit their space for making new creations. Lastly, many people do not pursue innovative business models due to the popularity of the original equipment manufacturer (OEMs) model in Taiwan. This discourages the independent development of new technologies. With limited ideas about how firms can operate in Taiwan, co-working spaces and service-oriented businesses do not flourish, which reduces the resources available to local Makers. Aside from rare corporate sponsorship opportunities, most makers must fund their own initial investments in learning and research. It is important to provide actual resources, expand funding channels, and improve communication with the community at-large. This study is a rare academic study on Makers and the Maker community in Taiwan. The issue of how to effectively foster Maker communities has become increasingly important to modern society. This analysis provides an academic perspective on Makers and suggests a need for further investigation of creative collaboration processes. It aims to fill a gap in the current MOA theory framework. Gathering feedback from Makers and community operators allows for triangulation of diverse perspectives, which can inform the process of designing public policy, educational curriculum, and community ecosystems. This study can also be used to promote the Maker movement, which imbues daily life and education with new possibilities involving technology, design, medicine, entertainment, and cultural preservation. Overall, incorporating maker culture into local communities can encourage new Makers to collaboratively build creative works and foster deeper connections.

REFERENCES

Anderson, C. (2012). *Makers: The Industrial Revolution*. Crown Business.

Bajarin, T. (2014). "Why the Maker Movement Is Important to America's Future." *TIME*. Retrieved from http://time.com/104210/maker-faire-maker-%20movement/ on May 19, 2014.

Lanham & MD & Rowman & Littlefield. (2014). *Makerspaces: a practical guide for librarians*,Vol. 8.

Cavalcanti, G. (2013). “Is it a Hackerspace, Makerspace, TechShop, or FabLab?” *MAKE*. Retrieved from http://makezine.com/2013/05/22/the-difference-between-hackerspaces-makerspaces-techs hops-and-fablabs/ on May 22, 2013.

Chen, B. H. (2014). “Maker Faire: Infinite Creativity in Huashan 1914 park.” *Liberty Times Net.* Retrieved from http://news.ltn.com.tw/news/life/breakingnews/1015592 on May 24, 2014.

Cheng, C. K. (2014). “Hoping that micro-entreprise can flourish everywhere.” *CommonWealth Opinion*. Retrieved from http://opinion.cw.com.tw/blog/profile/60/article/1834 on September 10, 2014.

Fried, B. & Wetstone, K. (2014). “The White House Maker Faire: Today’s D.I.Y. Is Tomorrow’s 'Made in America'”*.* Retrieved from https:/obamawhitehouse.archives.gov/blog/2014/06/18/president-obama-white-house-maker-faire-today-s-diy-tomorrow-s-made-america on June 18, 2014.

Fitzwillie (2013). “Why the Maker Movement is Here to Stay.” *MAKE.* Retrieved from http://makezine.com/2013/06/03/why-the-maker- movement-is-here-to-stay/ on June 3, 2013.

Frow, P. & Storbacka, K & Payne, A.F. (2007). “Managing the co-creation of value”. *Journal of Academy of Marketing Science*, Vol. 36, 83-96

Gray, M. (2014). “Apple’s garage startup roots are really a ‘myth’.” *NEW YORK POST.* Retrieved from http://nypost.com/2014/12/07/apples-garage-startup-roots-are-really-a-myth/ on December 7, 2014.

Guan, X. S. (2010). *Research Methods* (3rd Edition). Chuan Hwa Publishing Ltd.

Gruen, T.W. & Osmonbekov, T. & Czaplewski, A.J. (2006), “eWOM: The Impact of Customer-to-Customer Online Know-how Exchange on Customer Value and Loyalty”. *Journal of Business Research*, Vol. 59, No. 4, 449-456.

Hoyer, W.D. & MacInnis, D. (1997). *Consumer Behavior*. Boston, MA Houghton Mifflin.

Lawshe, C.H. (1945). “Eight Ways to Check the Value of a Training Program.” *Factory Management and Maintenance*, Vol. 103, 117-120.

Lipson, H. & Kurman, M. (2012). *Fabricated: The New World of 3D Printing.* Wiley.

Lin, L. (2010). “Kickstarter: A Stage for Dreamers and Entrepreneurs to Raise Funding Globally.” *INSIDE*. Retrieved from https://www.inside.com.tw/article/636-kickstarter-funding-platform-for-creative-projects.

Lee, T. (2014). “Like lego blocks, Singapore government to double size of the Block 71 startup experiment.” *TECHINASHIA*. Retrieved from https://www.techinasia.com/singapores-block-71-double-capacity-2014 on March 12, 2014.

Maclnnis, Deborah J, & Jaworski, Bernard J. (1989). “Information Processing from Advertisements: Toward an Integrative Framework.” *Journal of Marketing*.

Maclnnis, D.J, Moorman, C.& Jaworski, B. J.(1991). “Enhancing and Measuring Consumers Motivation, Opportunity, and Ability to Process Brand Information from Ads.” *Journal of Marketing.*

MacMillan, T. (2012). “On State Street, Maker Movement Arrives.” Retrieved from http://www.newhavenindependent.org/index.php/archives/entry/make_haven/ on April 30, 2012.

Owyang, J. (2014). “Maker Movement and 3D Printing: Industry Stats." Retrieved from http://www.web-strategist.com/blog/2014/02/13/maker-movement-and-3d-printing-industry-stats/ on February 13, 2014.

Ranjan, K.R, Read, S. (2014). “Value co-creation: concept and measurement.” *Journal of Academy of Marketing Science*, Vol. 44, No. 3, 290-315.

Stewart, M. (2013). “Meet USA's New Entrepreneurs.” Retrieved from http://www.usatoday.com/story/money/business/2013/10/14/martha-stewart-column-meet-the-makers/2980701/ on October 14, 2013.

Sawyer, K. (2007). *Group Genius: The Creative Power of Collaboration.* Basic Books.

Schiffman, L.G. and Kanuk, L. L. (2000). *Consumer Behavior*. Upper Saddle River, N.J: Prentice Hall.

The Economist. (2011). “More than just Digital Quilting.” Retrieved from https://ww.economist.com/technology-quarterly/2011/12/03/more-than-just-digital-quilting on December 3, 2011.

The Economist. (2013). “London’s Tech City: Start me up.” Retrieved from https://www.economist.com/britain/2013/10/03/start-me-up on October 3, 2013.

Wang, S. C. (2013). “Making in Public and Globalization in Maker Movement.” *CTIMES COMPONENTS & CONVERGENCE*, Vol. 263 (September), 52-54.

Rothschild, M. (1999). “Carrots, Sticks, and Promises: A Conceptual Framework for the Management of Public Health and Social Issues Behavior.” *Journal of Marketing*, Vol. 63, 24-37.

Shopping Design. (2013). “From Promoting Social Design to Entering a Designer Society.” Retrieved from http://www.bnext.com.tw/article/view/id/26234 on January 25, 2013.

DECISION SCIENCES INSTITUTE

An Empirical Analysis of the Role of Corporate Strategic Resources in Product and Process Innovation Optimization

Young Sik Cho
Jackson State University
Email: dsmith@exsu.edu

Kevin Linderman
University of Minnesota
Email: linde037@umn.edu

ABSTRACT

From a resource-based view, this research examined the role of corporate strategic resources in innovation optimization. Based on data analysis from 257 U.S. companies, key findings are that (i) firms that heavily rely on knowledge-based resources tend to focus their efforts on process innovation; (ii) process innovation serves as an order winner in a dynamic market condition, while product innovation functions as the order winner in the stable market. This study contributes to the Operations Management literature by providing an alternative new resource-based innovation paradigm, rather than the existing market-based PLC innovation perspective, to optimize product and process innovation decisions.

KEYWORDS: Product innovation, Process innovation, Strategic resources, Market dynamism, Cluster analysis

DECISION SCIENCES INSTITUTE
Integration of building information modeling to the construction project management

Ricardo Machado
Pontifical Catholic University of Goias, Postgraduate Program in Industrial and Systems Engineering, Goias, Brazil, + 55 62 39461119
Email: drrmachado@gmail.com

Priscilla Freitas
Pontifical Catholic University of Goias, Brazil, + 55 62 39461119
Email: prisbf@hotmail.com

ABSTRACT

This research concerns the use of building information modeling (BIM) in building project management. Design science research was adopted to design and develop a model to solve a project management problem. The integration model was built by the development of a plug-in for Revit software, capable of providing information in a database powered by a planning software. An empirical study on a large building site was carried out to evaluate the model applicability and utility. Integration between BIM and project management through verification of constraints, visual management, and union of the parties involved with the project make both more efficient.

KEYWORDS: Project management, Information systems and technology, Manufacturing and production planning, Construction management, Building information modeling

INTRODUCTION

The Last Planner System (LPS) is a planning model, which aims to improve workflow reliability, reduce variability and losses in construction processes, and promote the reduction of costs and construction time (Ballard, 2000; Sacks et al., 2010). LPS consists of collecting sets of tasks to be performed and selecting those that can be efficiently assigned to executing teams. The plan´s approval occurs after the removal of constraints related to labor, materials, project specifications, and equipment.

Machado & Heineck (2006) suggest performing the planning of anticipations, which consists of raising requirements to ensure efficient work, involving actions to ensure elimination of restrictions and maintaining those elements that positively interfere in the uninterrupted workflow. The complexity of this proposal refers to the structuring of an information system capable of supporting the anticipation actions management in production planning.

The proposed changes in construction planning processes affect the quality and the amount of information required for decision-making. Building Information Modeling (BIM) emerges as an essential element in this scenario by providing the advantage of aggregating information into a single database. Furthermore, BIM supports the decision-making process in construction management systems. BIM allows configuring a reliable basis for decisions to support and improve processes of the project lifecycle, from digital representations of physical and functional characteristics of buildings (National Building Information Modeling Standard, 2007).

According to Sacks et al. (2010), lean thinking applied to construction has contributed to the development of production planning and control systems. According to these authors, the proposal presented by Koskela (2000), which conceptualizes production concerning transformation, flow, and value, clearly shows the repercussion of information maturity on benefits generated by the workflow. In this context, the use of LPS appears to reduce variability in construction processes and improve coordination workflow. According to Sacks (2010), a new conception of technology capable of providing required information to the planning system is necessary. Leite et al. (2016) state projects deal with a significant amount of data generated in a great variety of different formats. BIM emerges as a technological option that comprises all necessary functions for the project´s lifecycle. When adopted correctly, BIM contributes to a more integrated design and construction process, which generates higher quality projects with lower costs and durations (Eastman et al., 2011). Other benefits of BIM adoption can be mentioned, such as improving communication, facilitating logistics management, analyzing implementation plans, and controlling physical construction progress.

Information requirements that arise in the application of LPS are essential to establish the link with benefits provided by the BIM model. Sacks and Harel (2006) showed LPS reduces the information gap between project managers and subcontractors with the improvement of resources assignments decisions. According to Sacks, Treckmann, and Rozenfeld (2009), computer-aided visualization of construction product and processes facilitate reporting of project status and provide a support decision making to achieve stable production flow. Sacks et al. (2010) analyzed the existence of convergences between theoretical principles of lean construction and BIM and identified 56 possible interactions. Mendes Junior et al. (2014) presented a synthetic framework determining seven interactions associated with production planning and control after analyzing the results obtained by Sacks et al. (2010). Biotto (2012) presented a method for building site planning and control using BIM 4D generating information used in the decision-making process at all hierarchical levels plans. Bhatla and Leite (2012) verified the synergy of lean construction and adoption of BIM in a case study on medium and short-term planning processes in electrical, hydraulic, and mechanical work packages. Amongst the positive findings, they highlighted the functionality for removal of constraints, verifying of conflicts among projects, reducing cycle time, and support in communication between project participants in understanding the construction process. Mendes Junior et al. (2014) presented a theoretical framework of integration of BIM to LPS, based on the proposal developed by Bhatla and Leite (2012). Moreover, they found it is possible to obtain more information for production using integrated documents into a single database. Because of the research, they observed the reduction of variability, improvement of communication among those people involved with a project, reduction in cycle time, and decreasing of rework. Sacks et al. (2013) developed and applied an integrated BIM model to LPS, called KanBim. The authors elaborated a proposal in which production teams can check online the information about the current stage of work concerning the removal of its restrictions until the tasks that constituted them are ready to begin.

In analyzing the studies presented previously, there is a need to improve integration between BIM and LPS through visual management and automatic verification of constraints. The general objective of this research is to operationalize the integration of BIM to the LPS, focusing on medium and short-term plans managed visually in the 3D model. Given this primary goal, three secondary objectives emerged: to check constraints automatically, to generate medium and short terms plans (through the development of a plug-in), and to apply the model in an empirical study to assess their usefulness and applicability.

LITERATURE REVIEW

Last Planner system

Ballard and Howell (1998) proposed the LPS in the mid-1990s and constituted the primary application of the philosophy of lean manufacturing in the construction industry. LPS endeavors to apply lean manufacturing planning principles to construction projects. LPS emphasizes the need to manage production flows and commitment of work teams and provides a more reliable production, by promoting the reduction of the variability of the workflow (Ballard, 2000).
According to Laufer and Tucker (1987), LPS involves elaboration of long-term planning, the look-ahead planning (for the medium-term), and weekly or commitment planning (for the short-term). Long-term planning, also called master plan, should establish overall objectives that guide the enterprise and define the project's macro stages sequence, duration, and rhythm (Ballard, 2000). The master plan serves several purposes, such as expenditure projection and disbursement (Ballard, 1997). As stated by Fernández, Cárdenas, and Armiñana (2011), the master plan incorporates planning of all project activities, establishing deadlines and spatial relationships between different scheduled activities. On the other hand, look-ahead planning has the primary function of controlling workflow, through systematic identification and removal of constraints (Ballard, 1997). At this level of planning, the master plan is detailed and adjusted according to the available information (Ballard and Howell, 1998). The primary function of the look-ahead is the link between long and short-term planning. Look-ahead planning serves as a production protection mechanism in the short-term horizon, making it more efficient since it releases to execution only activities whose constraints have been removed (Ballard, 2000). The third level of LPS comprises commitment planning (or weekly work scheduling). In this plan, conciliation occurs between what should and what can be done. In this regard, the decision refers to what will be accomplished, considering work packages and those responsible for running them, based on available resources and prerequisites fulfilment (Ballard and Howell, 1998). According to Ballard (2000), weekly work scheduling occurs at the short-term planning level, by guiding assignments of teams responsible for each work package, after removing constraints in look-ahead planning. Priven and Sacks (2015) showed that weekly work plans (WWP) meetings promote a social network among the subcontractors involved with the project, with improved communication, which enhances coordination and improve workflow.
The essence of LPS is on information management regarding constraints to generate a reliable and uninterrupted workflow. Modeling information by BIM's proposal appears as a significant potential opportunity for applying LPS in the scenario discussed earlier.

Building information modeling (BIM)

According to Eastman et al. (2011), BIM is a modeling proposal associated with the processes of producing, communicating and analyzing the design of a building. Succar (2009) states BIM as a set of policies, procedures, and technologies that generate a methodology to manage building projects, from digital format data, throughout the life cycle of the building.
The main difference between BIM and 3D modeling software is the ability to generate parametric objects. According to Leite et al. (2016), information modeling concerns different sources of information to be shared throughout the project life cycle, interchangeability enabled by designed data formats and interoperability, and gathering data on work packages performed, among others. Parametricity allows generating editable objects, which can be changed automatically and support the BIM platform. Without this capability, a BIM software is just a three-dimensional modeler. Eastman et al. (2011) classify the uses and benefits of BIM considering the phases of the project life cycle, as follows:

(a) project design, involving a preliminary study and its viability;
(b) design, encompassing visualization, correction of elements in the model, generation of 2D drawings, multidisciplinary work, automatic extraction of quantitative, improvements in the process of energy analysis and sustainability;
(c) execution, including synchronization of the planning of work with objects of the model, the discovery of interferences between elements of the building, implementation of lean construction, synchronization of phases of acquisition, design, and construction;
(d) and operation, promoting improvement in managing operation of building systems.
Multiple dimensions discriminate BIM. According to Rodas (2015), the most explored model is BIM 3D. The dimension known as BIM 4D refers to time, which can offer temporal planning, enabling manage the building throughout its life cycle. In this aspect, one can visualize and evaluate the progress of activities to be carried out during the construction of the building. BIM 5D is associated with cost estimates, assigning the model ability to attribute values to the project elements. In this dimension, BIM becomes an aid to budgeting, automatically reflecting any change in the plan in costs. BIM 6D and 7D involve aspects related to sustainability, management and preventive maintenance of the enterprise.

RESEARCH METHODOLOGY

In this research, the authors adopted the methodology of design science. This approach establishes a systematic process, whose purpose is to design and develop innovative artefacts with problem-solving conditions relevant to the practice field (Van Aken, 2004; March and Storey, 2008; Dresch, Lacerda, and Antunes Jr, 2015). The adherence to the primary objective presented before determined the selection of design science research, as it required the creation of a practical model to improve interaction between project and building site medium and short-term planning. The research was carried out in four phases, as presented in Table 1.

Table 1: Research Design		
PHASE	DESCRIPTION	PERFORMED ACTIVITIES
1	Literature review	Identification and delimitation of the research problem
2	Exploratory studies	Search of knowledge about Revit, database, and software intended to accomplish the LPS Choice of organization adopted as study object Understanding the production planning and control process and implementation stage of BIM modeling in the organization adopted for study
3	BIM model development	Choice of building site Preliminary design project Model Application Model Evaluation
4	Consolidation of the BIM model	Communication on the results achieved Evaluation of the theoretical / practical contribution of the model

The literature review sought to ensure adoption of a problem with relevance to production planning and control within architecture, engineering and construction industry. In phase 2 of the exploratory study, the authors performed the analysis and evaluation of alternatives for BIM. According to Bilal et al. (2016), while there are many BIM design software on the market five products are the most considered for the users: Autodesk REVIT, Bentley MicroStation, Graphisoft ArchiCAD, Vectorworks, and Digital Project. REVIT is the most popular BIM design software among architects, engineers, designers, and contractors. The information generated by REVIT is stored in a centralized database allowing information sharing and collaboration among stakeholders of the project. REVIT supports many building simulations (for this research, the authors focused on construction planning and control). The authors adopted the REVIT software platform for research and training of researchers. After, it was selected a company involved in the implementation of BIM and accessible to interventions sought by the analysis. An organization that was carrying out an initial BIM implementation process was adopted for the study. The phase 3, regarding modeling, involved empirical research in the building site of the chosen company.

The empirical study was carried out in three stages, aiming to understand the problem studied, to develop and test the proposed model and to identify the practical and theoretical contribution of the solution obtained, as shown in Table 2.

Table 2: Stages of the Empirical Study

STAGE 1	STAGE 2	STAGE 3
Preliminary meeting with company managers	Plug-in adaptation	Model application
Participation in planning meetings	Planning software development	Model evaluation and improvement
Development of a data flow diagram	Performing model tests	Presentation of the proposed model to the project managers
Proposal of a preliminary model	Model refinement	Model application

The selected object of study was a building site from a construction company located in a large city, in the Midwest region of Brazil. The building site studied was chosen due to the process of implementation of management and planning, using BIM. The building site involved a project of a shopping center with a constructed area of 52,134.09 m² and an estimated value of US$ 32.845.823,00. The project was started in November 2014 and completed in April 2017.

In the phase 4 the generated model was evaluated considering the constructs presented in Table 3.

For each sub-construct, a question was generated and applied in the interview. The evaluation process took place through the configuration of a focal group, consisting of the planning engineer, the project-coordinating architect and the engineer responsible for the production, as well as trainees who use the model.

Initially, interviewees were informed about results achieved by the project and the peculiarities of the model generated. After this introductory phase, participants in the focused group were asked to evaluate the effects of the project. The authors used focused group interviews to verify the validity and applicability of the proposed model. According to Bilal et al. (2016) focused group interviews allows to understand multiple viewpoints and to collect the real-life experience of construction industry practitioners. At the end of the interviews, the research team analyzed the theoretical and practical contribution produced by research.

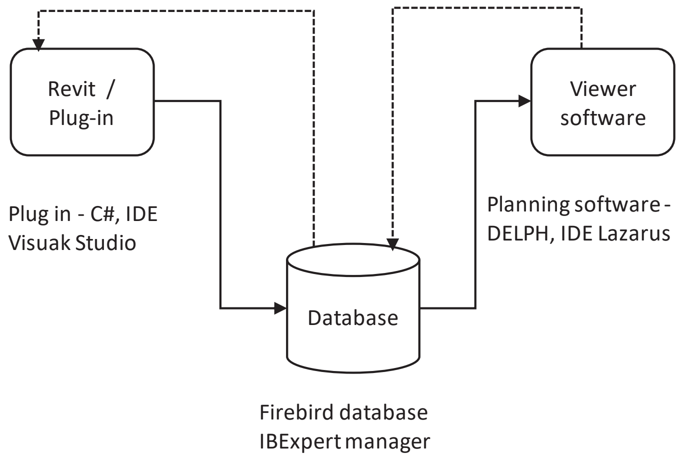
Table 3: Questions for Evaluation of Constructs		
CONSTRUCT 1	SUBCONSTRUCT	RESEARCH QUESTION
Utility	Information reliability	Is the information generated by the proposed model faithful to reality? Can information be viewed in time to allow its use?
	Increase of communication	Does the information generated by the model increase communication among stakeholders?
	Use of generated information	Is the information provided by the model used in production planning and control to make it more efficient?
	Integration of the information	The information that was previously in different databases have been aggregated into a single database?
Functionality	Easiness to use	Is the model's operation easy to learn?
	Easiness to understand information	Is the information generated by the model understandable?
	Effective feedback	Is it easy to feed the model with information from the construction site?
	Interest in continuing using the proposed model	Is the company interested in continuing using the proposed model?

INTEGRATION MODEL DEVELOPMENT

Introduction

Figure 1 shows the communication flow between the 3D model and LPS. The data connection between elements of the integration model was accomplished through classification of identification codes of geometric structures of the 3D model in Revit.

Figure 1: Data Communication



As can be seen in Figure 1, a database served as an intermediary element to establish communication between Revit modeling software and L7 planning software. Revit software in version 2017 was adopted in the model. The plug-in software was developed in C # programming language, and Visual Studio IDE environment was used in its construction. The database was structured in Firebird 3.0 environment, and IBExpert was used for its management. The planning software was built on Lazarus IDE platform and encoded in DELPHI programming language. The work of Rodrigues (2017) presents more details about configuration of the structure shown in Figure 1.

Development of a plug-in for the planning process

In order to structure medium and short-term plans, a plug-in was developed, as presented in Table 4.

Table 4: Functionalities, Parameters, and Goals of the Integrated Model

FUNCTIONALITY	PARAMETER	GOAL
Building site link to 3D model	Service Group of services	Assign geometric element to the structure of the building site
Look-ahead	PW – Productivity weight LI – Linked input LIC –Linked input consumption CD – Constraint Description	Assign labor consumption Assign necessary input to service Assign input consumption Assign project constraint
Weekly work plan	WWP	Assign commitment date Assign service completion
Service progress	Percentage executed	Assign service completion

The plug-in also received functionalities that enabled visualization of what was executed in the building site, to allow the monitoring of the project progress. According to Bilal et al. (2016), in the context of construction industry visualization provides a better understood of the building site allowing taking better decisions to resolve critical issues before fieldwork.

Among the features developed in the plug-in, the most relevant for this research was related to look-ahead planning. This functionality was created to select the part of the project that was intended to be carried out within four to six weeks and to check its restrictions regarding labor inputs, materials, and project specifications.

For deployment of the look-ahead functionality, it was necessary to create some parameters in the REVIT model. A numerical design parameter to verify labor constraint (PW - productivity weight) was created, which made it possible to assign different weights about the sections that involved different levels of labor consumption for the same service.

To verify constraints related to the materials needed to execute the programmed service, two design parameters were created: one of numeric type (LIC - linked input consumption), and another one of type text (LI - linked input). The parameter LI was created to associate the geometric element belonging to service to part name required to perform it. The parameter LIC was created to associate to the geometric element the consumption of material related to service to which it is linked.

In order to verify constraints related to project specifications, it was created the parameter CD – constraint description, whose function was to register pendency associated with a specific part of the project, if it was selected for inclusion in look-ahead planning. This parameter was intended to record some change in the design of a constructive element.

With the creation of parameters relating to project information and required inputs to work packages, it was possible to select a part of the 3D model, assign a period of running, check inputs available in building site and to analyze the existence of restrictions.

When checking through the plug-in, a procedure was activated that made the selected part to be included in the look-ahead receives green color, if needs of workforce, materials, and design were met. When the chosen section did not have its needs met concerning some restriction, it was presented in red color.

Another feature included in the plug-in was related to short-term plan - WWP, which showed visually what work packages were able to be carried out in the coming week, after removal of constraints and commitment from the team. It was enough to select the part of the project and assign it a date for its execution, which received the green color in a tone darker than the color assigned to look-ahead. Plug-in also allowed selecting in the 3D model a part of the building site that had already been executed and to assign it the date of execution. In the 3D model, the completed part of the building site changed from orange to blue.

Development of the planning software

For monitoring of the building site and insertion of data regarding available materials and workforce that fed the database, a desktop software was developed.

The first functionality inserted in the software allowed to create the managed building site and to segment it into groups of services, which were decomposed into work packages, to enable the formation of a more detailed look-ahead than the master plan. This procedure allows for insertion of information about unit compositions of work packages created in the software. In this way, each geometric element can be associated with the various inputs required for its execution.

In order to control the progress of building site, a functionality in the software was developed to generate comparative charts between scheduled and performed work packages. Another feature designed to monitor building site was to associate team’s settings (expressed in the number of employees) with labor consumption obtained.

Another functionality created in the software was related to management of existing materials in the building site. After the creation of this functionality, all material that entered the building site inventory was registered in the model, which allowed verifying availability of this input in the analysis of constraints in look-ahead.

Integrated model between BIM and LPS

BIM was supported by coupling of a plug-in to the REVIT software, to provide connection of geometric elements to logical structure of the project (transforming data generated by Revit with .rvt extension in data that fed a database created in Firebird 3.0 system with .fbd extension, and vice versa). Another association necessary to integrate with planning was the insertion of data referring to inputs (materials, workforce, and design specifications) necessary to carry out the scheduled work packages, making possible visual production planning and control within a 4D model.

The artefact developed by this research, involving integration of LPS with 3D modeling, started to function according to the model presented in Figure 2.

As indicated in figure 2, the detailed analytical structure of the project that allows the formation of medium and short-term plans is inserted in the software developed in the research and defines the master plan to serve as a reference of the contractual term. At that moment, unitary compositions of each work package provided in the project are determined, concerning material and labor inputs required for its execution. This structure is recorded by numerical codes generated in the Firebird 3.0 database, in which information created for communication between the 3D model and the planning software will be stored.

Next step is the preparation of 3D model for planning. To execute each of the geometric elements of the 3D model is necessary to connect them to the logical structure of the project and consumption of inputs required. This procedure requires creating project parameters described previously. Thus, the database is linked to the 3D model by plug-in added to Revit software.

Through plug-in, it is possible to transform .rvt data from Revit software into data with .fbd extension of Firebird database, and vice versa, making planning possible.

From this stage, the model is ready for application of LPS logic, which will be performed visually in Revit software. The process starts with look-ahead, where, through the functionality added by the plug-in, it is possible to select parts of the project to run.

Work packages selected for inclusion in planning undergo a procedure for checking material requirements, workforce and design specifications, informed by planning software to identify existing constraints.

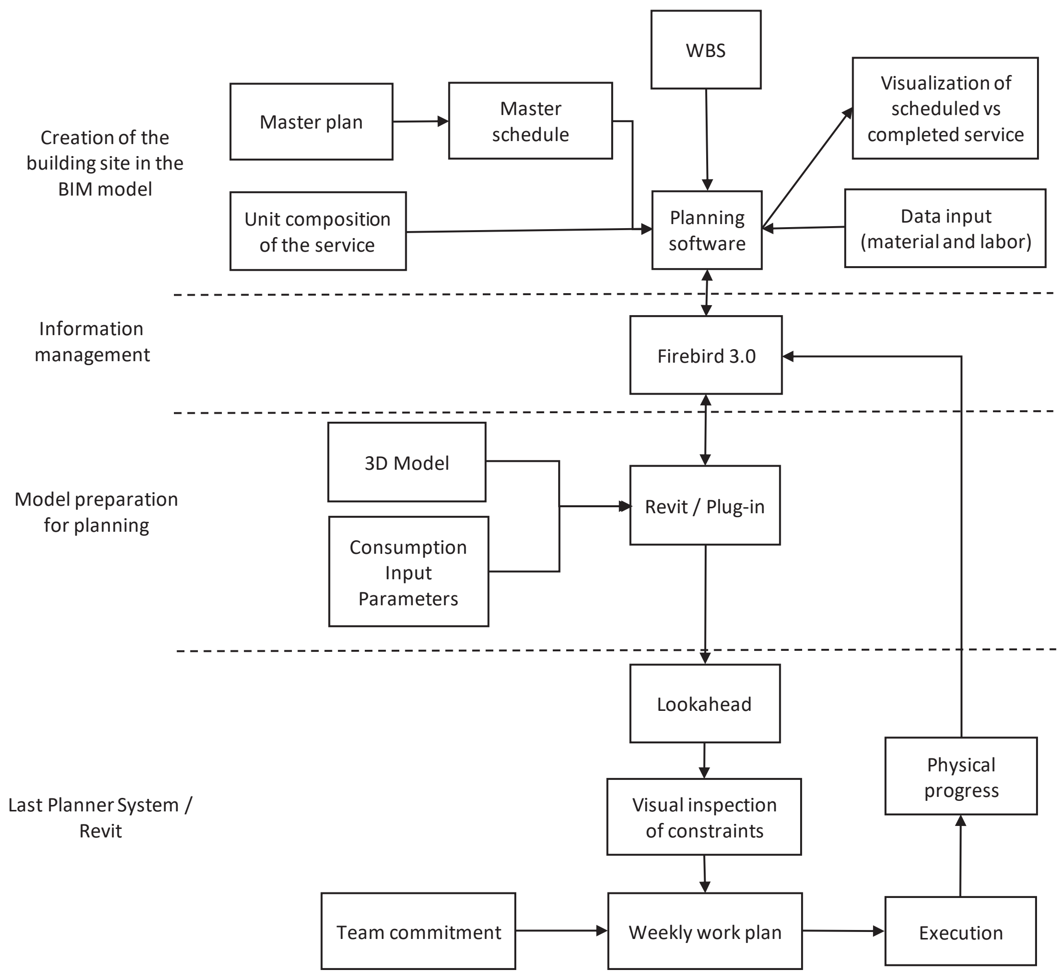
Management is done visually, since, with constraints, part of the building site in the 3D model will be displayed in red color, issuing a warning about what constraints should be removed.

When there is no constraint, the part selected is presented in light green color, indicating tasks available to be included in the WWP. After commitment approval session with teams responsible for work packages, the plan is selected to take place in the next week. Work packages approved at this stage will receive visual accompaniment to green in a darker tone that indicates approval for the look-ahead.

Following execution of the service, a physical advance is assigned indicating the date of implementation of the section selected in the 3D model. Subsequently, the executed part will receive a blue color for visual management within the Revit software. Automatically, through the plug-in, this information goes to Firebird 3.0 database and can be viewed in the planning software. Constructive steps controlled in separate files can be visualized together in graphs and spreadsheets, with dates and schedules, as well as the comparison between expected and achieved progress.

In the planning software information referring to the workforce and existing materials in the building site is entered. This control can be done weekly, daily or when necessary, allowing updating of information regarding current inputs in the building site. In this way, planning cycles are performed until services are fully completed.

Figure 2: BIM 4D Model developed in the Research



IMPLEMENTATION OF THE PROPOSED MODEL

Implementation of the proposed model occurred in four weeks, in January 2017. This temporal delimitation coincided with the completion period of monitored work packages. Throughout this period, data were collected to monitor building site progress concerning work packages performed, labor consumption and materials, to compare them with data generated by the model, to verify their potentialities and vulnerabilities. Work packages of fire installations and coating granite were studied.

In the preparation of the 3D model, weights were assigned to Revit software design parameters created to compare productivity, to control material consumption and to associate specific materials for each part of executed work.
After inserting work breakdown structure and unit compositions of each service in the planning software and defining parameters of projects used in definitions of quantities of labor and materials required in parts of work to be executed, the model BIM 3D was considered prepared to be used in the integration model proposed in this research.
There was a need to update visible progress to follow work packages applying the proposed model. Also, it was necessary to enter information on the workforce used in the planning software. For that, the working document titled "Information Diary" was analyzed, in which work of employees who performed services was systematically recorded. Information regarding materials has been entered by approved purchase orders.
The model was considered ready to run after fed with information regarding labor, materials purchased, the physical progress of parts of building site already executed, and projects details.
The first look-ahead plan performed referred to January 2017. Then, WWP and the monitoring of executed service were performed. With the application of the model in the coating granite work package, it was possible to identify it brought relevant information about the workforce, previously recorded in another worksheet, with no connection to planning. In this sense, managers were able to learn where contractor's workforce was being applied and to understand the reason for the delay in service, as well as to identify the best place to work for the team.
Constraints related to design specifications needed to be identified before the start of planning. In this sense, the model promoted the interaction of stakeholders, increasing communication between sectors of the company.
Changing in the plan of last week was only possible by identifying the causes occurred in previous weeks, providing better management and, consequently, improving the efficiency of short-term planning and making it formal, with use of the model.
With the group of services of the fire-fighting facility, it was clear the model could anticipate a lack of material and avoid delay in execution of services.
If the execution did not follow the project, the model would be useless since the report generated would not be compatible with reality. Therefore, verification of constraints in the look-ahead functionality would not be accurate either, concerning materials.
Another point verified was the number of sections of the 3D model pointed out as concluded that needed to be finished. The model had an option of registering partial advance of work packages in percentage values. This procedure could have been used to signal the formation of teams for these finalizations was necessary, and to adjust the schedule to include new work packages, or not to initiate parts of building site without completion of previous ones.

EVALUATION OF THE PROPOSED MODEL BY ITS USERS

After application of the model in the experimental study, a meeting was held with collaborators of the building site where the research took place, to evaluate its usefulness and functionality.

Usefulness evaluation

In the evaluation of the reliability of information, the authors noticed the divergence between the material requested from the supply department and the one required by the model. This problem was caused by actions that were executed deliberately different from what was predicted in the project.
Information regarding workforce proved reliable, as confirmed by building site production sector, which ratified data on productivity and quantity of employees at each work package front.

Another element that presented reliability was monitoring of parts of the building site executed, which could be identified in the model with visual management, through different colors adopted.
Increased communication was the most useful result generated by the model, according to interviewees. Improvement occurred when the project department inserted visual design changes directly on the model.
It was found those involved in planning envisaged possibility of using information integrated into a single database for management and execution of building site. Anor aspect emphasized by the planning team was the use of visual management with different colors for services executed, planned with constraints, and without constraints. Insertion of look-ahead in the planning process was also considered as positive, since it anticipated constraints, to eliminate them before the beginning of services.
Automatic search of the workforce and material needs in the model, provided by plug-in and database functionalities, was also pointed out as a facilitator of scheduling. Planning of construction site was improved since such verification was not performed systematically before.

Functionality evaluation

Visual management simplified understanding of information generated by the model since it facilitated analysis of the progress of scheduled work packages compared to executed one. Planning sector, which operated Revit software, had a very satisfactory interaction with the proposed model, presenting no difficulty in running it. However, the production sector showed no interest in using the model.
Most significant difficulties encountered in operation of the model were related to insertion of material data since different suppliers often adopted different identifications for the same input, which, in turn, were different from what was registered in geometric element modeled in Revit software. Also, it was also observed a restriction regarding the collection of labor data for each work package, since construction site management team registered only the number of employees per aggregate service, while in the model there was a need to record the number of employees per service.

FINAL REMARKS

Through proposed model, a single database gathered information necessary for LPS logic, promoting interaction between departments and stakeholders and adding information that was previously in dispersed documents and sometimes not used in the production planning and control.
The developed plug-in added functionality to Revit software, capable of promoting visual management of look-ahead planning, automatically checking constraints on the workforce, materials, and project specifications. By issuing visual alerts with different colors for parts of the building site with unrestricted tasks and other with constraints, it became possible to generate a task bank to be included in WWP short-term planning.
During the implementation of the proposal, it was possible to observe approximation and improvement of communication between departments of planning, production, projects, and supplies. The model concentrated information necessary for the formation of medium and short-term plans, which were isolated in various departments.
Automatic verification performed by installed plug-in's look-ahead function achieved constraints identification and visual management. Visual management through colors proved very attractive to users, improving interpretation of verified information.

Finally, synergy resulting from the integration of BIM 4D with LPS, through automatic verification of constraints in the design of look-ahead, visual management, and integration of involved parties with planning, generated a more efficient medium and short-term production planning and control system.

An existing limitation in integration model occurs when there is a need to change some geometric element in the 3D model at the moment of its execution. Integration model still cannot automatically promote inclusion or elimination of geometric element in its update database, requiring a manual procedure. Another limitation of the integration model is the lack of verification of constraints regarding predecessor work packages.

This research emerged questions that offer the following suggestions for future work: a) improve integration model so it can receive automatic changes in the 3D model; and b) developing applications for mobile devices (like tablets or smartphones), to assist in the collection of field data directly in the model.

REFERENCES

Ballard, G. (1997). Lookahead planning: The Missing Link in Production Control. “Proc. Annual Conference of the International Group for Lean Construction.” Gold Coast. Australia: IGLC, p. 13-26.

Ballard, G., and Howell, G. (1998). “Shielding Production: Essential Step in Production Control”. J. Constr. Eng. Manage., 124(1), https://doi.org/10.1061/(ASCE)0733-9364(1998)124:1(11).

Ballard, G. (2000). “The last planner system of production control.” Ph.D. dissertation, Univ. of Birmingham, Birmingham, U.K.

Bilal, M., Oyedele, L.O., Qadir, J., Munir, K., Akinade, O. O., Ajayi, S. O., Alaka, H. A., and Owolabi, H.A. (2016). “Analysis of critical features and evaluation of BIM software: towards a plug-in for construction waste minimization using big data.” Int. J. of Sustain. Build. Technol. and Urb. Devel., DOI: 10.1080/2093761X.2015.1116415 .

Bhatla, A., and Leite, F. (2012). Integration Framework of BIM with the Last Planner System. “Proc. of the 20th Annual Conf. of the IGLC”, USA.

Biotto, C. N. (2012). Method of management of production on civil construction with the use of modeling BIM 4D. Master thesis, Univ. Fed. Rio Grande do Sul, Porto Alegre - RS. (in Portuguese).

Dresch, A., Lacerda, D. P., and Antunes Jr, J.A.V. (2015). Design Science Research: A Method fos Science and Technology Advancement. Springer International Publishing, UK, ISBN: 978-3-319-07374-3, DOI: 10.1007/978-3-319-07374-3.

Eastman, C., Teicholz, P., Sacks, R., and Liston, K. (2011). BIM Handbook: A guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors. 2nd e. John Wiley & Sons, Inc., USA.

Fernández, A. D. R., Cárdenas, L. F. A., and Armiñana, E. P. (2011). “La gestion de la obra desde la perspectiva del ultimo planificador.” Rev. Obr. Públ., 3518, ISSN: 1695-4408.

Koskela, L. (2000). “An exploration towards a construction theory and its application to construction.” Espoo 2000. Technical Research Centre of Finland. VTT Publications 408, 296 p.

Laufer, A., and Tucker, R. L. (1987). “Is construction project planning really doing its job? A critical examination of focus, role and process.” Constr. Manage. Econ., 5(3), https://doi.org/10.1080/01446198700000023 , 243-266.

Leite, F., Cho, Y., Behzadan, A.H., Lee, S., Choe, S., Fang, Y., Akhavian, R., and Hwang, S. (2016). “Visualization, Information Modeling, and Simulation: Grand Challenges in the Construction Industry.” J. Comput. Civ. Eng., 10.1061/(ASCE)CP.1943-5487.0000604, ISSN 0887-3801, 04016035(16).

Machado, R. L., and Heineck, L.F.M. (2006). A Contribution to Production Planning Improvement in the Construction Industry through the Concept of Anticipation. “Third International Conference on Production Research”, ICPR Americas 2006, Curitiba – PR.

March, S. T., and Storey, V. C. (2008). “Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research”. MIS Quaterly, USA.

Mendes Junior, R., Scheer, S., Garrido, M. C., and Campestrini, T. F. (2014). “Integrating Construction Modeling (BIM) with Production Planning and Control”. Proc. XV ENTAC. Maceió – AL.

National Institute of Building Sciences. (2007). “National Building Information Modeling Standard - Version 1.0 - Part 1: Overview, Principles, and Methodologies”. USA.

Priven, V., and Sacks, R. (2015). “Effects of the Last Planner System on Social Networks among Construction Trade Crews.” J. Constr. Eng. Manage., 141(6), 10.1061/(ASCE)CO.1943-7862 .0000975, 04015006(10).

Rodas, I. A. R. F. (2015). Application of the BIM methodology in building management. Master thesis, Fac. Eng. Porto. (in Portuguese).

Rodrigues, P. B. F. (2017). A proposal for integrating BIM model to the Last Planner System. Master thesis, PUC Goias, Goiania - GO. (in Portuguese).

Sacks, R., et al. (2010). “Introducing a new Methodology to Develop the Information Delivery Manual for AEC Projects”. Proc. 27th International Conference – Applications of IT in the AEC Industry & Accelerating BIM research Workshop, Cairo, Egypt.

Sacks, R., and Harel, M. (2006). “An economic game theory model of subcontractor resource allocation behavior.” Constr. Manage. Econ., 24(8), https://doi.org/10.1080/01446190600631856, 869–881.

[Sacks, R., Barak, R., Belaciano, B., Gurevich, U., and Pikas, E. (2013). “KanBIM workflow management system: prototype implementation and field testing.” Lean Constr. J., 19-35.

Sacks, R., Treckmann, M., and Rozenfeld, O. (2009). “Visualization of Work Flow to support Lean Construction.” J. Constr. Eng. Manage., 135(12), https://doi.org/10.1061/(ASCE)CO.1943-7862.0000102, 1307-1315.

Sacks, R., Koskela, L., Dave, B. A., and Owen, R. (2010). “Interaction of Lean and Building Information Modeling in Construction.” J. Constr. Eng. Manage., 136(9), https://doi.org/10.1061/(ASCE)CO.1943-7862.0000203, 968-980. .

Succar, B. (2009). “Building information modeling framework: a research and delivery foundation for industry stakeholders.” Autom. Constr. 18(3), 357-375.

Van Aken, J. E. (2004). “Management research based on the paradigm of the design science, the quest for field-tested and grounded technological rules.” J. of Manage. Studies, 41(2), 219-246.

DECISION SCIENCES INSTITUTE

Drivers, Challenges, Moderators and Outcome: An End-User Perspective on Business Analytics Adoption

Zhezhu Wen
The University of Toledo
zhezhu.wen@utoledo.edu

Xiyue Deng
The University of Toledo
xiyue.deng@rockets.utoledo.edu

Suvankar Ghosh
The University of Toledo
suvankar.ghosh@utoledo.edu

ABSTRACT

The objective of this research is to understand the adoption driver, challenge and outcome of business analytics for less technically-oriented users. We identify three challenges that may have a complicated impact over adoption: cognizant engagement, bounded rationality, and meaningful use pressure. We also hypothesize IT mindfulness and knowledge acquisition efforts can mitigate challenges and lead to substantial use. The performance implications of different usage patterns are also hypothesized. We develop and validate an instrument to test our theory of business analytics adoption. The results of Q-sort revealed a hit ratio of 82% with an inter-rater reliability of 75%. The future work and the potential contributions are discussed.

KEYWORDS: Business analytics, system adoption, Beyond dominant paradigm, micro-level decision-making process, mindfulness, knowledge acquisition

INTRODUCTION

Portrayed as “the next frontier for innovation” (Shollo & Galliers, 2016), Big Data and Analytics (BDA) continues to drive imaginations among practitioners and academics. One key promise that this stream of innovation holds is its ability to create actionable insights that can inform future strategy and day-to-day operations (Grover et al., 2018). In a worldwide survey of more than 3,000 business executives and analysts revealed that the top-performing organizations are twice as likely to apply analytics to various corporate activities (LaValle et al., 2011). Influenced by success stories of early adopters and the competitive pressures, increasingly, many private and public organizations are showing their interest in implementing similar systems (Wixom & Watson, 2010). According to one statistic, the global business investment in big data infrastructure and analytics software will continue to grow and will surpass $200 billion a year by 2020 (Carande et al., 2017).

In this work, we focus on a subset of BDA – business analytics. While BDA refers to a system of infrastructure, tools (software), skillsets, and capabilities, the business analytics is concerned with “the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions" (Davenport & Harris, 2007). When compared with BDA, business analytics has more end-user implications.

The diffusion of business analytics happens both in scale and depth. Thus far, the analytics related tasks are mostly performed by professionals with an extensive computer, statistics, and quantitative modeling background. But many moves indicate that the use of business analytics will be expected from employees with no or little technical training in analytics. Gartner in a 2018 report indicated that the analytics output of business users with self-service capabilities will surpass that of professional data scientists (Violino, 2018). GM is gradually moving toward a platform where employees in many levels (e.g., executives, engineers, business analysts, etc.) can query through a Google-like search interface for information about specific business needs (Boulton, 2018). Colleges are also actively embracing the trend by revising or/and reconfiguring existing curriculums to maximize the number of students who can gain hands-on experiences in business analytics (Wilder & Ozgur, 2015). This trend of "analytical empowerment" (Alpar & Schulz, 2016) is to fully realize the value creation potential of business analytics by allowing "casual users" to engage with business analytics without adding extra burdens to already busy IT professionals (Lennerholt et al., 2018).

This changing landscape in the user base calls for more scholarly attention to the end-user adoption processes of business analytics. Thus far, majority of research in business analytics studied the organizational level phenomena, such as, firm performance implications (Müller et al., 2018; Torres et al., 2018; Song et al., 2018; Wamba et al., 2017), business value creation (Krishnamoorthi & Mathew, 2018; Xie et al., 2016; Chen et al., 2015; Grover et al., 2018), business analytics and intelligence capabilities building (IşıK et al., 2013; Gupta & George, 2016; Wang et al., 2018). Only a handful of research studied the end-user adoption-related issues (Li et al., 2013; Grublješič & Jaklič, 2015). However, the unique combination of socio-technical characteristics and occupational context that are embedded in and accompanied by the business analytics can potentially help IS researchers to expand and to enrich existing understanding about the innovation adoption processes.

In addition to well-known system adoption driver constructs such as effort and performance expectancy, a theoretical model should consider following three challenges that average users will encounter during the adoption process: 1) cognizant engagement challenge, 2) bounded rationality challenge, and 3) meaningful use challenge. The cognizant engagement challenge requires users to be able to cognitively differentiate the use cases between business analytics and other similar enterprise business intelligent systems (e.g., DBMS, SAP report, etc.). The bounded rationality challenge expects users to keep up with a substantial body of knowledge that is necessary for accurately identifying the input and output parameters (e.g., specification of distributions, coefficient significance, etc.). And lastly, as placed within the work environment, the use of the system is expected to produce meaningful outcomes (e.g., whether use the system is less important than the quality of work). Therefore, business analytics adoption research should embrace these challenges in their research design.

In this work, we attempt to develop a theoretical model and measurement instruments by extending the reasoned-behavior theory framework (Fishbein & Ajzen, 1977; Ajzen, 1991) with mindfulness, knowledge complementarity, and IS system usage literature. Through the model, we seek to capture users' underlying micro-level decision-making processes that lead to meaningful use of business analytics in the work environment. We believe our work has the following contributions: 1) our work attempts to approach the business analytics adoption processes more holistically than existing theoretical frameworks by considering task, user, and system characteristics; 2) it highlights the role of a nascent attention/state construct, namely, IT mindfulness, during the process of innovation adoption, and 3) it seeks to explore and explain the adoption beyond the dominant paradigm (Fichman, 2004).

This paper is organized in the following order: In the next section, we scan background and extant literature landscape related to business analytics adoption. Then, we present the theory development and the research model. In the subsequent section, we discuss the measurement items and the refinement process. We conclude this article after discussing future research plan and result-contingent managerial implications.

BACKGROUND AND LITERATURE REVIEW

Business Analytics and Competitive Advantage

Increasingly, firms are exposed to a competitive environment where data-driven decision-making cultures and practices are greatly rewarded (McAfee & Brynjolfsson., 2012; LaValle et al., 2011). Strategic management literature suggests that valuable, rare, inimitable, and non-substitutable (VRIN) resources and capabilities are the basis of competitive advantage and economic rent (Barney, 2001). Under this view, when a harmonious integration of firms' data infrastructure, organizational culture, and human capital is achieved, big data and analytics can become the source of competitive advantage (Erevelles et al., 2016)

Recent empirical evidence supports the view. Based on the results of an econometric analysis of 814 firms, Müller et al. (2018) report that a firm's productivity growth is positively associated with the firm's big data and analytics assets. The magnitude of increase in productivity was on average 3~7 percent each year. Based on a study of 179 sample firms from large publicly traded firms, Brynjolfsson et al. (2011) find that firms that adopt data-driven decision-making practices have output and productivity that is 5~6 percent higher than that would be expected given their other investments and information technology usage.

There are several distinct but highly related concepts in this context. The Institute for Operations Research and the Management Sciences (INFORMS) defines *analytics* as *the scientific process of transforming data into insights for the purpose of making better decisions* (Rose, 2016). *Business analytics* is concerned with "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions" (Davenport & Harris, 2007). The term *business analytics* is often associated with the term *business intelligence* (Chen et al., 2012) as both concepts are concerned with making a more informed decision based on data. According to Adair (2018), while the former is primarily more concerned with the prediction of trend/pattern and offering a prescription to a problem, the latter is primarily concerned with creating a summary of historical data. While *business analytics* is a relatively new concept, *business intelligence* is around since the 1950s. The recent advancement in data capturing and storage technology, coupled with much improved computation power, both techniques together contribute greatly in extracting insights from a large amount of structured and unstructured data.

Bughin et al. (2016) present a data and analytics transformation model with five different phases: 1) defining clear business needs, 2) gathering data from data ecosystem, 3) modeling insights, 4) redesigning of workflow, and 5) managing change for adoption. They assert the importance of the adoption by arguing that the value of analytics depends on its ultimate use.

Business Analytics Task Characteristics

Adapted from Rose (2016), we categorize users into four groups based on their strategic orientation and the primary purpose of use. This research focuses on decision-centric users whose main decision goals are related to their day-to-day operations.

Table 2.1 Areas of interests in Big data and analytics

	Data-centric camp	Decision-centric camp
Strategic decision	Chief data officer/architect	C-level executives
Operational decision	Data scientists	Data-driven users

(Source: Adapted from Rose, 2016)

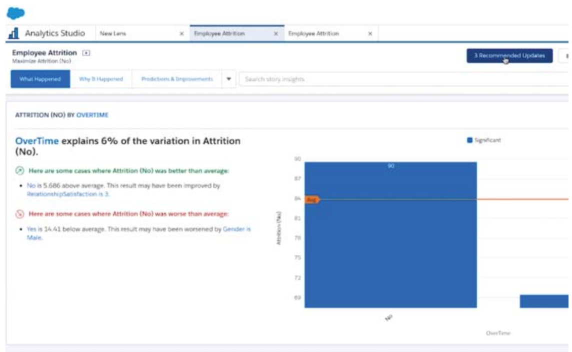
Power et al. (2018) provide interesting examples of business analytics that simulate what data-driven users may expect from using business analytics.

> "Banks and other businesses use business analytics to *classify* and *segment* customers based on credit risk, usage, and other characteristics, and use algorithms to *match* customer characteristics with appropriate product offerings. An airlines manager who is considering investing in a new route uses analytics to *predict* future travel demand for the destination. Managers prepare revenue *forecasts*. Operations analytics might *examine* product cost, quality control, and the throughput of resources in production lines. Machine *monitoring* and machine failure *tracking* are common. *Optimizing* production and work schedules has been ongoing for many years. … Marketing groups use *predictive analytics* for sales lead scoring. Patient costs and care metrics are *monitored* in many hospitals." (Power et al., 2018, pp.47)

These examples illustrate the predictive and prescriptive nature of business analytics. Solution providers and firms' IT departments are gradually introducing more "*user-friendly*" business analytics suites over the years. *Einstein Analytics* by salesforce.com *Cognos Analytics (formerly Watson Analytics)* by IBM, *Business Intelligence* by Oracle are those examples. These solution suites serve as *intelligent*, *self-service analytics platform* by providing an intuitive dashboard and *what-if* analysis features through which users can capture the status of their business and identify upcoming patterns. This move will allow more users to engage with business analytics without relying too much on data scientists or IT professionals. Some solutions like *Oracle Analytics* also offer user adoption status monitoring function to help organizations to maximize the value of analytics implementation. According to one survey involved 2,679 firms, *self-service business analytics* is considered the third most important trend among the respondents (Durcevic, 2018).

Figure 2.1 *Einstein Analytics* helps to identify a significant factor

Figure 2.2 *Oracle Sales Analytics* helps to provide a market win/loss analysis



(Source: *Salesforce.com*)

(Source: Oracle.com)

Business Analytics and Adoption Challenges

Although the *self-service analytics* efforts lower the level of technical knowledge to utilize business analytics, the acceptance and usage process will not be totally frustration-free. Below, we synthesize three challenges that average workers will encounter during the use of business analytics. Burton-Jones and Grange (2012) emphasize three key areas that are concerning the effective use of a system: users' competence and challenges, task structure, and system characteristics. Our following synthesis seeks to understand these three areas.

The *cognizant engagement* challenge arises when users attempt to determine whether the use of business analytics platform is a necessary condition for the given task. For the average users, the use of business analytics is not a necessary condition to complete the traditional lines of task, but the information is. Business analytics is a means for ends. There will be established ways of gathering information such as making phone calls to the floor level managers, digging through a spreadsheet to identify a pattern, and attending the staff meeting to discuss an issue. When a user does not know how to take full advantage of business analytics, he/she may under-use, over-use, or even abuse business analytics systems. Users can cope with this challenge only when they are clearly aware of the strength of business analytics and its difference with other similar enterprise systems.

Bounded rationality is another important challenge that users will need to cope with. Even when a user has a firm belief that business analytics is appropriate for the given task, he/she may still find it challenging to navigate through the mountain load of information being displayed throughout the use process. Especially, jargon, terminologies and underlying concepts will continue to challenge the user's limited knowledge base. An analytical process typically relies on data, model, and algorithm to deliver a solution. Considering the sheer number of options and variations tied to each artifact, the number of factors the user needs to consider is combinatorially limitless. Another area of the challenge will arise when the user has to confront the results delivered by an algorithm that he/she is not sufficiently familiar. For professionals, additional robustness checking process can add confidence to the results, but for decision-driven average users, they will need to carefully make sense of the results within his/her understanding.

The *meaningful use* challenge is relevant because the use of the system is not the ultimate measure for his/her job performance. An employee is judged by job performance. In other words, the use of business analytics is not a sufficient condition to receive favorable evaluations from the superior, but an informed decision is. After all, a user will have to find a meaningful way to use the system to ensure the quality of the decision.

In sum, users will feel *cognizant engagement* challenge when determining *whether* to use or not. They will also experience *bounded rationality* challenge when seeking how to use it properly. And lastly, they will feel *meaningful use* challenge when they realize *what* to deliver is the ultimate concern, not the use of business analytics.

Adoption Decision as Micro-level Decision-Making Process

Although many scholars have studied business intelligence systems acceptance (Li et al., 2013; Grubliešič & Jaklič; Yoon et al., 2017), business analytics research remains as a vacuum.

Grubliešič and Jaklič (2015) conducted a business intelligence acceptance study using case study methods. Their conceptual framework highlights the importance of organizational factors, such as result demonstrability, social influence, and facilitating conditions with sufficient resources that help create an adequate information culture. The artifact in their study is ERP system. Yoon et al., (2017) investigate the end-user acceptance of business intelligence applications by using survey data collected among ERP system users in large to medium-sized organizations. Their study found that individual motivation to learn determines the usage intention. Also, they identified the relative advantage, situational constraints, and organizational learning climate serve as the driver of individual motivation to learn. Li et al. (2013) identify perceived usefulness, intrinsic motivation toward accomplishment, intrinsic motivation to know, and intrinsic motivation to experience stimulation are the antecedents of system use. Their work highlights the importance of personal innovativeness with IT as the moderator that amplifies paths to innovative use.

Business intelligence system shares many similarities with business analytics, but we see the needs of complementing acceptance studies from both areas as business intelligence is primarily descriptive in nature, and business analytics seeks to conduct predictive and prescriptive analytics. We also felt the needs of enriching existing constellation (Yoon et al., 2017 provide a good summary of the list) of antecedents of acceptance by adding constructs that help address users' micro-level decision-making process (Fiol and O'Conneor, 2003).

Recently, the academic discourse of being mindful (Dane, 2011) has expanded its reach to information systems research (Sun et al., 2016; Thatcher et al., 2018). Essentially, mindfulness is a psychological state of consciousness. Dane (2011) defines it as a state of consciousness in which attention is focused on present moment phenomena occurring both externally and internally. Therefore, mindfulness in its original form is a *state* or an *attention* construct (not intention or traits). When mindfulness states are activated, people will attempt to identify novel distinctions of the current artifact within the current context. Therefore, it leads to a number of consequences: 1) a greater sensitivity to one's environment, 2) more openness to new information, 3) the creation of new categories for structuring perception, and 4) enhanced awareness of multiple perspectives in problem-solving (Langer & Moldoveanu, 2000; Dane, 2011). The mindfulness concept has been studied under many contexts and it generally improves subjects' health, creativity, productivity, learning outcome, and decision-making quality (Langer & Moldoveanu, 2000).

In IS research, Sun et al. (2016) seem to be the first group of researchers who adapted the concept in the adoption research. They specified the mindfulness concept as a second-order construct with four sub-dimensions and named it as MTA (mindfulness technology adoption. Together with TTF (task-technology fit), they formulated an MTA-TTF framework through which users' adoption and continued use decisions are observed. Thatcher et al. (2018) followed similar formative-measure of mindfulness and used it to predict users' active and automatic system use. Both their studies use a fairly simple technology artifact: wiki and office suites.

THEORETICAL DEVELOPMENT/MODEL

User's Reaction and Usage Intention

In this research, we see users as rational beings who adjust behavioral intentions based on the various responses within the social context. There are three specific layers of social contexts we would like to specify in our model: macro-, meso-, and micro-level contexts. Prior to the actual use, the macro-level social context influence users' perceptions on the ease of use (or effort expectancy) and the usefulness (or performance expectancy) through media, press, word of mouth, or sometimes through knowledge-entrepreneurs (Fichman, 2004). The significance of these two constructs are well-established through seminal works of Davis (Davis, 1989) and Venkatesh (Venkatesh & Davis, 2000; Venkatesh et al., 2003; Venkatesh et al., 2012).

In the meso-level context, a user's usage intention will be influenced by the firm's strategic approach with regard to the use of business analytics. If a firm has developed a significant business analytics capability through implementing effective data governance policy, business analytics policy, and decision-making culture that is favorable to the use of business analytics (Gupta & George, 2016), the user will demonstrate a higher level of usage intention. In the micro-level, a user will adjust his/her usage intention based on the amount of pre-requisite knowledge he/she possesses, whether the use of BA is consistent with the career goal and whether the use of BA enhances his/her image within the peers (Moore and Benbasat, 1991).

Vroom's expectancy theory (1964) suggests that the level of motivation is the product of valence (i.e., attractiveness of the outcome), expectancy (i.e., a person's belief that he/she will be able to reach the desired outcome) and instrumentality (i.e., a belief that a strong performance will be rewarded) (Vroom, 1964). Therefore, our explanation on the formation of a user's intention in meso-level and micro-level context seem plausible.

Thus, we hypothesize,

- *H1. (a) Effort expectancy and (b) performance expectancy are positively associated with the intention to use.*
- *H2. The level of organizational BA maturity is positively associated with the intention to use.*
- *H3. The level of individual motivation is positively related to the intention to use.*

For the measurement, we specify BA maturity as a formative measure of data governance, BA policy, and BA culture, and specify individual motivation as a reflective measure of the level of pre-requisite knowledge, career path congruence, and image.

Usage Patterns and Performance

To understand the impact of usage on the job performance, following Delone and McLean (1992), we conceptualize the usage as an antecedent of job performance. This connection between the *drivers of use* and the *impact of use* will help picture the complete adoption process.

Burton-Jones and Straub (2006) assert researchers measured the usage in various ways including binary scale (use/not use), continuous scale (frequency, extent), or nominal scale (nature of use), and the measurement of usage should be specific to the context of research. In this research, we envision two different types of usage: *substantial use* and *symbolic use*. In the former, a user's engagement with a system is more substantial and is considered more meaningful from performance achieving perspective. In the latter, the usage pattern is more routinized, superficial, and passive. Saeed and Abdinnour (2011) modeled a similar idea using a linear path of usage stage: routinization (Users use the various features that IS has to offer) to infusion (users integrate the IS in their work) to extension (user explore the potential of the IS in novel contexts). But, following Li et al., (2013), we differentiate *symbolic use* and *substantial use* to highlight the qualitative, not quantitative, differences between two type of usage.

Thus, we hypothesize,

- *H4. There will be two distinct usage patterns: substantial and symbolic use*
- *H5. The substantial use will have a greater contribution to the task outcome than symbolic use.*

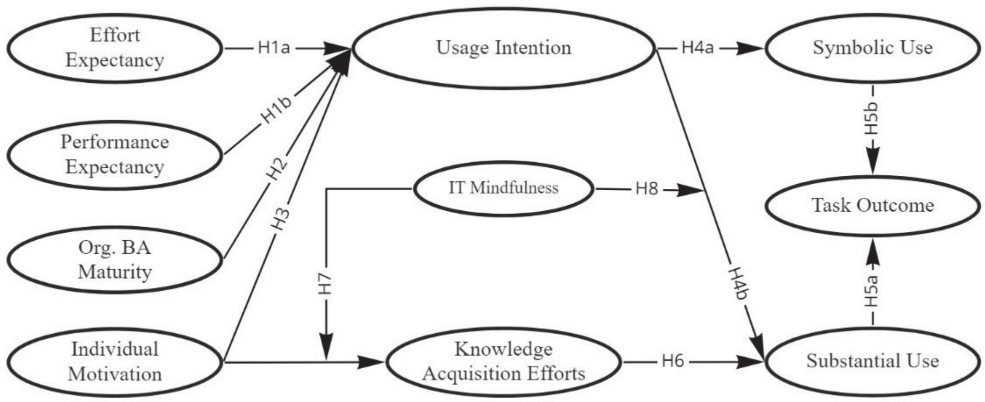
Knowledge Complementarity (Knowledge Acquisition Efforts)

Due to the complicated nature of analytics tasks, a user will find his/her current level of knowledge is inadequate. Cohen & Olsen (2015) asserts that both codified knowledge management capability and employee's knowledge management capability serve as the antecedents of service outcome. Their view is consistent with what innovation management literature sees about the knowledge complementarity, that is internal and external knowledge together form a “whole” knowledge that contributes to the innovation outcome (Fang, 2011; Cassiman & Veugelers, 2006). Thus, we believe when a critical piece of knowledge flowed through users' efforts, it will contribute toward more meaningful use.

Thus, we hypothesize,

- *H6. Knowledge acquisition efforts will mediate the relationship between the level of individual motivation to substantial use.*

Figure 3.1 Research Model



IT Mindfulness

To capture the users' micro-level psychological engagement during the adoption process, we incorporate the notion of mindfulness in this research. Followed by Thatcher et al. (2018) and Sun et al. (2016), we name it as IT mindfulness and specify the construct as a formative measure with four sub-dimensions. Based our existing understanding about the role of mindfulness (Dane, 2011; Langer & Moldoveanu, 2000), we believe a user with a higher level of IT mindfulness will have better ability to amplify its usage intention. Thus, the path under the influence of IT mindfulness will be more likely leading to a substantial use while the path without the IT mindfulness will lead to a symbolic use. We also believe it amplifies the users' motivation to engage with business analytics, so that the users will increase their knowledge acquisition efforts.

Thus, we hypothesize,

- *H7. The level of a user's IT mindfulness will positively moderate the relationship between individual motivation and knowledge acquisition efforts.*
- *H8. The level of a user's IT mindfulness will positively moderate the relationship between the usage intention to substantial use.*

Since mindfulness is a state and/or attention construct, it does pose some challenges in terms of measuring with a satisfactory level of validity in empirical research. However, since there are people who tend to exercise mindfulness more often (Dane, 2011), and also the state is malleable through training (Thatcher et al., 2018), it is also appropriate to measure it as a trait construct.

Table 3.1 List of Key Constructs

Construct	Definition	References
Effort Expectancy	The degree to which a person believes that using a system would be free of effort.	Davis, 1989

Performance Expectancy	The degree to which a person believes that using a particular system would enhance his or her job performance.	Davis, 1989
Organizational BA Maturity*	An organization's overall readiness in leveraging from business analytics through the key artifact (data), surrounding processes, and supportive organizational culture.	
Individual Motivation	The extent to which an individual is motivated to use BA.	
Usage Intention	Level of intention to use the system.	
Knowledge Acquisition Efforts	The level of efforts to obtain relevant information to use BA from outside of the work context.	
IT Mindfulness*	The degree to which a person compares the technology with existing technologies so that the individual is more aware of its uniqueness.	Sun et al., 2016
Symbolic Use	A usage pattern that is passive and is reluctant to explore beyond initial instruction.	
Substantial Use	A usage pattern that is active and seeks to maximize the returns from using BA.	
Task Outcome	Immediate task outcome from using the BA.	

(Notes: * indicates a higher-order construct)

Figure 3.1 exhibits our research model and table 3.1 lists all the key constructs of the model. Next, we discuss the measurement development and the result of validity test.

INSTRUMENT DEVELOPMENT

Since many constructs used in this research do not have readily available measurements, one of the major objectives of this research is to develop valid and reliable measurement instruments for the constructs used in the proposed research model. To ensure the validity of constructs and identify ambiguity in items, three steps of the instrument development process were conducted: 1) item generation, 2) pre-test of content validity and 3) Q-sort study. Specifically, items generation is a process that produces measurement items for newly developed constructs through a comprehensive literature review. The pre-test of content validity has been done through a structured interview with scholars from the information systems field. The Q-sort study is used to ensure the convergent and discriminant validity of these measurement instruments (Churchill, 1979; Moore & Benbasat, 1991).

Item Generation

In this research, there are a total 10 of the main constructs, two constructs are the higher-order constructs which formed by second-order constructs, and 65 measurement items generated through a comprehensive literature review. Measurement items were developed through A combination method of self-developing and adapting of existing research. Initial items were tested through pre-test of content validity and Q-sort study.

Pre-test

In order to ensure content validity, a group of academic experts was invited to pre-test the initial measurement instruments. They were asked to evaluate the appropriateness, clarity, and accuracy of the instruments. We analyze their thoughts and suggestions by the discrepancies between their interpretations and our expectations. Based on their feedback, the initial measurement items were modified.

Q-Sort

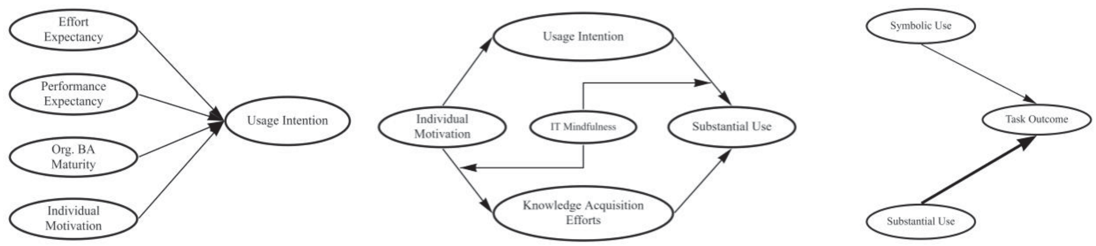
In this research, the Q-sort method was employed to develop the measurement instrument, which is widely used in psychology and other social sciences disciplines to investigate the individual's subjectivity. A structured interview was conducted among two information technology scholars. In the beginning, the general research idea, the conceptual model and definitions of each construct were presented to the interviewees. Then, the interviewees were given a set of randomized index cards that contain the candidate items. Interviewees were asked to categorize these cards based on the construct definitions. A *non-applicable* category is provided so that the interviewees can store the items that they think do not belong to any of the provided constructs. The first-round inter-rater reliability is 53%. A further discussion of the discrepancy between our expectations and interviewee's interpretation was conducted. Based on the discussion and communication, measurement items were modified, deleted, or combined. The second iteration of Q-sort was done after a week of the first iteration. The inter-rater reliability was increased to 75%. Cohen's Kappa method is the most robust statistical method that has been widely used to test the measurement of inter-rater agreement (Cohen, 1960; Cohen 1968). As suggested by Landies and Koch (1997), the coefficient greater than 0.6 is considered as substantially acceptable (Table A.1). Cohen's Kappa coefficient based on the analyze of second-round Q-sort results is 0.728 for this study.

FUTURE WORK AND DISCUSSION

We are planning to conduct a pilot test among a small group of people. In the full test, it is preferable to have two separate sources of respondents to avoid common method bias: 1) managers or supervisors can answer the organization-related questions such as level of BA maturity and task outcome about a user; 2) their subordinates who are the decision-driven users of business analytics can answer the rest of the survey. This paired sample approach can greatly increase the rigor of the research.

If the hypothesized model receives empirical support, then this research will be able to provide rich insights regarding the end-user adoption of business analytics systems. Those insights can be categorized into three areas: 1) the factors (figure 5.1) that determine the general level of usage intentions, 2) the mechanisms (figure 5.2) through which substantial use is facilitated, and 3) the usage pattern that is more desirable (figure 5.3). Managerial implications pertaining to the first and the third area are clear and intuitive.

Figure 5.1	Figure 5.2	Figure 5.3
Key antecedents to intention	Mechanisms to substantial use	Desirable usage pattern



Along with users’ perception (i.e., effort and performance expectancy) of the technical aspect of a business analytics system, organizational readiness and individual’s motivations also contribute to the formation of users’ usage intention.

Business-analytics-friendly *data practices*, well-defined *organizational processes*, and supportive *decision-making culture* will create an environment conducive to the use of business analytics by lowering users’ interaction cost. This line of thought encourages the management to review the current states of readiness in data management and analytics related practices. This necessitates the existence of a well-developed business analytics maturity model (BAMM) or a similar normative framework that can help management identify the firm’s current standing and/or future directions. In the literature, there is much discourse, both from practitioners and academics, regarding such frameworks. For details, please see Halper & Stodder (2014), Comuzzi & Patel (2016), Rajterič (2010), and Cosic et al. (2012). Meanwhile, a user may be motivated to use a business analytics system independent of a system’s technical characteristics and surrounding organizational environment. This research attempts to establish three antecedents that could be used to predict the formation of individuals’ motivation. These three antecedents together demonstrate the path-dependent nature of such motivation: 1) skillsets a user accumulated from the past, 2) incentives that could stimulate the user’s behavior today, and 3) the career planning that focuses on long-term future will all together determine the level of motivation to engage with business analytics systems. Business analytics workforce management should become aware of this fact.

Overall, the first area of insights emphasizes the fact that the formation of a user’s usage intention is influenced by multiple forces; thus, it suggests that management should use more holistic perspectives in terms of planning for the end-user adoption.

The third area of insights is based on the anticipation that the path coefficient from *substantial use* to *job performance* will be much greater than that from *symbolic use*. Supply chain research often suggests that “last mile delivery” is the constraint or the bottleneck of the entire system performance (Boyer et al., 2009; Hübner et al., 2016). Likewise, in this research, we conceptualize that the end-user adoption is “the last mile delivery” in terms of business analytics value creation process. The key take-away for managers, those who work through other people and coordinate (Robbins et al., 2008), is that the insight clarifies the area that needs constant attention from management – *it is to help users facilitate and achieve substantial use of business analytics system*.

The second area of insights in this research seeks more attention from management as it is related to a nascent field – *mindfulness*. More attention is needed because the *attitudinal/state nature* makes it somewhat abstract and ambiguous, yet it carries much performance implications. The concept of mindfulness has several overlapping dimensions across many fields, including management, education, clinical therapy, Buddhism philosophy, and religious studies (Dane, 2011; Sun et al., 2016). Historically, mindfulness is seen as cultivated through meditative practice. Therefore, confusion between mindfulness and meditation arises as people often use both terms interchangeably (Conze, 2003; Dane, 2011).

However, as demonstrated in this research design, we follow Sun et al. (2016) and Thatcher et al. (2018) by implementing it as a conscientious attitude/psychological states construct. These efforts somewhat demystify "Zen-like" appearance of the mindfulness concept. Essentially, in this research design, the practical utility of mindfulness emanates from being a moderation factor: 1) It augments users' intention to engage with the system, and 2) it amplifies users' motivation to spend greater efforts in seeking the necessary knowledge to use the system. Then the relevant question will be: *How to induce such attitude/states* in *the current context*?

In business research in general, and the IS field in particular, the systematic activation mechanisms of mindfulness are an area that is open for on-going exploration. Thus, based on the assumptions, definitions and dimensions used throughout this research, we propose management to review the following three ways that firms can help users to be more mindful in using business analytics system. Although we anticipate that mindfulness will lead to substantial use, we also believe that not all users will demonstrate mindfulness. Thus, our following discussion presumes the fact.

1) *Activate through general mindfulness*: Firms in a fast-moving environment have implemented mindfulness therapy and meditation sessions to help employees overcome stress and become more productive (Levin, 2017; Dane 2011). Since mindfulness is malleable (i.e., can be changed through training) and is also enduring (Thatcher et al., 2018), general mindfulness training may have a positive impact over domain-specific mindfulness (e.g., IT mindfulness in this research).
2) *Activate through substitute mechanisms*: The key role IT mindfulness plays during the adoption process is through invoking constant awareness of the distinctive characteristics of business analytics given the current context. Inspired by social cognitive theory (Bandura, 2001), we propose that firms implement business analytics usage experience sharing programs through online or offline media to help users become aware of how other people are using the system. This effort not only helps users learn from others but also makes them aware of their own usage patterns. We expect this substitute mechanism may help active mindfulness.
3) *Activate through system design and technical features*: Firms can request the developers of business analytics systems to implement technical features that can help users to be more mindful.

CONCLUSION

In this work, we attempt to model the drivers, challenges, coping mechanisms and outcomes of business analytics adoption. We identify three levels of drivers. Our research design seeks to understand users' micro-level psychological engagement process when they encountered challenges while using the business analytics platform. We hypothesize IT mindfulness and knowledge acquisition efforts can help users cope with the challenges effectively. Such mechanisms help users to form a meaningful use pattern that precedes a superb task outcome.

To re-iterate our contributions in innovation adoption research: 1) our work attempts to approach the business analytics adoption processes more holistically than existing theoretical frameworks by considering task, user, and system characteristics; 2) it highlights the role of a nascent attention/state construct, namely, IT mindfulness, during the process of innovation adoption, and 3) it seeks to explore and explain the adoption beyond the *dominant paradigm*, and lastly 4) we develop and validate instruments that can be used to test hypothesized relationships.

APPEDIX

Table A.1 Cohen's Kappa value and the Interpretation

Cohen's Kappa	Meaning
<0	No agreement
0-0.19	Poor agreement
0.20-0.39	Fair agreement
0.40-0.59	Moderate agreement
0.60-0.79	Substantial agreement
0.80-1.00	Almost perfect agreement

(Adapted from Landis and Koch, 1977)

Table A.2 Measurement Items

Effort Expectancy and Performance Expectancy

Constructs	Code	Measurement Items
Effort Expectancy	EE1	Learning to operate the system would be easy for me.
	EE2	I would find it easy to get the system to do what I want it to do.
	EE3	My interaction with the system would be clear and understandable
	EE4	I would find the system to be flexible to interact with.
	EE5	It would be easy for me to become skillful at using the system.
	EE6	I would find the system easy to use.
Performance Expectancy	PE1	Using the system in my job would enable me to accomplish tasks more quickly.
	PE2	Using the system would improve my job performance.
	PE3	Using the system in my job would increase my productivity.
	PE4	Using the system would enhance my effectiveness on the job.
	PE5	Using the system would make it easier to do my job.
	PE6	I would find the system useful in my job

Organizational BA Maturity (A Second-order Construct)

Construct	First-Order	Code	Measurement Items
Organizational BA maturity	Data Governance	DG1	Our organization has a well define data gathering policy.
		DG2	Our organization pays great attention to data gathering process.
		DG3	Our organization knows the importance of good quality data.
		DG4	Our organization has a well-defined administrative policy with regard to the collection and the use of data.
	BA Policy	BAP1	In our organization, the policy is clear on how to initiate a BA related project.
		BAP2	In our organization, the policy is clear on how to secure a budget for a BA related project.

		BAP3	In our organization, the policy is clear on how to be coordinated for a BA related project.
		BAP4	In our organization, the policy is clear on the reporting structure on a BA related project
	BA Culture	BAC1	In our organization, people generally believe BA is important for our business.
		BAC2	In our organization, executives generally believe BA is important for our business.
		BAC3	In our organization, functional managers generally believe BA is important for our business.
		BAC4	In our organization, colleagues generally aware of the importance of BA.

Individual Motivation

Construct	Code	Measurement Items
Individual Motivation	IM1	I am motivated in using BA on my job.
	IM2	I want to learn more about BA.
	IM3	I wish I can be more fluent in using BA.
	IM4	I am looking for opportunities to improve my BA skills whenever possible.

IT Mindfulness

Construct	First-Order	Code	Measurement Items
IT Mindfulness	Technological Novelty Seeking	ITM1	I paid attention to differences of this new technology from any other technology I previously used.
		ITM2	I tended to figure out how this tool was unique in relation to the tools that I am currently using.
		ITM3	I was mindful about how this tool differs from similar tools I had used.
	Engagement with the Technology	EWT1	I was engaged in investigating this tool when making the adoption decision.
		EWT2	I gathered factual information about this tool before making the adoption decision.
		EWT3	I got involved in exploring this tool before I adopted it.
	Awareness Of Local Contexts	ALC1	When making the decision to adopt this tool, I thought about how this tool might help my work.
		ALC2	When making the decision to adopt this tool, I thought about how this tool might change the way my work was done.
		ALC3	When making the decision to adopt this tool, I thought about how this tool may be compatible with my work requirements
	Cognizance Of Alternative Technologies	CAT1	I attended to alternative views regarding the tool before making the adoption decision.
		CAT2	I was aware of other tools than this tool before deciding to adopt it.

		CAT3	I paid attention to equivalent tools to fulfill my needs before deciding to adopt this tool.
		CAT4	I thought about alternative tools to address my demands when deciding to adopt this tool.

Usage Intention

Construct	Code	Measurement Items
Usage Intention	UI1	I intend to use BA to accomplish my work
	UI2	To the extent possible, I would use BA to do different things.
	UI3	I intent to use BA as much as I can in my work

Knowledge Acquisition Efforts

Construct	Code	Measurement Items
Knowledge Acquisition Efforts	KAE1	I reach out to vendors when I find the generic guideline is not clear.
	KAE2	I search internet when I find the vendor solution is unconvincing.
	KAE3	I find the vendor's training uses example that is too simple.
	KAE4	I find the real-world situation is much more sophisticated than the vendor's training materials.
	KAE5	I will use external sources of knowledge (other than the vendor) to solve a problem that I encountered while using BA.

Symbolic Use and Substantial Use

Construct	Code	Measurement Items
Symbolic Use	SU1	When I use BA, in general, I often navigate randomly to see if there is anything that matches with the given task.
	SU2	When I use BA, in general, I feel I have a groundless hope that it can do some magic for me.
	SU3	When I use BA, I often find myself wondering without knowing what to do.
	SU4	When I use BA, I use vendor provided standard procedure of the BA.
	SU5	When I use BA, I use it in the way the vendor trained me.
	SU6	When I use BA, I use the BA features and functions that are familiar to me.
Substantial Use	SBU1	When I use BA, I first make clear what is the nature of my inquiry through BA.
	SBU2	When I use BA, I purposefully navigate the interface to look for something I need for that task.
	SBU3	When I use BA, I put a greater emphasis on understanding the problem than worry about software.
	SBU4	When I use BA, in general, I use it to complement my educated guess.
	SBU5	When I use BA, I only use it when I believe it is more meaningful than other tools available.
	SBU6	When I use BA, I use it in novel ways to support my work.

Task Outcome

Construct	Code	Measurement Items
Task Outcome	TO1	My supervisors perceive my work outcome through BA reliable.
	TO2	My supervisors use my work outcome through BA in their decision-making process.
	TO3	My supervisors often find surprising insights from the work I generated through BA.
	TO4	My BA related tasks now have better quality because of BA.

References

Adair, B. (2018). The difference between business intelligence vs. business analytics solutions, *Selecthub.com,* URL: https://selecthub.com/business-intelligence/business-intelligence-vs-business-analytics/, accessed: May 10, 2019

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Alpar, P., & Schulz, M. (2016). Self-service business intelligence*. Business & Information Systems Engineering*, *58*(2), 151-155.

Barney, J. B. (2001). Is the resource-based “view” a useful perspective for strategic management research? Yes. *Academy of management review*, *26*(1), 41-56.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual review of psychology*, *52*(1), 1-26.

Brynjolfsson, E., Hitt, L., & Kim, H. (2011). How does data-driven decision-making affect firm performance. In *Workshop for Information Systems and Economics at the 9th Annual Industrial Organization Conference, April* (pp. 8-10).

Boulton, C. (2018). At GM, self-service analytics drives business results, CIO.com, URL: https://www.cio.com/article/3289771/at-gm-self-service-analytics-drives-business-results.html, accessed: April 25, 2019

Boyer, K. K., Prud'homme, A. M., & Chung, W. (2009). The last mile challenge: evaluating the effects of customer density and delivery window patterns. *Journal of business logistics*, *30*(1), 185-201.

Burton-Jones, A., & Grange, C. (2012). From use to effective use: a representation theory perspective. *Information systems research*, *24*(3), 632-658.

Burton-Jones, A., & Straub Jr, D. W. (2006). Reconceptualizing system usage: An approach and empirical test. *Information systems research*, *17*(3), 228-246.

Bughin, J., Manyika, J., Woetzel, J. (2016). The age of analytics: competing in a data-driven world, *McKinsey global institute*

Cassiman, B., & Veugelers, R. (2006). In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management science*, *52*(1), 68-82.

Carande, C., Lipinski, P., & Gusher, T. (2017). How to integrate data and analytics into every part of your organization. *Harvard Business Review*.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, *32*(4), 4-39.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4).

Cosic, R., Shanks, G., & Maynard, S. (2012, January). Towards a business analytics capability maturity model. In *ACIS 2012: Location, Location, Location: Proceedings of the 23rd Australasian Conference on Information Systems 2012*(pp. 1-11). ACIS.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46.

Cohen, J. (1968). Weighed Kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70 (4), 213–220.

Cohen, J. F., & Olsen, K. (2015). Knowledge management capabilities and firm performance: A test of universalistic, contingency and complementarity perspectives. *Expert Systems with Applications*, *42*(3), 1178-1188.

Comuzzi, M., & Patel, A. (2016). How organisations leverage big data: A maturity model. *Industrial management & Data systems*, *116*(8), 1468-1492.

Conze, E. (2003). *Buddhist meditation*. Courier Corporation.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. Journal of Business Research, 70, 379–390. doi:10.1016/j.jbusres.2016.08.011

Dane, E. (2011). Paying attention to mindfulness and its effects on task performance in the workplace. *Journal of management*, *37*(4), 997-1018.

Davenport, T. H. & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Boston, MA: Harvard Business Press.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, *3*(1), 60-95.

Durcevic, S. (2018). Top 10 analytics and business intelligence trends for 2019. *The datapine blog*, URL: https://www.datapine.com/blog/business-intelligence-trends/, Accessed: 05/11/2019

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904.

Fang, E. (2011). The effect of strategic alliance knowledge complementarity on new product innovativeness in China. *Organization Science*, *22*(1), 158-172.

Fichman, R. G. (2004). Going beyond the dominant paradigm for information technology innovation research: Emerging concepts and methods. *Journal of the association for information systems*, *5*(8), 11.

Fiol, C. M., & O'Connor, E. J. (2003). Waking up! Mindfulness in the face of bandwagons. *Academy of management review*, *28*(1), 54-70.

Fishbein, M., & Ajzen, I. (1977). Belief, attitude, intention, and behavior: An introduction to theory and research.

Grover, V., Chiang, R. H., Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, *35*(2), 388-423.

Grublješič, T., & Jaklič, J. (2015). Business intelligence acceptance: The prominence of organizational factors. *Information Systems Management*, *32*(4), 299-315.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, 70, 308–317. doi:10.1016/j.jbusres.2016.08.004

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Halper, F., & Stodder, D. (2014). TDWI analytics maturity model guide. *TDWI Research*, 1-20.

Hübner, A., Kuhn, H., & Wollenburg, J. (2016). Last mile fulfilment and distribution in omni-channel grocery retailing: a strategic planning framework. *International Journal of Retail & Distribution Management*, *44*(3), 228-247.

IşıK, Ö., Jones, M. C., & Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management*, *50*(1), 13-23.

Krishnamoorthi, S., & Mathew, S. K. (2018). Business analytics and business value: A comparative case study. *Information & Management*, *55*(5), 643-666.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33 (1), 159–174.

Langer, E. J., & Moldoveanu, M. (2000). The construct of mindfulness. *Journal of social issues*, *56*(1), 1-9.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, *52*(2), 21.

Lennerholt, C., van Laere, J., & Söderström, E. (2018). Implementation challenges of self service business intelligence: A literature review. In *51st Hawaii International Conference on*

System Sciences, Hilton Waikoloa Village, Hawaii, USA, January 3-6, 2018 (Vol. 51, pp. 5055-5063). IEEE Computer Society.

Levin, M. (2017). Why Google, Nike, and Apple Love Mindfulness Training, and How You Can Easily Love It Too. *Inc. com, June*, *12*.

Li, X., Hsieh, J. P. A., & Rai, A. (2013). Motivational differences across post-acceptance information system usage behaviors: An investigation in the business intelligence systems context. *Information systems research*, *24*(3), 659-682.

McAfee, A., Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-68.

Moore, G. C. & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. Information Systems Research, 2(3), 192–222.

Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. *Journal of Management Information Systems*, *35*(2), 488-509.

Power, D. J., Heavin, C., McDermott, J., & Daly, M. (2018). Defining business analytics: an empirical approach. *Journal of Business Analytics*, *1*(1), 40-53.

Rajterič, I. (2010). Overview of business intelligence maturity models. *Management: Journal of Contemporary Management Issues*, *15*(1), 47-67.

Robbins, S. P., DeCenzo, D. A., & Coulter, M. K. (2008). *Fundamentals of management: Essential concepts and applications*. Upper Saddle River, NJ: Pearson Prentice Hall.

Rose, R. (2016). Defining analytics: A conceptual framework. *OR/MS Today*, *43*(3).

Saeed, K. A., & Abdinnour, S. (2013). Understanding post-adoption IS usage stages: an empirical assessment of self-service information systems. *Information Systems Journal*, *23*(3), 219-244.

Shollo, A., & Galliers, R. D. (2016). Towards an understanding of the role of business intelligence systems in organisational knowing. *Information Systems Journal*, *26*(4), 339-367.

Song, P., Zheng, C., Zhang, C., & Yu, X. (2018). Data analytics and firm performance: An empirical study in an online B2C platform. *Information & Management*, *55*(5), 633-642.

Sun, H., Fang, Y., & Zou, H. M. (2016). Choosing a fit technology: Understanding mindfulness in technology adoption and continuance. *Journal of the Association for Information Systems*, *17*(6), 2.

Thatcher, J. B., Wright, R. T., Sun, H., Zagenczyk, T. J., & Klein, R. (2018). Mindfulness in information technology use: definitions, distinctions, and a new measure. *MIS Quarterly*, *42*(3), 831-847.

Torres, R., Sidorova, A., & Jones, M. C. (2018). Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. *Information & Management*, *55*(7), 822-839.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, *36*(1), 157-178.

Violino, B. (2018). 7 keys to self-service BI success, *CIO.com,* URL: https://www.cio.com/articles and articles /3300559/7-keys-to-self-service-bi-success.html accessed: April 25, 2019

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*(1), 64-79.

Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education*, *15*(2), 180-187.

Wixom, B., & Watson, H. (2010). The BI-based organization. *International Journal of Business Intelligence Research (IJBIR)*, *1*(1), 13-28.

Vroom, V. H. (1964). *Work and motivation* (Vol. 54). New York: Wiley.

Xie, K., Wu, Y., Xiao, J., & Hu, Q. (2016). Value co-creation between firms and customers: The role of big data-based cooperative assets. *Information & Management*, *53*(8), 1034-1048.

Yoon, T. E., Jeong, B. K., & Ghosh, B. (2017). User acceptance of business intelligence application: motivation to learn, technology, social influence, and situational constraints. *International Journal of Business Information Systems*, *26*(4), 432-450.

DECISION SCIENCES INSTITUTE

Citation Analysis of Operations Management Journals

Charles Petersen
Northern Illinois University
Email: cpetersen@niu.edu

Xinyan Cao
Northern Illinois University
Email: xcao@niu.edu

Gerry Aase
DePaul University
Email: gaase@depaul.edu

ABSTRACT

Several published studies have ranked journals based on perceived quality according to operations management (OM) researchers. This study presents a citation analysis of OM research articles published in *International Journal of Operations and Production Management, Journal of Operations Management*, *Manufacturing and Service Management*, and *Production and Operations Management* between 2013 and 2017. The results of this analysis compared to previous research show some striking differences suggesting the evolution of OM research away from operations research and engineering. The results also illustrate the diversity of OM research ranging from analytical modeling to empirical studies influenced by other business disciplines, economics, and the behavioral sciences.

KEYWORDS: Operations Management, Journal Ranking, Citation Analysis

INTRODUCTION

The field of operations management (OM) has been evolving for many years and may broadly be construed as incorporating supply chain management, quality management, product and process design, project management, and other topical areas. OM is defined as the design, operation, and improvement of the systems that create and deliver the firm’s primary products and services (Jacobs & Chase, 2018). Given the breadth of the discipline and the overlap with other fields of inquiry, it is often difficult to distinguish between OM research and research related to the fields of general management, industrial engineering, and management information systems. This is particularly true of the overlapping, but distinct field of operations research/management science (OR/MS), which is defined as the application of quantitative methods to decision making in all fields (Jacobs & Chase, 2018). Operations management is moving in the direction of placing greater emphasis on the management aspects of problems and less emphasis on the techniques of OR/MS (Fuller & Martinec, 2005). As a logical consequence, there are many viewpoints about what constitutes high-quality OM research. To help identify quality OM research, several studies have examined journal rankings according to OM researchers.

There are many reasons for evaluating the quality of journals (Coe & Weinstock, 1984; Barman, Hanna, & LaForge, 2001; Meredith, Steward, & Lewis, 2011) beyond promotion and tenure (P&T) decisions and faculty merit evaluation including:

- Nominations for professorships, endowed chairs and fellowships
- Editors seeking feedback on their article selection process
- Research grants and awards
- Library administrators making journal acquisition decisions
- Business school/program rankings.

The next section examines recent journal ranking studies focused on the OM field. This review focuses on identifying the survey respondent or citation analysis base journals (i.e., basis for study), the measure of journal ranking used by each study (e.g., quality, influence, or relevance), and the degree to which authors distinguish the OM and OR/MS fields. The authors examine OM journals using a citation analysis based on articles published in four highly regarded OM journals. Results of the analysis are then discussed.

LITERATURE REVIEW

Opinion surveys, author affiliation indices (AAI), and citation analyses are well established methodologies for evaluating journals within the various business and related fields such as accounting (Chan, Tong, & Zhang, 2012), finance (Oltheten, Theoharakis, & Travlos, 2005; Chen & Huang, 2007), information systems (Peffers & Ya, 2003; Katerattanakul, Han, & Hong, 2003; Barnes, 2005; Ferratt, Gorman, Kanet, & Salisbury, 2007), management (Podsakoff, MacKenzie, Bachrach, & Podsakoff, 2005), and marketing (Baumgartner & Pieters, 2003; Guidry, Hollier, Johnson, Tanner, & Veltsos, 2004; Pan & Chen, 2011). This section examines similar studies focused on the OM field.

Literature Review of Surveys, AAI, and Other Methods

The first published studies of OM journal ranking are Saladin (1985) and Barman, Tersine, and Buckley (1991). Saladin surveyed members of the Operations Management Association to rate various journals based on perceived quality. Barman et al. (1991) surveyed Decision Sciences Institute members who listed OM as their primary interest by asking them to rate journals according to their perceived quality and relevance to OM research. The base list of journals considered within these studies has influenced subsequent studies within the OM field. Three survey studies have since been published where each study focuses on a different respondent population.

Soteriou, Hadjinicola, and Patsia (1999) sampled European members of INFORMS and EurOMA and had respondents rate journals based on “perceived quality and relevance to OM research.” These authors suggested the uniqueness of the operations management by omitting from their survey several industrial engineering and OR/MS journals considered in previous studies.

Barman et al. (2001) provided a “10-year update on the rankings of perceived relevance and quality of selected POM journals” by surveying U.S. members of the Production and Operations Management Society whose primary interest is OM. Results showed the perception of journals had changed since Barman et al. (1991) study They also found several journals were perceived as high quality but were not perceived as being relevant to OM research. For example, *Management Science* and *Operations Research* were the top ranked journals according to perceived quality, but they were ranked much lower regarding relevance to OM. This was an interesting outcome, yet the authors did not separate OM and OR/MS as being different domains. The authors also identified *Manufacturing and Service Operations Management* with its first issue in 1999 as being relevant to OM based on survey write-ins.

Theoharakis, Voss, Hadjinicola, and Soteriou (2007) examined how the diversity or background of self-identified OM researchers affects their “perceived quality and relevance of a journal.” The authors clarified relevant research as that which stems from real industry problems, rather than being relevant to the OM field. Their definition of quality remained consistent with prior studies as being the perceived quality of OM related articles that the journal publishes. The authors found the journal ratings differed significantly based on the region of the world and by whether the researchers identified themselves as an empiricist or as a modeler. This research also made great strides towards distinguishing the OM and OR/MS fields by reporting results for only the 11 academic journals perceived as being most relevant (i.e., based on real industry problems). This decision resulted in the exclusion of journals focused primarily on OR/MS, but they stopped short of explicitly distinguishing the two fields. The authors provided a complement paper that looked at the demographic, geographic, and type of research in more detail (Voss, Soteriou, Hadjinicola, & Theoharakis, 2014).

Harless and Reilly (1998) conceived author affiliation index (AAI) as a way to develop an objective indicator of journal quality and one that is highly correlated with who publishes in these journals which is where researchers from leading U.S. universities publish their scholarly efforts. Gorman and Kanet (2005) were the first to use author affiliation index (AAI) to 27 OM journals and compare their results to survey and citation analysis. Gorman and Kanet (2007) updated their research by applying AAI to 23 journals ranked by their previous research and Olson’s (2005) survey rankings to show that top researchers from top universities publish in top journals.

Other methods have been used to measure journal quality or to examine OM or OM-related research. Xu, Cheang, Lim, and Wen (2014) evaluated 31 OR/MS journals using Google’s PageRank and found that their results matched survey and AAI rankings. Watson and Montabon (2014) developed a ranking of supply chain management journals based on departmental journal ranking lists and discussed some differences in perceived quality based on geographical location. Holsapple and Lee-Post (2010) addressed some of the limitations of surveys, AAI and citation analysis, and developed a behavior-based approach for assessing OM journal importance based on publishing behaviors of 90 tenured OM professors at 31 leading U.S. universities. This led to a set of 27 journals for which they developed three scores to rate these journals in terms of their importance: 1) publishing breadth. 2) publishing intensity, and 3) publishing mode.

Literature Review of Citation Analysis

Vokurka (1996) conducted the first citation analysis study of OM journals and used *Decision Sciences, Journal of Operations Management*, and *Management Science* as the base journals from 1992 to 1994. After culling non-OM articles, Vokurka identified 146 OM articles with a total of 4049 citations. Of the 332 unique journals cited, his analysis identified 25 journals that accounted for nearly 80% of the total journal citations.

Goh, Holsapple, Johnson, and Tanner (1996) described two basic methodologies for conducting citation analysis. The first method inspects data published in the Social Science Citation Index (SSCI), but the authors note this is problematic since some OM-related journals are not indexed in the SSCI. A second approach involves selecting a set of base journals and then manually collecting citation data for all OM-related articles within those base journals. Goh, Holsapple, Johnson, and Tanner (1997) use *International Journal of Production Research, Journal of Operations Management,* and *International Journal of Operations and Production Management* as the base journals for their analysis of citations from 1989 to 1993. The results are normalized to avoid undue weighting towards one of the three base journals. This was accomplished by dividing the number of times a journal is cited in a base journal by the total journal citations in that base journal. The results of this study illustrate the breadth of the OM field by identifying citations of a sizable number of engineering, OR/MS and management journals.

Pilkington and Meredith (2009) used citation analysis combined with a network analysis of co-citation data from *International Journal of Operations and Production Management*, *Journal of Operations Management,* and *Production and Operations Management* to show statistically significant changes in the intellectual structure of the OM field between 1980 and 2006.

Petersen, Aase, & Heiser (2011) used meta-analysis of five journal ranking studies and performed a citation analysis using *International Journal of Operations and Production Management*, *Journal of Operations Management, and Production and Operations Management* from 1999 to 2005. Their results show that OM research has evolved away from operations research and engineering, and that the diversity of OM research ranges from analytical modeling to empirical studies which are influenced by behavioral nature of management and marketing research.

All these studies are subject to inherent limitations associated with the respective methodology used citation analysis. MacRoberts and MacRoberts (1989 & 1996) and Vastag and Montabon (2002) identify and discuss several problems with citation analysis including: uncited primary research, biased citing, self-citing, and other problems associated with the type of publication, nationality, specialty areas, and coverage of literature.

Survey studies also have several potential problems when a list of journals is provided. Beed and Beed (1996) found a typical economist knows no more than eight journals in their relevant field and asking them to rank a long list of journals often reveal personal biases to certain journals. They also identified a study where economists claimed knowledge of journals that did not exist. An example of this is found in a study of electronic commerce journals where a journal was ranked in the top ten even though its first issue was not printed for more than two years after the survey study was completed (Bharati & Tarasewich, 2002). Another issues with survey research is the low response rate. Author affiliation index (AAI) also has drawbacks related to composition of the universities and the numerical nature of the AAI scores (Agrawal, Agrawal, & Rungtusanatham, 2011)

Several citation studies and survey-based journal-ranking studies have been published for the OM field. Although these methodologies have some limitations, they are well established methods for evaluating journals. We also believe the subtle evolution of the OM field based on the discussions in these journal ranking studies is quite interesting. Beginning with the first study of Saladin (1985), authors gradually distinguished OM from the OR/MS and other related fields. However, none of the prior studies explicitly distinguish OM and OR/MS as distinct fields of inquiry.

CITATION ANALYSIS METHODOLOGY

The authors selected four OM journals as the base journals for this citation analysis study: *International Journal of Operations and Production Management (IJOPM), Journal of Operations Management (JOM)*, *Manufacturing and Service Operations Management (MSOM)*, and *Production and Operations Management (POM)*.

These OM journals were selected for several reasons. Articles published in these journals tend to focus solely on the OM field. Craighead and Meredith (2008) stated that these journals are moving toward the OM paradigm, while *Decision Sciences* and *Management Science* do not fully focus on the OM paradigm. Including these two highly regarded journals in the base journal list would require us to determine which articles are OM-related. We prefer to avoid any bias associated with this type of subjective decision.

The four journals selected also represent the broad cross section of research methodologies employed by OM researchers, which will likely instill some heterogeneity into the study. Results of this study support this premise. Over time and with encouragement of many leading academics, the OM field has grown to include both empirical and modeling research. OM journals such as *IJOPM* and *JOM* currently cater more to empirical research, while *MSOM* and *POM* tends to publish more modeling-based OM research that some OM researchers may feel falls into the OR/MS domain. Consequently, these base journals accurately represent a breadth of the best work published in the OM field.

Citations from articles published in the four base journals between 2013 and 2017 were entered into a common database. According to *Academic Journal Guide* (2018), these base journals are all categorized as OM journals, thus we assume the 1,415 articles published in these journals are all OM-related and consequently are included in our analysis. Table 1 summarizes the journal and citation statistics for the four base journals. These articles contained a total of 78,337 citations of which 60,316, or 81.3%, were research journal articles. Results show that citations from all four base journals are comprised of similar percentages between journal citations, books and other categories. These values are consistent with previous studies (Vokurka, 1996; Goh et al, 1997; Petersen et al, 2011).

Table 1: Journal and Citation Statistics for *Journal of Operations Management (JOM)*, *International Journal of Operations and Production Management (IJOPM), Manufacturing and Service Operations Management (MSOM)*, and *Production and Operations Management (POM)* for the years 2013 through 2017

	IJOPM	JOM	MSOM	POM	Total
Number of issues	60	30	20	54	164
Number of articles	343	200	213	659	1,415
Articles per issue	5.7	6.7	10.7	12.2	8.6
Total citations	27,060	15,932	8,391	26,954	78,337
Citations per article	78.9	79.7	39.4	40.9	55.4
Journal citations	21,894	12,422	5,669	20,331	60,316
as a percentage	80.9%	78.0%	67.6%	75.4%	77.0%
Book (etc.) citations	5,089	3,401	2,359	5,873	16,722
as a percentage	18.8%	21.3%	28.1%	21.8%	21.3%
Misc. citations	77	109	363	750	1,299
as a percentage	0.3%	0.7%	4.3%	2.8%	1.7%
Number of unique journals cited	942	601	356	883	1,805

Note: Misc. citations include Working papers, dissertations/theses, and Non-traditional sources.

Based on the procedure used by Goh et al. (1997), we normalized the data for each base journal by dividing the number of citations for each cited journal by the total number of citations for that base journal. These normalized indices essentially represent the percentage of citations for a journal in each base journal. For example, *JOM* was referenced a total of 1,952 times within the issues published by *JOM* between the years of 2013 to 2017. This value is normalized by the total journal citations of 12,422 yielding a normalized index of 15.7%. The four normalized indices for a journal are then averaged across the four base journals. For *JOM* as an example, these indices are 9.6% for *IJOPM*, 15.7% for *JOM*, 0.6% for *MSOM*, and 3.5% for *POM* to yield an average normalized index of 7.3%.

Citations within a journal are often biased toward a few select journals. This phenomenon is common, and perhaps expected, particularly in diverse fields such as OM. Moreover, journals commonly contain a larger percentage of self-citations. As pointed out by Goh et al. (1997), normalizing the data helps compensate for biases associated with differences in the number of journal issues published during a year and the number of articles published annually. It also mitigates the bias associated with the self-citation phenomenon for the base journals.

CITATION ANALYSIS RESULTS

A total of 1,805 cited journals are ranked based on the average normalized index across the based journals, with the top 40 most cited journals listed in Table 2. The articles cited from these journals account for 48.96% of all the journal citations across the four base journals, while they account for 46.42%, 52.68%, 46.60%, and 50.04% of the journal citations within *IJOPM*, *JOM*, *MSOM*, and *POM*, respectively. These un-normalized rankings are based on the total citations a journal received from the base journals. Results of a Sign Test show that there is no significant difference between the normalized and un-normalized ranking at the 10% level. The normalized and un-normalized rankings share the same top 14 most cited journals. *Management Science* is ranked the first in both normalized and un-normalized ranking, followed by *JOM* as the second place in both rankings. One may observe the significant differences in the total citations between two closely ranked journals based on the normalized index. For example, *JOM* is ranked the second and *Operations Research (OR)* the third in the normalized ranking, but *OR* received a much smaller total (2,002 citations) from the base journals which is only 41.73% of total citations of *JOM* (4,798 citations).

Interestingly, the *Wall Street Journal* (*WSJ*) is ranked 36th in the normalized ranking. Although *WSJ* is not an academic journal, its appearance in the top 40 list evidently demonstrates the importance of real-world applications amongst the OM research. Hence, we decide to keep it in the list.

Consistent with Petersen et al. (2011), *JOM* is ranked in the top ten cited journals for most base journals except *MSOM*. Similarly, *POM* is amongst the top ten for all base journals except *IJOPM*. Furthermore, *MSOM* is highly cited in *MSOM* and *POM*, but relatively low in *IJOPM* and *JOM*; whereas *IJOPM* is ranked in the top ten for in *IJOPM* and *JOM* citations, but not in *MSOM* and *POM*. These observations echo the phenomenon that *MSOM/POM* and *IJOPM/JOM* emphasize on different streams of research, respectively.

Tables 1 and 2 together facilitate the detection of existence/extent of self-citation issues. The results show that the percentage of self-citing is 11.2% for *IJOPM*, 15.7% for *JOM*, 9.7% for *MSOM*, and 8.4% for *POM*. Although *JOM* has the highest self-citing rate of 15.7%, the self-citing rates are considered relatively low across all base journals, therefore do not raise a concern in our study.

The list in Table 2 indicates that all four base journals cite research works from a wide variety of domains in addition to OM domain. Following *Academic Journal Guide* (2018), Table 3 categorizes the journals in Table 2 into seven domains: Operations Management, Operations Research/Management Science (OR/MS), Management (MGMT), Marketing (MKT), Economics

(ECON), Information Management (IM), and Sector/Miscellaneous. Ten out of 40 journals fall in the OM domain, the largest group, followed by the OR/MS domain with eight journals. This is attributed to the close relationship and ambiguous boundary between OM and OR/MS field. Compared with Petersen et al. (2011), the number of management journals decreased from ten to eight, whereas the number of marketing journals stayed the same although the list is slightly different. Note that such change of the journals in each domain may be partially due to the changes in the categorization. For example, *IISE* (former *IIE Transaction*) was categorized as an Industrial Engineering/Other Engineering journal by *Academic Journal Quality Guide* (2009) but has been recently re-categorized as an OR/MS journal by *Academic Journal Guide* (2018). Interestingly, the economic and information management journals have emerged with six journals that fall into these two domains and are all considered top journals in their respective fields. This observation is aligned with the trend of the application of economic theories in OM studies and the rising interdisciplinary research between the OM and IM fields.

Sign Tests are performed to compare the normalized rankings of the cited journals between every two base journals (Siegel & Castellan, 1988). The results, presented in Table 4, reveals that the OR/MS journals are consistently more important for the modeling-oriented journals (i.e., *MSOM* and *POM*) than for the empirical-oriented journals (i.e., *IJOPM* and *JOM*), whereas the management journals are more important for *IJOPM* and *JOM* than for *MSOM* and *POM*. Within the empirical-oriented stream, the number of higher-ranked journals cited in *JOM* is significantly more than that in *IJOPM* at the 5% level; the differences are not significant on the domain level. Within the modeling-based stream, journals in OM and management domains are both reported to be more important for *POM* than *MSOM*. One explanation may be that *MSOM* tends to be more OR/MS oriented compared with *POM* since *MSOM* appears to cite relatively more OR/MS journals, although the difference is not significant.

DISCUSSION AND CONCLUSIONS

This study is the first citation analysis study of OM journals in eight years. Interestingly, the results of this study share the same top OM journals (i.e., *Management Science* and *Journal of Operations Management*) with the most recent study by Petersen et al. (2011). In fact, seven of the top ten journals and 28 of the top 40 journals are the same as the prior study. Our results show that there are five OM journals, three OR/MS journals, and two MGMT journals in the top ten.

The results of this study show that OM researchers cross boundaries into other business domains, OR/MS, engineering, and the behavioral sciences, but unlike the results shown by Petersen et al. (2011), OM research has been recently drawing upon the research in Information Management and Economics. This study affirms the increase in the breadth of research conducted in the OM field, particularly in empirical research. Fisher (2007) suggested that increasing the empirical elements of operations research would be highly beneficial to OM research. Meredith (2009) advocates for expanding our OM research paradigm from not only empirical and modelling research to one consisting of interpretivism by working with business professionals to help them solve relevant business problems. Operations management is a diverse field that continues to embrace both empirical and modeling research (Krajewski, 2002; Theoharakis et al, 2007). The results of this citation analysis show that this idea is quite valid.

A limitation of this study is the selection of our base journals for the citation analysis: *International Journal of Operations and Production Management*, *Journal of Operations Management*, *Manufacturing and Service Management,* and *Production and Operations Management*. The authors chose these journals because they are highly regarded OM journals that publish a breadth of both empirical and modeling research. Even though self-citation is sometimes an issue with citation analysis, these four base journals had self-citation rates between 8.4% and 15.7% which are low enough to not raise concern over self-citations skewing the results of this study (Nisonger, 2000).

Table 2: Citation Analysis Ranking of Journals for OM Research

Normalized Objective Rank	Un-normalized Objective Rank	Cited Journal	IJOPM Citations	JOM Citations	MSOM Citations	POM Citations	Total Citations	Normalized Average
1	1	Management Science (MS)	393	727	1,238	3,229	5,587	11.3%
2	2	Journal of Operations Management (JOM)	2,111	1,952	33	702	4,798	7.3%
3	5	Operations Research (OR)	29	49	683	1,241	2,002	4.7%
4	4	Production and Operations Management (POM)	353	418	276	1,704	2,751	4.6%
5	6	Manufacturing and Service Operations Management (MSOM)	56	138	549	906	1,649	3.9%
6	3	International Journal of Operations and Production Management (IJOPM)	2,447	253	0	95	2,795	3.4%
7	7	Strategic Management Journal (SMJ)	621	643	22	328	1,614	2.5%
8	10	European Journal of Operational Research (EJOR)	189	63	154	643	1,049	1.8%
9	8	International Journal of Production Economics (IJPE)	699	155	45	225	1,124	1.6%
10	9	Academy of Management Review (AMR)	497	358	10	197	1,062	1.6%
11	11	Harvard Business Review (HBR)	472	246	56	219	993	1.6%
12	12	Academy of Management Journal (AMJ)	390	384	13	187	974	1.5%
13	14	Journal of Marketing (JMK)	260	309	45	180	794	1.3%
14	13	Organization Science (OS)	222	278	29	275	804	1.3%
15	17	Marketing Science (MKS)	12	60	106	457	635	1.2%
16	18	Journal of Marketing Research (JMR)	166	216	41	193	616	1.0%
17	16	Decision Sciences (DS)	307	204	15	171	697	1.0%
18	15	International Journal of Production Research (IJPR)	510	94	17	134	755	1.0%
19	23	American Economic Review (AER)	23	70	90	252	435	0.9%
20	19	Administrative Science Quarterly (ASQ)	209	206	9	136	560	0.9%

Note: Citations of articles published from 2013 to 2017.

Table 2: Citation Analysis Ranking of Journals for OM Research (Cont'd)

Normalized Objective Rank	Un-normalized Objective Rank	Cited Journal	IJOPM Citations	JOM Citations	MSOM Citations	POM Citations	Total Citations	Normalized Average
21	20	Journal of Supply Chain Management (JSCM)	335	172	0	34	541	0.8%
22	22	MIS Quarterly (MISQ)	154	223	2	100	479	0.8%
23	27	IISE Transactions (IISE) (*Former IIE Transactions*)	25	13	77	252	367	0.7%
24	31	Naval Research Logistics (NRL)	0	8	90	210	308	0.7%
25	21	Supply Chain Management: An International Journal (SCM)	396	63	6	35	500	0.6%
26	32	Econometrica (ECMA)	2	53	63	183	301	0.6%
27	26	Journal of Management (JM)	179	142	2	59	382	0.6%
28	29	Information Systems Research (ISR)	44	154	7	143	348	0.6%
29	25	Journal of Product Innovation Management (JPIM)	232	65	4	84	385	0.5%
30	36	The RAND Journal of Economics (RJE)	4	19	54	187	264	0.5%
31	30	Journal of Applied Psychology (JAP)	105	122	7	93	327	0.5%
32	24	International Journal of Physical Distribution and Logistics Management (IJPDLM)	281	67	0	43	391	0.5%
33	28	Industrial Marketing Management (IMM)	282	63	2	18	365	0.5%
34	42	INFORMS Journal on Applied Analytics (IJAA) (*Former Interfaces*)	43	40	47	97	227	0.5%
35	48	Quarterly Journal of Economics (QJE)	13	26	49	129	217	0.4%
36	50	Wall Street Journal (WSJ)	23	34	46	101	204	0.4%
37	35	Journal of International Business Studies (JIBS)	134	106	0	29	269	0.4%
38	33	Journal of Business Logistics (JBL)	149	92	0	37	278	0.4%
39	40	Journal of the Operational Research Society (JORS)	37	36	23	138	234	0.4%
40	34	Journal of Business Research (JBR)	157	72	0	41	270	0.4%

Note: Citations of articles published from 2013 to 2017

Table 3: Research Domain Classification of Journals Ranked by Citation Analysis

Operations Management (OM) Journals	Operations Research/ Management Science (OR/MS) Journals	Management (MGMT) Journals	Marketing (MKT) Journals	Economics (ECON) Journals	Information Management (IM) Journals	Sector/ Miscellaneous Journals
IJOPM (6)	DS (17)	AMJ (12)	IMM (33)	AER (19)	ISR (28)	JIBS (37)
IJPDLM (32)	EJOR (8)	AMR (10)	JMK (13)	ECMA (26)	MISQ (22)	JAP (31)
IJPE (9)	IISE (23)	ASQ (20)	JMR (16)	QJE (35)		JPIM (29)
IJPR (18)	IJAA (34)	HBR (11)	MKS (15)	RJE (30)		WSJ (36)
JBL (38)	JORS (39)	JBR (40)				
JOM (2)	MS (1)	JM (27)				
JSCM (21)	NRL (24)	OS (14)				
MSOM (5)	OR (3)	SMJ (7)				
POM (4)						
SCM (25)						

Note: Number in parentheses is the normalized ranking from Table 2.
Journals are alphabetically ordered in each domain.

Journal Citation Analyses of Operations Management
Journals

Table 4: Sign Test Results of Direct Comparisons between Base Journal Citation Rankings

Citation Ranking Comparisons Between Base Journals A and B	Journal *A* Rank is higher than *B*	Journal A Rank is lower than B	No Change in Rank	*p-value*
IJOPM (A) vs. JOM (B)*	12	27	1	0.012
OM domain subgroup	6	4	0	0.377
OR/MS domain subgroup	2	6	0	0.145
MGMT domain subgroup	2	5	1	0.227
MKT domain subgroup	1	3	0	0.313
ECON domain subgroup	0	4	0	0.063
IJOPM (A) vs. MSOM (B)	24	16	0	0.134
OM domain subgroup	8	2	0	0.055
OR/MS domain subgroup*	1	7	0	0.035
MGMT domain subgroup*	8	0	0	0.004
MKT domain subgroup	2	2	0	0.688
ECON domain subgroup	0	4	0	0.063
IJOPM (A) vs. POM (B)	21	19	0	0.437
OM domain subgroup	8	2	0	0.055
OR/MS domain subgroup*	1	7	0	0.035
MGMT domain subgroup*	7	1	0	0.035
MKT domain subgroup	2	2	0	0.688
ECON domain subgroup	0	4	0	0.063
JOM (A) vs. MSOM (B)	25	14	1	0.054
OM domain subgroup*	8	1	1	0.020
OR/MS domain subgroup*	1	7	0	0.035
MGMT domain subgroup*	8	0	0	0.004
MKT domain subgroup	3	1	0	0.313
ECON domain subgroup	0	4	0	0.063
JOM (A) vs. POM (B)	24	16	0	0.134
OM domain subgroup	7	3	0	0.172
OR/MS domain subgroup*	1	7	0	0.035
MGMT domain subgroup*	8	0	0	0.004
MKT domain subgroup	3	1	0	0.313
ECON domain subgroup	0	4	0	0.063
MSOM (A) vs. POM (B)	14	25	1	0.054
OM domain subgroup*	1	9	0	0.011
OR/MS domain subgroup	5	2	1	0.227
MGMT domain subgroup	1	7	0	0.035
MKT domain subgroup	2	2	0	0.688
ECON domain subgroup	4	0	0	0.063

Note: p-values are one-tailed probabilities for binomial test with q = p = 0.50
** denotes statistical significance at 5% level*

REFERENCES

Academic Journal Guide (2018), published by Chartered Association of Business Schools (https://charteredabs.org/academic-journal-guide-2018/).

Academic Journal Quality Guide (2009), published by The Association of Business Schools (www.the-ABS.org.uk/).

Agrawal, V.K., Agrawal, V., & Rungtusanatham, M. (2011). Theoretical and interpretation challenges to using the author affiliation index method to rank journals. *Production and Operations Management*, *20*(2), 280-300.

Barman, S., Tersine, R.J., & Buckley, M.R. (1991). An empirical assessment of the perceived relevance and quality of POM-related journals by academicians. *Journal of Operations Management*, *10*(2), 194-212.

Barman, S., Hanna, M.D., & LaForge, R.L. (2001). Perceived relevance and quality of POM journals: A decade later. *Journal of Operations Management*, *19*(3), 367-385.

Barnes, S.J. (2005). Assessing the value of IS journals. *Communications of the ACM*, *48*(1), 110-112.

Baumgartner, H., & Pieters, R. (2003). The structural influence of marketing journals: A citation analysis of the disciplines and its sub-areas over time. *Journal of Marketing*, *67*(2), 123-139.

Beed, C., & Beed, C. (1996). Measuring the quality of academic journals: the case of economics. *Journal of Post-Keynesian Economics*, *18*(3), 369-396.

Bharati, P., & Tarasewich, P. (2002). Global perceptions of journals publishing e-commerce research. *Communications of the ACM*, *45*(5), 21-26.

Chan, K.C., Tong, J.Y., & Zhang, F.F. (2012). Accounting journal rankings, authorship patterns and the author affiliation index. *Australian Accounting Review*, *22*(4), 407-417.

Chen, C.R., & Huang, Y. (2007). Author affiliation index, finance journal ranking, & pattern of authorship. *Journal of Corporate Finance*, *13*(5), 1008-1026.

Coe, R.K., & Weinstock, I. (1984). Evaluating the management journals: A second look. *Academy of Management Journal*, *27*(3), 660-666.

Craighead, C.W. & Meredith, J. (2008). Operations management research: evolution and alternate paths. *International Journal of Operations & Production Management*, *28*(8), 710-726.

Ferratt, T.W., Gorman, M.F., Kanet, J.J, & Salisbury, W.D. (2007). IS journal quality assessment using the author affiliation index. *Communications of the Association for Information Systems*, *19*(23), 710-724.

Fisher, M. (2007). Strengthening the empirical base of operations management. *Manufacturing & Service Operations Management*, *9*(4), 368-382.

Fuller, J.A., & Martinec, C.L. (2005). Operations research and operations management: from selective optimization to system optimization. *Journal of Business & Economics Research*, *3*(7), 11-15.

Goh, C.H., Holsapple, C.W., Johnson, L.E., & Tanner, J.R. (1996). An empirical assessment of influences on POM research. *Omega*, *24*(3), 337-345.

Goh, C.H., Holsapple, C.W., Johnson, L.E., & Tanner, J.R. (1997). Evaluating and classifying POM journals. *Journal of Operations Management*, *15*(2), 123-138.

Gorman, M.F., & Kanet, J.J. (2005). Evaluating operations management-related journals via author affiliation index. *Manufacturing & Service Operations Management*, *7*(1), 3-19.

Gorman, M.F., & Kanet, J.J. (2007). OM forum: Evaluating operations management-related journals via author affiliation index— Do professors at top U.S. business schools do what they say? *Manufacturing & Service Operations Management*, *9*(1), 51-53.

Guidry, J.A., Hollier, B.N., Johnson, L.E., Tanner, J.R., & Veltsos, C. (2004). Surveying the cites: A ranking of marketing journals using citation analysis. *Marketing Education Review*, *14*(1), 45-59.

Harless, D., & Reilly, R. (1998). Revision of the journal list for doctoral designation. Unpublished report, Virginia Commonwealth University, Richmond, VA, http://www.bus.vcu.edu/economics/harless/.

Holsapple, C.W., & Lee-Post, A. (2010). Behavior-based analysis of knowledge dissemination channels in operations management. *Omega*, *38*, 167-178.

Jacobs, F.R., & Chase, R.B., (2018). *Operations and Supply Chain Management, 18*th edition, McGraw-Hill, New York, NY.

Katerattanakul, P., Han, B., & Hong, S. (2003). Objective quality ranking of computing journals. *Communications of the ACM*, *46*(10), 111-114.

Krajewski, L.J. (2002). Reflections on operations management research. *Journal of Operations Management*, *20*(1), 2-5.

MacRoberts, M.H., & MacRoberts, B.R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, *40*(5), 342-349.

MacRoberts, M.H., & MacRoberts, B.R. (1996). Problems of citation analysis. *Scientometrics*, *36*, 435–444.

Meredith, J.R. (2009). Issues in the modeling-empiricism gap. *Journal of Supply Chain Management*, *45*(1), 44-48.

Meredith, J.R., Steward, M.D., & Lewis, B.R. (2011). Knowledge dissemination in operations management: Published perceptions versus academic reality. *Omega*, *39*, 435-446.

Nisonger, T.E. (2000). Use of the journal citation reports for serials management in research libraries: An investigation of the effect of self-citation on journal rankings in library and information science and genetics. *College & Research Libraries*, *61*(3), 63-275.

Olson, J.E. (2005). Top-25-business-school professors rate journals in operations management and related fields. *Interfaces*, *53*(4), 323-338.

Oltheten, E., Theoharakis, V., & Travlos, N.G. (2005). Faculty perceptions and readership patterns of finance journals: A global view. *Journal of Financial and Quantitative Analysis*, *40*(1), 223-239.

Pan, Y., & Chen, C.R. (2011). Author affiliation index: A new approach to marketing journal ranking. *Marketing Education Review*, *21*(3), 275-287.

Peffers, K., & Ya, T. (2003). Identifying and evaluating the universe of outlets for information systems research: Ranking the journals. *Journal of Information Technology Theory and Application*, *5*(1), 63-84.

Petersen, C.G., Heiser, D.R. & Aase, G.R. (2011). Journal Ranking Analyses of Operations Management Research. *International Journal of Operations & Production Management*, *31*(4), 405-422.

Pilkington, A., & Meredith, J. (2009). The evolution of the intellectual structure of operations management — 1980–2006: A citation/co-citation analysis. *Journal of Operations Management*, *27*, 185-202.

Podsakoff, P.M., MacKenzie, S.B., Bachrach, D.G., & Podsakoff, N.P. (2005). The influence of management journals in the 1980s and 1990s. *Strategic Management Journal*, *26*, 472-488.

Saladin, B. (1985). Operations management research: Where should we publish. *Operations Management Review*, *3*(4), 3-9.

Siegel, S., & Castellan, Jr., N.J. (1988). Nonparametric Statistics, 2nd edition, McGraw-Hill, New York, NY.

Soteriou, A.C., Hadjinicola, G.C., & Patsia, K. (1999). Assessing production and operations management related journals: The European perspective. *Journal of Operations Management*, *17*(2), 225-238.

Theoharakis, V.T., Voss, C.A., Hadjinicola, G.C., & Soteriou, A.C. (2007). Insights into factors affecting production and operations management (POM) journal evaluations. *Journal of Operations Management*, *25*(4), 932-955.

Vastag, G., & Montabon, F. (2002). Journal characteristics, rankings, and social acculturation in operations management. *Omega*, *30*, 109-126.

Vokurka, R.J. (1996). The relative importance of journals used in operations management research: A citation analysis. *Journal of Operations Management*, *14*(4), 345-355.

Voss, C.A., Soteriou, A.C., Hadjinicola, G.C., & Theoharakis, V.T. (2014). The diversity of the POM field as reflected in the evaluation of its journals. *International Journal of Management, Business, and Economics*, *1*(1), 23-52.

Watson, K., & Montabon, F. (2014). The ranking of supply chain management journals based on department lists. *International Journal of Production Research*, *52*(14), 4364-4377.

Xu, Z., Cheang, B., Lim, A., & Wen, Q. (2011). Evaluating OR/MS Journals via PageRank. *INFORMS Journal on Applied Analytics*, *41*(4), 375-388.

DECISION SCIENCES INSTITUTE

The Disparity in Performance Criteria Used in College Ranking

Jean-Claude Pedjeu
Tennessee State University
Email: jpedjeu@tnstate.edu

Aurore J. Kamssu
Tennessee State University
Email: akamssu@tnstate.edu

Jeffrey S. Siekpe
Tennessee State University
Email: jsiekpe@tnstate.edu

ABSTRACT

This study analyzes the factors affecting college rankings. The question is whether rankings signal quality. Findings reveal that ranking systems use different parameters; but how do those parameter relate to student learning, faculty research and service to the community. The results should explain why institutions emphasize on some performance measures.

KEYWORDS: University ranking, school rating, Teaching, Student performance, Faculty performance, university service, school quality, Information Technology.

INTRODUCTION

Every day, we are called to make decisions that affect our lives. For instance, we have to choose restaurants when opting to eat out, cars and cars dealers when selecting a mode of transportation, house style and location when buying or renting a home, schools when deciding on education and so forth. However, some choices are more important than others in the sense that they have greater impacts on our lives. The choice of the type of education one get can shape the future health and wealth of an individual. Therefore, during the decision-making process, one usually takes into account inputs from a variety of sources. One critical decision that will affect many people life is their choice of institution of higher education (Mårtensson & Richtnér, 2015). The Economist (2010) reported that higher education seems like a market for an expensive product, where students, in order to spend money on out-of-state tuition or private school, want to make sure the school is ranked better than the public schools in their state. School ranking has become an important tool used by students, faculty, and other stakeholders to compare institutions. Rankings start as early as nursery schools and continue all the way beyond to post-graduate programs.

Studies on rankings of educational institutions have been conducted at different levels and in many disciplines including but not limited to management, economics, law, organizational sociology, strategy, and corporate reputation. Professional programs particularly, business and Law schools are very driven by ranking. It is believed that rankings may (1) help set the standard for what is considered a good school in terms of both 'identity' (Albert & Whetten, 985; Devinney, et al., 2008; Rindova & Fombrun, 1999) and competitive ‘landscape’ (Devinney, et al., 2008; Wedlin, 2006); (2) confer authenticity, legitimacy , and positional status on the schools that are ranked (Devinney et al., 2008; Gioia & Corley, 2002); (3) help expedite changes when key stakeholders show concerns about the school image the school (Elsbach & Kramer, 1996; Martins, 2005; Zemsky, 2008); (4) adjust programs to fit local needs while opening doors to distant opportunities and challenges. There is a need to understand the methodology used to rank schools. Especially because there has been little evidence or attempt to assess rankings in a more systematic manner; given the attention attributed to these rankings in addition to the fact that people have divergent views about what they represent and how they work (Dichev, 2008). The current research examines as Devinney et al. (2008) suggests the sub-component scores of rankings and the performance criteria used by ranking organizations.

LITERATURE REVIEW

A multitude of rankings emerged over the years for different purposes. To understand the dependence/reliance on rankings, it is important to understand the origin, retrace the history of college rankings, then examine what college rankings are used for, finally, analyze the use of the rankings.

The Origin and Evolution of Ranking

In 1906, a distinguished American psychologist, James McKeen Cattell (Cattell, American Men of Science: A Biographical Directory, Second Edition 1910), laid the foundation of today's college and university rankings by first compiling and publishing a list of more than 4,000 men he deemed accomplished scientists based on the degrees earned, achievements, and socio-economic status. That publication was "*American Men of Science: A Biographical Directory*." The list was updated and resulted to another publication "*American Men of Science: A Biographical Directory, Second Edition*" at the end of 1910, which included the biographies of about 5500 prominent scientists, the institutions they attended and where they taught at post-graduation (Cattell, American Men of Science: A Biographical Directory, Second Edition 1910). The institutions were also ranked based on how many scientists were associated with them. This was the first published academic quality ranking of American universities. The five leading institutions were Harvard, Chicago, Columbia, Hopkins, and Yale. The third edition of the "*American Men of Science*" was published in 1921 and contained about 9600 scientists and the analysis further includes occupation/discipline and position (Cattell and Brimhall, American men of science; a biographical directory 1921). Today, the "American Men of Science" is published as "American Men and Women of Science". These early 1900's ranking were Outcome-based rankings as opposed to the reputation based ranking which began much later.

Outcome-based rankings

In 1910, The American Association of Universities (AAU), believing that an impartial ranking would be widely accepted, requested that the United States federal government through the Bureau of Education Study (EBS) rank colleges based on quality. The goal of the AAU was to help institutions select the best undergraduate students for their graduate programs. However, the release of the results of the study was prohibited by an executive order from the then United State President William H. Taft (Ranking Everything College 2018). Stephen S. Visher (1931) produced another example of Outcome-Based rankings by analyzing the institutions and regions most associated with the "Who's Who in America" leaders (Visher 1931). He claimed that New England produced about twice as many "Who's Who in America" as did the middle Atlantic states and about ten times as many as the south central states.
Between 1930 and 1951, Beverly Waugh Kunkel, a biologist at Lafayette College, and Donald B. Prentice, then president of the Rose-Hulman Institute of Technology, used a methodology

like that of Cattell and Visher to assess academic quality. They believed that “undoubtedly the most reliable measure” of a higher education institution was “the quality of product.” Therefore, they ranked universities based on the number of alumni listed in "Who’s Who in America" (Kunkel and Prentice 1951) (Myers and Robe 2009). These outcome-based rankings were not the only ranking methodologies with roots in the early 1920; other methods included hinge on peer-review opinion as well as the ever-popular reputational-based rankings.

Reputational-based Rankings

In 1924, Raymond Hughes compiled the first list ranking a department’s quality based solely on reputation rather than the popular outcome-based methodologies; faculty, from 36 institutions and 20 fields, rated instructors in their respective disciplines (Ranking Everything College 2018). Hayward Keniston (1959) compiled list of fifteen strongest departments in their discipline at twenty-five top universities, then gathered the results into institution-wide rankings (Keniston 1959); which contributed to the rise of reputational rankings.

Reputational rankings have grown in popularity and become the predominant method for generating academic quality rankings; as the reputation of institution has a greater impact on ranking than the eminence of students produced (Ranking Everything College 2018). Reputational-based methodologies considered faculty quality as a criterion in the assessment of program quality and fixed the research gap on the evaluation of graduate programs. The main emphasis has been on institutions with high national visibility (Conrad and Pratt 1985) (Hughes, A Study of the Graduate Schools in America 1925) (Hughes, Report of the Committee on Graduate Instruction. 1934).

The 1934-1936 biennial survey on education in the United States report that in the early 1900s, the educational standards in colleges and universities in the US differs greatly. It took about 30 years for the North Central Association of Colleges and Secondary Schools to set up quantitative standards for accreditation of institution of higher education. Overtime, it was noted that accreditation standards put a lot of weight on basic factors, such as endowments or income, infrastructures, faculty education, magnitude of the library, rather than by the extent to which their programs achieve targeted objectives. Hence, in April 1934, after a 3-year study, a new program was adopted proposing that a higher education institution be judged according to the purposes it seeks to serve or based on the total pattern it presents as an institution of higher education (John 1938). Furthermore, the Educational Policies Commission of the National Education Association (NEA) (1935-1968) met to develop a long-term plan, spark new thinking to education, and continuously appraise the teaching profession (Ebben 1982).

Cartter (1966) assessing the quality of graduate education of 29 fields of studies, evaluated (1) the quality of graduate faculty, (2) the effectiveness of the doctoral program, and (3) the probable changes in relative positions of departments in the following 5-10 years. The study polled 4,000 student scholars and department chairs of 106 institutions, used a very comprehensive methodology and was well received by higher education officials; it therefore, lead reputational rankings to become the norm (Cartter 1966). Thereafter, reputational rankings for graduate programs evolved due to tremendous hard work of Kenneth Roose and Charles Andersen from the American Council on Education; however, upon analyzing the second survey of the quality of graduate programs in the United States Johnson (Johnson 1970) argued that rankings are to some degree imperfect.
In 1981, The National Academy of Sciences with the National Research Council published the assessment of research-doctorate programs in the U.S because of rating 2,699 programs at

228 institutions. Fifteen years later, the National Research Council extended the scope by ranking 41 disciplines at 274 institutions resulting in data on 3,634 programs. The 1981-methodology improved when L. Solmon and A. Astin published the ranking of 80-150 departments in 7-fields based on 6-criteria. The 7-fields: biology, business, chemistry, economics, English, History, and sociology. The 6-criteria: overall quality of education, preparation for graduation, preparation for job, faculty commitment, faculty accomplishments, and innovativeness of curriculum.

In 1983, US News published its first edition of ranking college "Rating the Colleges"; followed in 1986 by the US News second edition of ranking college titled " The Best Colleges in America "; then in 1988, the US News third edition titled "America's Best Colleges"; and from 1989 to today US News college rankings have been published every year. On the same note, in 2003, The Shanghai's Jiao Tong University published the rating of institutions based on their research quality. Likewise, in 2004, a British periodical, Times Higher Education Supplement, along with a provider of guides to higher education, Quacquarelli Symonds, countered by producing new rankings based on outsiders' view including employers that recruit graduates and other academics (The Economist, 2010). In 2008, Forbes published its first annual list of "America's Best Colleges".

Since then, some rankings emerged based on such criteria as, the number of links to the university's website from other institutions, the contents of a database of academic articles. Other rankings were created to sell advertisement space and news media; others such as the Shanghai's Jiao Tong University ranking, were intended to acquire more funding for university research. Quacquarelli Symonds created their ranking to give the affiliated countries international status. The *Financial Time* measure were, also motivated by a desire to strengthen and promote the European identity of management education (Wedlin, 2006). The desire to rank high translates in the assumption that a school is only as good as the rankings present it to be. In addition, most schools are willing to bear any high costs if it can improve their ranking. Therefore, the validity and practical relevance of rankings based on outsiders' opinion has raised concerns in academia (Devinney et al., 2008). In addition, DeNisi (2008) and Zemsky (2008) reported that rankings especially those by media outlet do not always signal 'quality' because noise dominates real news; and noise can account for at least half of observed variations in rankings (*Dichev*, 2008). To evaluate the determinants of school quality, it is important to understand the use of college ranking.

The Use of College Ranking

The various stakeholders of universities have a common goal: to promote and insure a quality education for students. The true value of a school must define the number of applicants on the input side as well as the number of employers eager to hire the graduates on the output side. However, the most controversial thing about ranking higher education institutions is that institutions will prioritize to improve their rankings because rankings have a significant influence on admissions, financial resources, and reputation. Some school may try to improve their rankings by redirecting resources toward rankings 'influencing strategies'; therefore, fostering imitation on the measured criteria and reducing uniqueness and heterogeneity (Liberman & Asaba, 2006). Schools may also engage in strategic manipulation to embellish the information given to ranking agencies. These manipulations are achieved by coaching and preparing students, staffs, faculty members and even alumni on the best way to answer questions or by changing the composition of the faculty. In law schools for instance, adjuncts are hired on contracts that end before the audit occurred (Espeland & Sauder, 2007).

With today's trend toward globalization, universities faced the challenge of the global market demand, including the need to broaden the scope and develop mobility for their graduates, in order to be competitive. This challenge is deepening as employers are willingly hiring graduates from foreign schools and the share of foreign students and professors are sharply increasing in the United States and the European Union. There is more than ever a need of a universal way of ranking schools specially in today's global environment, where not only students shop of school internationally, but also employers recruit in the global market. Fletcher (1970) argues that there is a need of a replicable measure that can transcend time and location.

Grewal et al. (2008) found that, over 8 years (1999–2006), ranking positions by the U.S. News for the top 50 National Universities is quite robust as the probability of a university being ranked within 4 points of its current rank is about 90%. Their findings also indicated that increasing expenses on educational resources measures is influence highly ranked institutions in future ranking, while low ranked institutions do well by improving in academic reputation (Grewal, Dearden and Lilien 2008) (Kim, The Functions and Dysfunctions of College Rankings: An Analysis of Institutional Expenditure 2018).

RESEARCH METHODOLOGY AND ANALYSIS OF RANKING PARAMETERS/CRITERIA

Throughout the course of our research, we identified more than a dozen ranking systems. These systems include, but are not limited to, U. S. News ((Dix 2018);), WalletHub.com ((Dix 2018);), Money magazine ((Dix 2018);), Forbes ((Dix 2018);), Kiplinger ((Dix 2018);), Washington Monthly ((Dix 2018);), The Wall Street Journal/Times Higher Education ((Dix 2018);), The Princeton Review ((Dix 2018);), Niche.com ((Dix 2018);), CollegeFactual.com ((Dix 2018);), WalletHub.com ((Dix 2018);), BestColleges.com ((Dix 2018);), CollegeRaptor.com ((Dix 2018);), CollegeAtlas.org ((Dix 2018);). The current study focusses on four of above rankings: U. S. News, The Wall Street Journal/Times Higher Education, Forbes, and the Washington Monthly. The current study analyzes the relationship between universities performances in teaching, research and service and their national rankings. Ranking agency would use different parameters to rank programs, usually based on the signals from the professional world. The question is how much of those signals be based on student performance, faculty resources and performance in research and the university service to the community.

College rankings have drawn significant attention worldwide and has become is a massive business in North America today. More than 100 different guidebooks and rankings are available, and millions of copies are sold each year (Kuh 2011 and references thereafter) (Kuh 2011). As the public's insatiable appetite for comparative information in the form of rankings grows, the U.S. News & World Report (U.S. News), widely viewed as the flagship of college rankings has the enormous yearly consumption of its "US-Best Colleges" ranking, has solidified its market position by expending its ranking business to hospitals among other entities. After three decades and massive criticism on their methodology, the U.S. News & World Report has made significant adjustment in their ranking parameters and associated weights. As commented above, the question has always been about how possible it is to assess the quality of teaching in institutions of higher education. Altbach (2010) argued that, teaching, one of the main functions of any university is largely ignored in all the rankings. He explained that "the quality and impact of teaching is virtually impossible to measure and quantify". To the credit of Critics of college rankings, early ranking flaws lie heavily on the variability in the ranking factors (P. Altbach 2010).

Ranking System (Organization)	Ranking Criteria	Weights
2019-US News & World Report College Rankings (USNWR) • Began – 1983 • Total – 301 • Source – Uses peer and high school counselors' reputation surveys and self-reported data from ranked institutions	**Graduation and retention rates (22%)**	
	Average graduation rate (80%)	17.6%
	Average first-year student retention rate (20%)	4.4%
	Undergraduate academic reputation (20%)	
	Peer assessment survey (75%)	15%
	High school counselors' ratings (25%)	5%
	Student selectivity for admission (10%)	
	Standardized tests (77.5%)	7.75%
	High school class standing (22.5%)	2.25%
	Faculty resources for academic year (20%)	
	Class size (40%)	8%
	% Faculty compensation (35%)	7%
	% Faculty with terminal degree in their field (15%)	3%
	Full time faculty (5%)	1%
	Student/faculty ratio (5%)	1%
	Financial resources/student (10%)	
	Financial resources/student (100%)	10%
	Average alumni giving rate (10%)	
	Average alumni giving rate (100%)	5%
	Graduation rate performance (8%)	
	Graduation rate performance (100%)	8%
	Social mobility (5%)	
	* Pell Grant graduation rate (50%)	2.5%
	* Pell Grant graduation rate vs Other Students (50%)	2.5%

Ranking System (Organization)	Ranking Criteria	Weights
2019-Niche • Began – 2002 • Total – 880 • Source – N/A	academic	N/A
	admissions	N/A
	financial	N/A
	student life	N/A

Ranking System (Organization)	Ranking Criteria	Weights
2018-Forbes • Began – 2008 • Total – 650 • Source – It incorporates Post-graduate success, Student debt loads, Student experience (retention rates reported by the Department of Education, and data from Niche), Graduation rates constitute, Academic success. It uses a three-year moving average is used to smooth out the scoring	**Graduation rate (12.5%)**	
	four-year rate	7.5%
	six-year rate	2.5%
	Pell Grant Recipients	2.5%
	Academic Success (12.5%)	
	Academic Award past 1-4 yrs.	6.25%
	NSF's earned Doctorates	6.25%
	Student Experience (20%)	
	Freshmen to Sophomore	15%
	Niche college ranking	5%
	American Leaders List (15%)	
	American Leaders List (100%)	15%
	Debts (20%)	
	Debts (100%)	20%
	Alumni Salary (20%)	
	Alumni Salary (100%)	20%

Ranking System (Organization)	Ranking Criteria	Weights
2019-Money Magazine • Began – 1990 • Total – 726 • Source – Takes into account 21 factors partitioned into three groups: educational quality, affordability, and alumni earnings	**Quality of education (1/3 of weighting i.e. 33.33%)**	
	Graduation rate	10%
	Value-added graduation rate	10%
	Peer quality [Academic preparation of students, Yield]	3.33%
	Instructor quality	3.33%
	Financial troubles	5%
	Pell Grant recipient outcomes	1.67%
	Affordability (1/3 of weighting i.e. 33.33%)	
	Net price of a degree	10%
	Educational Debt [Student borrowing, Parent borrowing]	6.67%
	Ability to repay: [Student loan default risk index (SDRI), Value-added SDRI, Federal student loan repayment, Value-added student loan repayment]	10%
	Affordability for low-income students	6.67%
	Outcomes (1/3 of weighting i.e. 33.33%)	
	Graduates' earnings	4.17%
	Earnings adjusted by majors	5%
	College Scorecard 10-year earnings	3.33%
	Value-added earnings	4.17%
	College Scorecard employment outcomes	8.33%
	Job meaning	1.67%
	Socio-economic mobility index	6.67%

Ranking System (Organization)	Ranking Criteria	Weights
World Street Journal (WSJ) and Times Higher Education (THE) • Began – 2004–2009 in collaboration with the Quacarelli Symonds & since 2010 independently • Total – 800 • Source – Uses two reputational surveys for research and teaching (previously used Thomson-Reuters Global Institutional Profiles project data but now plans to do its own data collection and analysis)	**Resources. Does the college have the capacity to effectively deliver teaching? (30%)**	
	amount of money that each institution spends on teaching per student	11%
	ratio of students to faculty members	11%
	number of published scholarly research papers per faculty	8%
	Engagement. Does the college effectively engage with its students? (20%)	
	student's engagement with learning	7%
	student's opportunity to interact with others	6%
	student recommendation	4%
	number of different subjects taught	3%
	Outcomes. Does the college generate good and appropriate outputs? (40%)	
	graduation rates for each institution	11%
	value added by the teaching at a college to both salary	12%
	student's ability to repay debt	7%
	academic reputation of the college	10%
	Environment. Is the college providing a learning environment for all students (diverse student body and faculty)? (10%)	
	proportion of international students on campus	2%
	student diversity – both racial and ethnic diversity	3%
	inclusion of students with lower family earnings	2%
	racial and ethnic diversity of the faculty	3%

Ranking System (Organization)	Ranking Criteria	Weights
Washington Monthly (2018) • Began – 2006 • Total – 316 • Source – Ranks College and universities based upon how well each enhances social mobility, fosters scientific and humanistic research and promotes an ethic of service.	**Research (25%)**	
	institution's research spending	5%
	number of sc. and eng. PhDs awarded by the university	5%
	Number of undergrad. alumni who end up with PhD	5%
	number of faculty receiving prestigious awards	5%
	number of faculty in the National Academies	5%
	Social mobility (50%)	
	*** Graduation rate (16.66%)**	
	**** reported graduation rate (RGR) (8.33%)**	
	Pell grant	5.56%
	first-generation	2.77%
	**** Predicted graduation rate PGr vs RGR (8.33%)**	
	Pell grant	5.56%
	first-generation	2.77%
	*** Financial success (16.66%)**	
	*** Financial success (100%)**	16.66%
	*** Student loan repayment rate (16.66%)**	
	raw repayment rate	8.33%
	regression-adjusted repayment rate	8.33%
	Ethic of service (25%)	
	military service	5%
	national service (alumni in Peace Corps)	5%
	percentage of FW-S in CSP	5%
	voting engagement	5%
	national service	5%

The variability makes tracking the performance of ranking over time rather impractical. As noted by Kuh (2011), the 1966's College-Rater weight distribution was as follows: high school rank (10%), SAT/ACT scores (2%), number of Rhodes, Wilson, Danforth, and NSF award recipients (5%), student-faculty ratio (4%), proportion of students entering graduate school (9%), library collection (15%), proportion of faculty with doctorates (18%), and faculty salaries (18%). Note that this came 30 years after a program was proposed that an institution of higher education be judged based on the total pattern it presents as an institution of higher education. However, since early 2000s, with the emergence of many ranking systems, the weaknesses in college rankings has seen significant consideration. While some ranking systems focus is concerned with affordability (Forbes, Money Magazine), overall campus and student life (Niche, Washington Monthly), academics (U.S. News & World Report), it worth noting that in the latest editions, the major ranking systems methodologies are comprehensive. Since the 2017 edition, U.S. News followed the Carnegie Classification of Institutions of Higher Education's Basic Classification system to determine schools' placement into the ranking categories. They reported that no schools changed categories between their 2018 rankings and their 2019 edition (Morse, Brooks and Mason 2018).

U.S. News & World Report (2010) based their ranking of Business schools on the calculation of the weighted average of eight indicators: (1) full-time average undergrad GPA, (2) full-time average GMAT score, (3) full-time acceptance rate, (4) average starting salary and bonus, (5) graduates employed at graduation, (6) Employed 3 months after graduation, (7) out-of-state tuition and fees, and (8) total full-time enrollment. The *Financial Time (FT)* ranking is created from 20 attributes that focus on the school and its graduates from different programs. The attributes include among other things, average salary after graduation, value for money and the latest FT ranking. Users can chart the rankings variables and view the location of business schools worldwide via an interactive map (Jacobs, 2009). Thus, the unit of analysis used by FT is open to different interpretations (Devinney et al, 2008).

As the impact of the ranking college grows, so is the critic for the ranking methodology and governmental education policies as well as the number of rankings systems which nowadays has exceeded a dozen in the USA and around the world. The variety of the rankings systems brings lots of confusion to stakeholders and serves as a powerful advertisement tool for well ranked universities or competitiveness measure for others (Brooks 2005). Every ranking agency or system has one or a few focus points of interest and even two of them may have common focus points, the parameters used may differ in type or in weights in the rankings. US News and World Report (Morse, Brooks and Mason 2018) put more weight on the academics which includes outcomes [35%: shared among social mobility (5 percent), graduation and retention (22 percent), and graduation rate performance (8 percent)], faculty resources [20%: class size (8 percent), faculty salary (7 percent), full-time faculty with the highest degree in their fields (3 percent), student-faculty ratio (1 percent) and the proportion of faculty who are full time (1 percent)], Expert Opinion [20%: the peer assessment (15 percent) and the high school counselor assessment (5 percent)], and Student Excellence [10%: standardized tests (7.75 percent) and High school class standing (2.25 percent)].

The means a total of 85% of weight in the 2019's US News and World Report US colleges and universities ranking is attributed to academics related items. It's important to note that although the same items in their 2018 ranking were accounted for 85%, the distribution of weight to these items was different. For instance, the student excellence item in the 2018 ranking was worth 12.5%. Hence, only 15% of ranking weight in the past two editions was attributed to nonacademic, namely Financial Resources (10%) and Alumni Giving (5%). Kiplinger, another ranking agency which debuted in 1998, put more emphasis on best value by contrasting academic quality with cost and financial aid measures. It splits the weights in terms of quality criteria (55%) and cost criteria (45%); the detailed weight distribution is not public (K. Pitsker 2018). This methodology appears to encourage low cost and high retention rate in higher education.

Ranking should reflect success in preparing students for their future job or for further education. Leiter, America's most prominent critic of law school ranking methodology used by the *U.S. News & World Report*, claims that the reputation and the quality of education provided by law schools can be seen by the picks of new graduates by elite employers (*Keller*, 2008). Being ranked by a high-profile organization such as *Business Week, U.S. News & World Report*, or *Financial Time* can signal membership of a somewhat exclusive club, and students and other stakeholders evaluate quality thru rankings (Devinney et al, 2008). Those rankings play a major role in schools decision making as well as its competitive power; therefore, being omitted from the rankings may be as worse as having a poor ranking. Moreover, the ranks of top schools are considerably more stable, while there is a lot of volatility for lower-ranked schools; (Devinney et al., 2008; Dichev, 2008; Morgeson & Nahrgang, 2008). The reason is that at the top the disparity is more substantial and therefore, noticeable; making it difficult to lower level school to break to the top. At the lower level on the other hand, there is a wider distribution of similar schools which can perform better from year to year.

SUMMARY AND CONCLUSION

In a nutshell, rankings matter and have an impact on the universities their surroundings; but some schools not fully disclose information about their school, other may engage in strategic manipulation by not only coaching and preparing students, staffs, faculty members and even alumni how to answer question in order to achieve a higher rank, but also by changing the composition of its faculty. This may be misleading especially to prospective students and recruiters who often rely on those rankings during their selection process. More and more, schools are taking a second look before participating in the ranking contests especially with the refusal to provide information to ranking agencies by Harvard and Wharton few years ago. So what information did these agencies use since both schools always find themselves on the list and among the best? Therefore, do rankings measure quality correctly?

We can agree on the notion that schools that ranked better attract quality students and therefore graduate striving students. Programs that attract very good students have better chances of producing high-quality graduates regardless of the quality of the program (Tracy & Waldfogel, 1997). Therefore, ranking may play more of a matching role for students and schools rather than be a determinant of quality. The question is that could those same schools graduate the caliber of students or if they started with students at a lower competence level. The quality of a program should be measured by the level of knowledge it graduates acquire at the school. Therefore, grouping school based on criteria such as the quality of the entering freshman or the allocation of funds may result in better ranking systems and better valuation of programs. This suggestion bring to an end the current discussion, but it does not put a stop on the research, it rather opens doors to more studies that hopefully be explored soon.

REFERENCES

Albert, S. and Whetten, D. A. (1985), "Organizational Identity". In L.L. Cunningham and B.M. Staw (eds.) *Research in organizational behavior*. Vol. 7, Greenwich, CT: JAI Press, pp: 263-295.

Altbach, Philip G. 2010. "University Ranking Season Is Here." *Economic and Political Weekly* 45 (49): 14-17.

Brooks, Rachelle. 2005. "Measuring university quality." *Review of Higher Education* 29 (1): 1-21.

Cartter, Allan M. 1966. Assessment of Quality in Graduate Education. Education, Washington, DC.: *American Council on Education*.

Cattell, James McKeen, and Dean R. Brimhall. 1921. American men of science; a biographical directory. Garrison, N.Y.: *The Science Press*.

Cattell, James McKeen. 1910. American Men of Science: A Biographical Directory, Second Edition. New York: *The Science Press*.

Conrad, Clifton F., and Anne M Pratt. 1985. "Designing for Quality." *The Journal of Higher Education* (Taylor & Francis, Ltd.) 56 (6): 601-622.

DeNisi, Angelo S., 2008, "Rain, snow and sleet are just different types of precipitation". *Academy of Management Perspectives*, 22 (1): 15-17.

Devinney, Timothy, Grahame R. Dowling and Nidthida Perm-Ajchariyawong, (2008). "The *Financial Times* business school ranking: What quality is this signal of quality?" *European Management Review*, Vol. 5, No.4 pg. 195-208.

Dichev, Ilia D. (2008) “The Financial Times business schools ranking: What quality is this signal of quality?” *European Management Review*. Vol. 5, No. 4; pg. 219-224.

Dix, Willard. 2018. How To Rank College Ranking Sites. August 29. https://www.forbes.com/sites/willarddix/2018/08/29/how-to-rank-college-ranking-sites/#17198dd91e82.

Ebben, Linda J. 1982. "American Council on Education Archives, Guide, 1918-1977." Accessed 3 13, 2019. https://files.eric.ed.gov/fulltext/ED219989.pdf.

Elsbach, Kimberly D. and Roderick M. Kramer, 1996, "Members' responses to organizational identity threats: Encountering and countering the *Business Week* rankings". *Administrative Science Quarterly*, 41 (3): 442-476.

Espeland, Wendy N. and Michael Sauder, 2007, "Rankings and reactivity: How public measures recreate social worlds". *American Journal of Sociology*, 113 (1): 1-40.

Fletcher, Colin, 1970, "On replication: Notes on the notions of a replicability quotient and a generalizability quotient". *Sociology*, 4 (1): 51-69.

Gioia, Dennis and Kevin G. Corley, 2002, "Being good versus looking good: Business school rankings and the circean transformation from substance to image". *Academy of Management Learning and Education*, 1 (1): 107-120.

Grewal, Rajdeep, James A. Dearden, and Gary L. Lilien. 2008. "The university rankings game: Modeling the competition among universities for ranking." *The American Statistician* 62 (3): 232-237.

Hughes, Raymond M. 1925. A Study of the Graduate Schools in America. Education, Oxford: Miami University Press.

Hughes, Raymond M. 1934. "Report of the Committee on Graduate Instruction." Educational Record 15: 192-234.

Jacobs Michael. (2009) “Business school rankings online”. *Financial Time* at FT.com

John, Walton C. 1938. Biennal Survey of Higher Education, 1934-1936. Government, Washington, DC: Department of Interior Bulletin, 1937, No. 2.

Johnson, D. Gale. 1970. A Rating of Graduate Programs. Education, Washington, DC: American Council on Fluration,

Keller, Tony. (2008) “Ranking Canada's law schools” *Maclean's* Vol. 121, Iss. 37; pg. 57-58

Keniston, Hayward. 1959. Graduate Study and Research in the Arts and Sciences at the University of Pennsylvania. Education, Philadelphia: University of Pennsylvania Press.

Kim, Jeongeun. 2018. "The Functions and Dysfunctions of College Rankings: An Analysis of Institutional Expenditure." *Research in Higher Education* 59: 54–87.

Kiplinger's Personal Finance. (2009) “How We Rank the Schools*” Kiplinger's Personal Finance,* Vol. 63, No. 12, pg. 52

Kuh, George D. 2011. "Rehabbing the Rankings: Fool's Errand or the Lord's Work?" College and University 86 (4): 8-19.

Kunkel, Beverly W., and Donald B. Prentice. 1951. "The Colleges in Who’s Who in America." October 20: 241-247.

Liberman, Marvin B. and Shigeru Asaba, 2006, "Why do firms imitate each other? *Academy of Management Review*, 31 (2): 366-385.

Mårtensson, Pär , and Anders Richtnér. 2015. "What parameters do students value in business school rankings?" *Journal of Higher Education Policy and Management* 37 (6): 646-658.

Martins, L. L., 2005, "A model of the effects of reputational rankings on organizational change". *Organization Science*, 16 (6): 701-720.

Morgeson, F. P. and Nahrgang, J. D. (2008), "Same as it ever was: Recognizing stability in the *Business Week* ranking". *Academy of Management Learning and Education*, 1 (1): 26-41.

Morse, Robert, Eric Brooks, and Matt Mason. 2018. "How U.S. News Calculated the 2019 Best Colleges Rankings." How the college rankings were calculated. September 9. Accessed March 30, 2019. https://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings.

Myers, Luke, and Jonathan Robe. 2009. College Rankings: History, Criticism and Reform. Washington, DC : Center for College Affordability and Productivity.

Pitsker, Kailin. 2018. Kiplinger's Personal Finance. February. https://www.kiplinger.com/article/college/T014-C000-S002-kiplinger-s-best-college-values-2018.html.

Ranking Everything College. 2018. The History of College Rankings. Accessed March 8, 2019. https://www.collegerank.net/history-of-college-rankings/.

Rindova, V. P. and Fombrun,C. J. (1999), "Constructing competitive advantage: The role of firm-constituent interactions". *Strategic Management Journal*, 20 (8): 691-710.

The Economist. (2010) “International: Leagues apart; University rankings” *The Economist,* Vol. 394, No. 8675, pg. 65

Tracy, Joseph and Joel Waldfogel, 1997, "The best business schools: A market-based approach". *Journal of Business*, 70 (1): 1-31.

U.S. News & World Report. (2010) “Schools of Business” *U.S. News & World Report,* Vol. 147, No. 5, pg. 66

Visher, Stephen Sargent. 1931. "The Comparative Rank of the American States." *American Journal of Sociology* 36 (5): 735–757. www.jstor.org/stable/2766938.

Wedlin, Linda, 2006, "The role of rankings in codifying a business school template: Classifications, diffusion and mediated isomorphism in organizational fields". *European Management Review*, 4 (1): 24-39.

Zemsky, R. (2008) "The rain man cometh - again". *Academy of Management Perspectives*, 22 (1): 5-14.

DECISION SCIENCES INSTITUTE

Educational Data Mining and Student Retention in Higher Education

Adrian Isache
Northern Kentucky University
Email: isachea1@nku.edu

Teuta Cata
Northern Kentucky University
Email: catat@nku.edu

ABSTRACT

The higher education system is dealing with a growing issue: decreasing student retention rates. Student retention could be improved by using a Student Relationship Management model based on data universities are collecting. The focus of such model should be on enriching the "Experience" that students have during their college years.

KEYWORDS: Educational Data Mining, Student Retention, Student Relationship Management

INTRODUCTION

The research field of educational data mining has at its foundation the concept of data mining. In simple theoretical terms, data mining represents one of the steps in the Knowledge Discovery in Databases (KDD) process concerned with the extraction of useful information from raw data. Data mining has been commonly defined as the process of identifying patterns and correlations within large datasets, using intelligent methods such as machine learning, database systems analysis and statistical analysis. As opposed to data analysis, which is concerned with the interpretation of data for descriptive purposes, data mining is focused on knowledge discovery and modelling for obtaining predictive insight. As large batches of data are collected, often times they are rendered useless if not for the discovery and interpretation performed through data mining techniques. By analyzing patterns and relationships in databases, big data is transformed into comprehensible information and predictive outcomes to facilitate the process of decision-making.

In practical terms, the knowledge discovery in databases process (KDD) can be described as interactive as well as iterative, as it consists of several steps that require constant involvement (sometimes repetitive) and active decision-making by the user. Brachman and Anand (1996) and Fayyad et al (1996) were the first to discover and examine the steps in the KDD process. The methods particular to the data mining process, as part of KDD, are used to translate the data into relevant information that is easy to understand. These generally agreed upon methods for performing data mining have been described in research by the same two groups of authors – (Brachman, Ronald J & Anand, Tej., 1996) and (Fayyad, Usama, Piatetsky-Shapiro, Gregory & Smyth, Padhraic., 1996).

Data mining is limitless in terms of its applications and the sectors in which it can be integrated. Figure 1, below, indicates some of the industries where data mining is being applied.

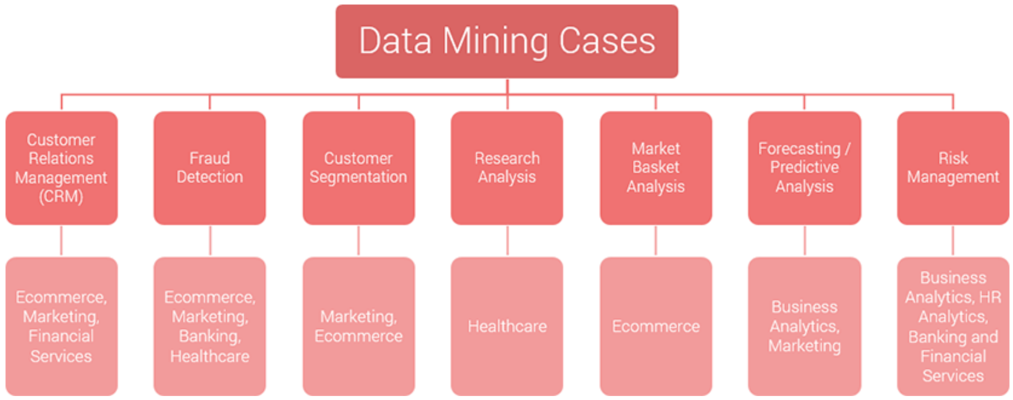


Figure 1. (The App Solutions, 2018)

Some sectors integrate multiple industries listed in the figure above. For example, data mining applications used in customer segmentation, research analysis, forecasting/predictive analysis, business analytics, marketing, etc. are commonly used in the education sector for various purposes.

LITERATURE REVIEW

The Field of Educational Data Mining

The integration of data mining applications, machine learning and statistical analysis, for use in educational settings (e.g., universities), translates into a specific research field: Educational Data Mining (EDM) (Education Encyclopedia).

In the context of educational settings, data mining refers to research, techniques and tools used to develop and improve methods for extracting meaningful information from large repositories of data related to learning activities performed in educational settings. The advances in educational technology, including the explosive increase in computing power and the shift towards computer-based learning environments, have created additional opportunities for analyzing data. Large amounts of data collected from computer-based learning activities are available for analysis and interpretation. Considering the new opportunities, educational data mining has raised continuous interest and in 2009, researchers have established the Journal of Educational Data Mining, as a platform for sharing and disseminating research findings. In 2011, the International Educational Data Mining Society was established, seeking to further grow the field by connecting EDM researchers worldwide (Education Encyclopedia).
These applications require the techniques used in data mining and explained in the first section. The sequence of steps for performing educational data mining is described in the figure below:

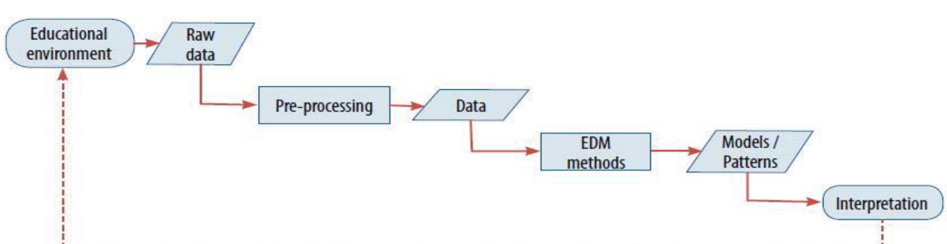


Figure 2. (Universities and Knowledge Society Journal, 2015)

The costs and challenges associated with the implementation of educational data mining are significant. Just storing the data can be costly, but also hiring staff qualified to manage the data systems and ensure the proper and seamless integration is likely to be difficult and add to costs. Also, selecting the data that can be used and useful for mining is a challenging process and requires intensive labor which again, translates into higher costs. Privacy and ethics can also represent challenges in educational data mining, requiring close supervision and attention to these issues from initial strategy to implementation.

In recent years, the impact of educational data mining and its applications has grown exponentially, and the education sector has been completely transformed. As students, teachers, schools and government officials have recognized the effectiveness of EDM and the value of the data collected through its applications, EDM technologies have become more and more prevalent in educational settings. Currently, most higher education institutions rely on EDM to optimize individual learning using educational software, computer-supported assisted learning and computer-adaptive testing resources, pedagogical environments, student success and retention in courses and programs, etc.

Earlier research studies have identified goals and opportunities for improvement of EDM applications in higher education, which are focused mainly on improving different aspects of learning and educational experiences. Such goals focus on the students' learning and behavior based on what students and educators bring to the learning process. These EDM applications consider a broad spectrum of variables related to students learning styles and knowledge along with pedagogical approaches and tactics to improve students' learning (Baker & Yacef, 2009). Improvements have been made in the EDM applications to expand on other students' success analysis than just learning experiences. A research published in 2013 by Romera and Ventura advises more advanced applications to be utilized by EDM. Such applications can be used to detect the overall future academic performance, concerns for academic failure (such as low motivation and poor academic performance), students' social structure match to the academic environment (serves as a predictor of social satisfaction), and curriculum mapping applications (helps matching the curriculum to the learning process). Similar studies suggest that a handful of undergraduate courses should be monitored as predictors for either poor or good performance for undergraduate students, and use the performance on those courses as an early alert to help lower achieving students (Asif at el., 2017).

Student Retention as a Success Index in Higher Education

Student retention in the higher education sector represents the length of time for which a student remains engaged in studies at a particular institution. A series of terms are used to describe specific situations in terms of student retention, such as "retained", "dropout", "transfer", "stopout", and "slowdown" student where each of them represent a group of students who need special consideration and attention in a university setting (Levitz, 2008).

Why do students decide to leave universities before a degree completion? This is a question of interest for higher education institutions and the administrations, but also for students, parents of students, prospective employers, and researchers in the field. From the students' perspective, not completing a degree will translate into less desirable employment opportunities and diminished lifetime earnings, in comparison to college graduates. Parents of students, who in most cases support a substantial amount of costs associated with attending college are put in the position to recognize the investment as a sunk cost if the students do not graduate. Employers are also concerned with student retention as the number of jobs requiring the completion of a higher education degree has been continually increasing, and job requirements are more advanced. Researchers in the field are interested in the concept of student retention, its impacting factors and potential strategies for improvement. Finally, other stakeholders may find value in understanding student retention (Cotter, 2013).

Universities find it increasingly necessary to prioritize efforts to improve student retention and seek effective initiatives and strategies to do so. Every higher education institution has specific conditions and factors determining its retention rate. However, the foundation for developing a strategic plan to be implemented is having a general understanding of what makes students stay enrolled in universities versus dropping out. Every year, universities analyze collected data to generate statistics regarding retention rates. Once they have established the level of retention, administrators look into the factors impacting the results. Then, they transition towards the decision-making process for improving the existing state. This progression requires an initial understanding of the issue (Fishman, Ludgate & Tutak, 2017).

For a university to successfully sustain its mission, student retention represents a fundamental goal. And besides the financial aspect, student retention is a key indicator of organizational success among higher education institutions. According to previous research data, here are some of the main aspects determining the importance of retention rate for institutions of higher education.

To gain insight into why student retention rates are declining is equivalent to understanding why students choose to leave colleges. According to researchers, only 20-30 percent of the students choosing to drop out, do so because of difficulties related to academics. Other factors which influence students to drop out of school are related to the increasing cost of education, finding the right social group to relate to, feeling isolated because of low engagement to different activities, and having unclear goals and expectations about their education (Education Encyclopedia). Among other various factors, these are just the most common challenges for student retention causing decreasing rates. By understanding these factors, universities can develop innovative and efficient programs to help boost retention rates. Statistical research studies show that among higher education institutions in the United States, the highest student retention rates are above 95 percent and the lowest rates may be as low as 10 percent. In the case of elite universities, typical graduation rates may be 85 percent or higher, while for the average level institutions, the rates drop to about 50 percent. The lowest graduation rates reported range between 15 to 25 percent. The students who are most likely to drop out of school are the first-year freshmen. Students in their senior year are the least likely to drop out. Looking at data for an average institution, the attrition rates by students' status are as following: freshman to sophomore year - about 25 percent, sophomore to junior year - about 12 percent, junior to senior year - about 8 percent and only about 4 percent of seniors might leave before their degree completion. For an average institution, about 50 percent of an incoming class of students graduate in 4-5 years (Education Encyclopedia).

While student retention rates may vary because of exceptionally high or low rates for individual institutions, retention is higher at the upper levels of education (masters, doctoral degrees); older institutions with well-established traditions and full-time students who reside on campus (Education Encyclopedia).

Using Educational Data Mining to Increase Student Retention

Data mining in EDM is used extensively to understand what factors are affecting learning and education. A complete review of over 500 research papers published in the last 20 years classified the factors which influence learning in the four main categories: traditional classroom methodologies, technology and media used in learning, educators' actions, and management actions (Rodrigues et al, 2018). A very similar research conducted by Aldowah et al. (2019) on research published between the years 2000 – 2017 identified over 400 studies on the topic of students' learning processes and outcomes. By using different data mining techniques such as clustering, visual data mining, statistics, association rule mining, regression, etc., Aldowah et al (2019) identified the different themes of research in higher education. The majority of the research papers reviewed in this metastudy is focused on students' learning approach, learning materials used by students, collaborative learning used in classes, self-learning, modeling of learning behavior, and retention factors.

Research shows that only 20-30 percent of dropout rates are influenced by difficulties with academics (Education Encyclopedia). The focus of previous research has been mainly on learning and classroom experience. Surprisingly, most of the studies in the last 20 years (over 900 research papers reviewed by Rogrigues et al, 2018 and Aldowa et al. 2019) focus mainly on the classroom teaching and learning styles, technology and learning behavior, but hardly address how students experience their college learning and how they integrate and emerge in the university environment. Not as much research is done on the other reasons why 70-80 percent of dropout rates which are related to: the educational cost, social difficulties, level of engagement and unclear educational expectations (Education Encyclopedia). The latest research is focused more on the “Experience” and “Bonding” that students need with the universities in order to successfully finish their studies.

Chopra and Deranek (2018) suggest that engagement with the local industry partners increases the students' learning and retention during their undergraduate studies. Experiential and active learning teaching style allows students to learn by gaining hands-on experience which is much more valuable for students. Such learning experiences do not only help students retain information and apply critical thinking, but they also help them make the transition to the real world much easier. Students engaged in such active learning activities are more satisfied by the entire learning process. Redirecting the focus of the learning process to the more active learning style, may affect students to continue their education and finish their studies rather than drop out of college.

The keyword for students' success is “Experience”. Students need to experience their education in a way that fits their intentions and backgrounds. A Student Relationship Management (SRM) system can push notifications and allow students to get involved in their education in a way that suits them. A Spanish speaker may want to use the school's portal in Spanish, engineering students can get information about technology labs, a graduate student get different information compared to a freshman. Having an SRM system in place which integrates all information rather than keep them as separate siloes, is very important to keep potential students interested in their education and finish their studies (Britt, 2018).
Social media integration to a Higher Education customer relationship management (CRM) helps with student retention and collaboration. Facebook is found as the most successful social media platform to engage undergraduate students with their academic communications and help with student retention and loyalty. Undergraduate students widely use Facebook versus LinkedIn or Twitter, and it can be used to improve students' experiences while attending college. It can also collect academic concerns shared by students, identify students' negative experiences, and post-academic services feedback (Wali and Andy-Wali, 2018).

Academic advisors have a significant impact of the students' satisfaction in their undergraduate studies. A strong student-advisor relationship can increase the students' perception of fit with the university and their loyalty to their institution. Academic advisors should build a strong interpersonal relationship and trust with the students, which will not only lead to an increase of student retention but also to students staying loyal to the university as alumni (Vianden and Barlow, 2015).

A very interesting program proposed by U-Planner, using data from researchers Ackerman and Schibrowsky, has introduced the concept of a “student relationship management system”, or SRM. Similar to a Customer Relationship Management system, the SRM would be implemented to help institutions build better relationships with their students, increase loyalty to the school and its programs and eventually increase retention rates. Ackerman and Schibrowsky support the SRM idea claiming that “students who perceive a mutual and strong commitment between themselves and the college are more likely to remain enrolled and are more likely to recommend the school to friends” (Ackerman & Schibrowsky, 2008) (U-Planner, 2017).

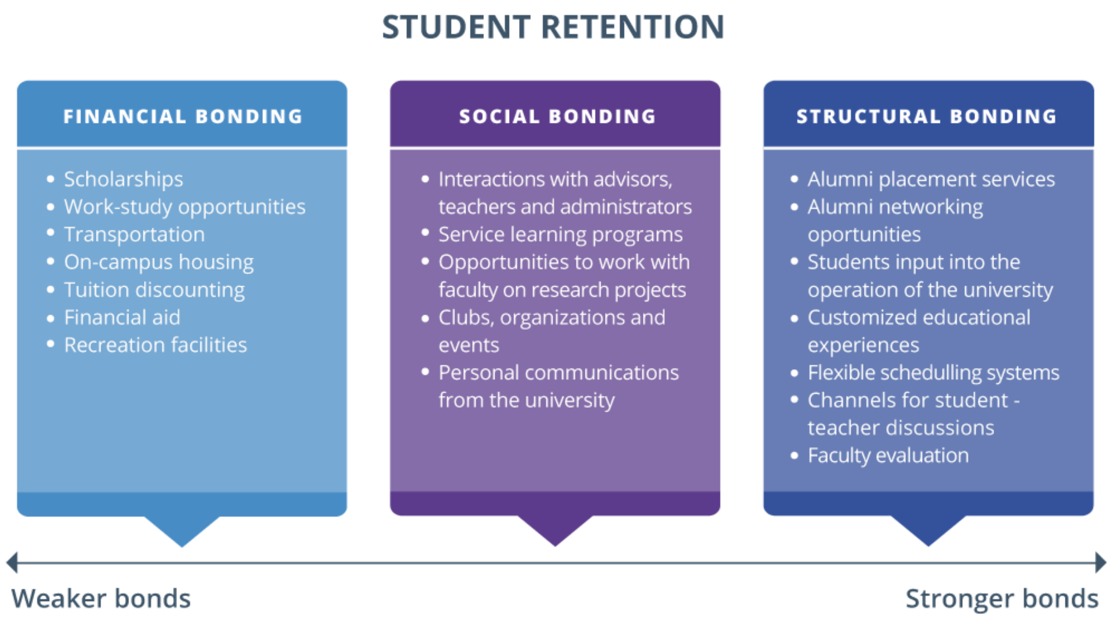


Figure 3. (U-Planner, 2017)

THEORETICAL DEVELOPMENT/MODEL

Student Relationship Management

According to B.J Shaver (2012) and Fontaine (2014), the following table illustrates a comparison between terms used in service marketing literature and their counterpart terms used in enrollment management literature.

Table 1: Glossary of Comparable Terms Used in the Service Marketing Literature and Enrollment Management Literature Service Marketing Terminology Enrollment Management Terminology

Term	Definition	Term	Definition
Customer experience	The period of time from when a customer first receives services from a provider to when he/she stops paying for services	Student experiences	The period of time from when a student matriculates to a university until he/she graduates from or transfers out of the university
Customer expectations	A desire the customer had for his/her service experience prior to the service transaction	Student expectation	A desire the student had for his/her university experience prior to matriculation
Customer changed expectation	A desire the customer had for his/her service experience prior to the service transaction that changed during his/her experience	Student changed expectation	A desire the student had for his/her university experience prior to matriculation that changed during his/her experience
Customer met expectatio	The fulfillment of a desire the customer had for his/her service experience prior to the service transaction	Student met expectations	The fulfillment of a desire the student had for his/her university experience prior to matriculation
Customer satisfaction	The degree to which customers are happy with aspects of their service experience	Student satisfaction	The degree to which students are happy with aspects of their university experience
Customer retention	The customer's act of continuing to patronize a service provider	Student retention	The student's act of remaining enrolled at a university
Customer retention behavior	The customer's act of continuing to patronize a particular service provider or switching to another service provider	Student retention Behavior	The student's act of remaining enrolled or transferring out of a university
Customer life cycle	The period of time beginning when a customer first has contact with a service provider, continuing through his/her customer experience, and lasting as long as he/she remains in contact with the service provider	Student life cycle	The period of time beginning when a student first has contact with a university, continuing through his/her student experience, and lasting as long as he/she remains in contact with the university as an alumnus/alumna
Customer loyalty	A customer's acts of allegiance to a service provider, such as recommending its services to others	Student loyalty	A student's acts of allegiance to one's university, such as applying to graduate school at the university or donating time and/or money to the university
Customer loyalty behavio	A customer's act of encouraging or discouraging others to use his/her provider's services	Student loyalty behavior	A student's act of encouraging or discouraging others to apply to his/her university. Also, a student's act of choosing to apply or not apply to graduate school at his/her university, or choosing to donate or not donate money to the university

Figure 4. (IGI Global)

According to IGI Global dictionary, Student Relationship Management (SRM) is "a specialized CRM concept applied to academic institutions, which involves automating and synchronizing a number of different processes such as academic advising and counseling, in aim to improve the student experience, reduce dropout rates, and improve organizational efficiency" (IGI Global).

In an SRM system, the student takes the role of the customer, as portrayed in how the Customer Relationship Management concept is successfully implemented in the business world. Each student has individual needs in terms of their academic progress and success, and social experience while in a higher education setting. These needs can be better addressed by faculty and administration, through the use of data that is being collected, analyzed and interpreted. An SRM system would represent the central database in which this data is stored and can be accessed to improve decision-making to positively affect students and universities. The concept of the SRM would include everything from the collection of data to the insight provided and the decisions made based on the new information. Higher retention rates could be achieved by touching every aspect of a complete SRM which includes but is not limited to: quality of guidance for selection of major, ability to enroll in desired courses, knowledgeable and approachable advisor, knowledgeable and approachable faculty, safe and secure campus; and much more (Elliot & Shin, 2010).

DISCUSSION AND CONCLUSIONS

The rising interest in the field of Educational Data Mining in the past decade, as well as the constant need for higher retention rates of students in universities, have prompted the discovery of novel opportunities to use big data for the benefit of all stakeholders in higher education. Colleges and universities have been developing and implementing plans to boost student retention, but very few of these initiatives have focused on analyzing the vast amounts of student related data that is already being collected by institutions. Data mining methods can provide an understanding of the causes behind decreasing retention rates. Such an understanding is the basis for building predicting models to prevent attrition and improve students' academic success. All higher education institutions have the opportunity to use already existing resources to develop their own successful retention strategies for the future. Just focus on the data.

The majority of institutions in the United States higher education system are already implementing plans for increasing their retention rate. But it is very often that these plans are not based on the data that is being collected from the students, instructors, alumni, etc. Thus, this valuable data is most likely to remain unused. The plans that are being developed are most times just sporadic initiatives to tackle the issue of retention, but with a vague understanding of the concept and without a project for using data and establishing metrics for tracking and analyzing results. But most universities admit that they are not aware of all the existing data, extra data that could be collected, the proper way to do it, and the way it could be utilized to achieve organizational goals.

Data should be the central pillar in any retention plan implemented at the higher education level. The retention should be focused on planning, implementing and monitoring an SRM where students are advised and mentored on building relationships and creating experiences with the university they are attending. As U-Planner suggests, Structural and Social Bonding help universities to increase student retention. Students need to participate and provide input in the events and structure of their universities. Such participation will increase the bonding and the "fit" with the school they attend. They need to have closer relationships with their advisors, connect with alumni, and participate in events and community activities, which will help them "live" and "experience" the college life in a more beneficial way than just taking classes or partying.

References:

Ackerman, Robert & Schibrowsky, John. (2008). *A Business Marketing Strategy Applied to Student Retention: A Higher Education Initiative.* Retrieved from researchgate.net.

Aldowah, H., Al-Samarraie, H., Fauzy W.M. (2019). *Educational data mining and learning analytics for 21st century higher education: A review and synthesis.* Telematics and Informatics. 37:13-49. Retrieved from science.direct.com.

Asif, R., Merceron, A., Ali, Abbas S. & Haider, Ghani N. (2017). *Analyzing undergraduate students' performance using educational data mining, Computers & Education.* 113, 177-194. Retrieved from sciencedirect.com.

Baker, Ryan S.J.D. & Yacef, Kalina. (2009). *The State of Educational Data Mining in 2009: A Review and Future Visions.* Journal of Educational Data Mining. Retrieved from jedm.educationaldatamining.org.

Brachman, Ronald J & Anand, Tej. (1996). *Advances in knowledge discovery and data mining*. American Association for Artificial Intelligence. Retrieved from dl.acm.org.

Britt Phillip. (2018). *Colleges Can't Cling to Old CRM Technology: Higher education today needs CRM systems that can integrate and adapt.* CRM Magazine. (8):28. Retrieved from destinationcrm.com.

Chopra S, Deranek Kimberly M. (2018). *Efficiently Teaching Engineering and Technology Students through Effective College-Industry Partnerships.* Journal of Engineering Technology. 35(1):10-15. Retrieved from lib.dr.iastate.edu.

Cotter, Brittany. (2013). *Student Retention*. Cobek Software. Retrieved from cobek.com.

Education Encyclopedia. *College Student Retention - Defining Student Retention, A Profile of Successful Institutions and Students, Theories of Student Departure*. Retrieved from StateUniversity.com.

Elliot, Kevin M & Shin, Dooyoung. (2010). *Student Satisfaction: An alternative approach to assessing this important concept*. Journal of Higher Education Policy and Management. Retrieved from researchgate.net.

Fayyad, Usama, Piatetsky-Shapiro, Gregory & Smyth, Padhraic. (1996). *From Data Mining to Knowledge Discovery in Databases*. Retrieved from aaai.org.

Fishman, Tiffany Dovey, Ludgate Allan & Tutak, Jen. (2017). *Improving outcomes in American higher education*. Deloitte Insights. Retrieved from deloitte.com.

Fontaine, Michael. (2014). *Student Relationship Management (SRM) in Higher Education: Addressing the Expectations of an Ever Evolving Demographic and Its Impact on Retention*. American Research Institute for Policy Development. Retrieved from researchgate.net.

IGI Global dictionary. *What is Student Relationship Management (SRM)*. IGI Global. Retrieved from igi-global.com

Romero, Cristobal & Ventura, Sebastian. (2013). *Data mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* Retrieved from onlinelibrary.wiley.com

Rodrigues, M. W., Isotani, S., & Zárate, L. E. (2018). *Educational Data Mining: A review of evaluation process in the e-learning.* Telematics and Informatics. 35, 1701–1717. Retrieved from sciencedirect.com.

Ruffalo Noel Levitz. (2008). *Student Success, Retention, and Graduation: Definitions, Theories, Practices, Patterns, and Trends.* Stetson University. Retrieved from stetson.edu.

The App Solutions. (2018). *Data Mining: The Definitive Guide To Techniques, Examples, And Challenges*. Retrieved from theappsolutions.com.

U-Planner. (2017). *How to Improve Student Retention in Higher Education in the United States.* Retrieved from u-planner.com.

Vianden, J. & Barlow, P. J. (2015). *Strengthen the Bond: Relationships between Academic Advising Quality and Undergraduate Student Loyalty.* NACADA Journal. 35(2), 15–27. Retrieved from nacadajournal.org.

DECISION SCIENCES INSTITUTE

Is Independent Director Independent? Evidence from Corruption Culture of Independent Director

Robin
Yuan Ze University
Email: robin@uib.ac.id

Chia-Wei Huang
National Taiwan Normal University
Email: d94723002@ntu.edu.tw

Huai-Chun Lo
Yuan Ze University
Email: hclo@saturn.yzu.edu.tw

ABSTRACT

This paper investigates whether the independent director's corruption culture influences the corporate innovation and compensation of independent and CEO. Evidence shows that a stronger independent director's corruption culture is decreasing corporate innovation and increasing the compensation of independent director and CEO. However, the poor external corporate governances are weakly supported our results. Such evidence implies that the independent director's corruption culture less likely to monitor activities and imply in inefficient decisions and quality of corporate innovation and compensation policies.

KEYWORDS: Corruption culture, corporate innovation, excess compensation, corporate governance

INTRODUCTION

The general consensus in both the capital market and academic research are that an independent director plays an effective role in corporate governance. Practitioners and researchers suggest that less independent directors, large board members and insider boards, director hold three or more directorship (busy board), a capacity of strengthened by official positions and the connection to other corporate leaders (powerful CEO), CEO who also holds chairman position (duality CEO), and lower institutional ownership are less effective in corporate governance. For example, CEO likely to involved in independent director selection and create more conflicts of interest with replacing more directly independent director into non-directly connected to an insider (gray independent director) that less likely to pressure from monitoring CEO activities (Shivdasani & Yermack, 1999). CEO-director ties potential influence on new director selection with their network connections create less effective of monitoring, and acquisition frequently high-connection firms but destroy the value of shareholders (Fracassi & Tate, 2012). A new member of directors joined into board members after the CEO served (co-opted board) create a friendly in boards and leads a less board monitoring included the sensitivity of CEO turnover-performance, increase sensitivity of pay-performance, and increase investment activities (Coles et al, 2014).

The Sarbanes-Oxley Act of 2002 (SOX) requires a majority of an independent director to increasing their roles in corporate governance. For example, SOX enhanced independent director's workload and effectiveness of monitoring (Coates, 2007; Linck et al, 2009; Bargeron et al, 2010). However, is independent director still independent? We focus on the independent director (namely, independent director's corruption culture). Specifically, we investigate whether the independent director's corruption culture enhances or impedes corporate innovation and compensation of independent director and CEO.

Independent directors have an important role in corporate governance (Adams & Ferreira, 2007; Coles et al, 2008; Hwang & Kim, 2009; Armstrong et al, 2014; Falato et al., 2014; Ishii & Xuan, 2014; Kim et al, 2014; Masulis & Mobbs, 2014; Schmidt, 2015; Masulis & Mobbs, 2016). Independent director with a corruption culture more than the median in the industry (strong corruption culture) tends to be less active on monitoring and indicates that agency problem arise from independent director's corruption culture.

In practical cases, the scandal of *WorldCom* and *Enron* are related to the independent boards. Both cases settlements are required to paid personally (payments by director's pocket) to settle securities class action lawsuits. First, around the years 1997 – 2001 *Enron* was used improper accounting trick to restating their financial statements and inflated stock price by against senior executives and directors and *Enron* is the second-largest bankruptcy in the U.S. at the time with total damage is approximately $25 billion. In 2004, tenth boards agreed to pay $13 million of litigation expenses. Second, around the years 1999 – 2002 *WorldCom* CEO, CFO, and independent directors manipulated a billion dollar virtual transactions and accounting fraud and *WorldCom* is the largest bankruptcy firms in the U.S. at the time. In 2005, twelve boards agreed to pay $24.75 million of a settlement and total settlements over $6 billion for all defendants in securities class action. Beside *WorldCom* and *Enron*, there are eleventh cases of out-of-pocket settlement by independent directors. A consequence of those scandals, the role of the director has been reviewed.

In classical models, firms with the more independent director are obtained more patents and claim more citations (Balsmeier et al, 2017). Independent director increased the monitoring role that reduces agency problem. Generally, firms received more patents from the familiarity with their industry and citation increase will come from the number of patents that firms have it. Thus, innovation efficiency is a way to measures innovation performance. By innovation efficiency, firms may generate the number of patents per dollar of R&D expenditures that allows to compares the output and the input of investment (Hirshleifer et al, 2013). However, the poor corporate governance such as entrenched boards to pursue their own interest rather than shareholders lead less innovation performance (O'Connor & Rafferty, 2012). Further, the independent director may impede their competence and ignore the output of corporate innovation.

In compensation pieces of literature, the independent director and CEO are two parties that have a bargaining advantage to influences their size and structure of the compensation packages (Hermalin & Weisbach, 1998; Ryan & Wiggins, 2004). They also mention that firms with more independent directors are likely to receive more compensation and firms with entrancement CEO and duality are likely to receive cash rather than equity compensation. Furthermore, the compensation of independent director and CEO are positive relationships (Brick et al, 2006).

To test our hypothesis, we measure the independent director's corruption culture as the firm's corruption by using their average level of corruption. For each independent director, we match their surname-ancestry with Corruption Perception Index (CPI) to calculate the average of independent director's corruption culture follow the previous studies (DeBacker et al, 2015; Liu, 2016). We use a sample of 12,332 firm-year observation in the U.S. between 1996–2015, we find that the independent director's corruption culture decreases corporate innovation and increases the compensation of independent director and CEO.

First, our results show that strong independent director's corruption culture is significant to CEO turnover-performance sensitivity. The interaction effect of independent director's corruption culture and *Prior Abnormal Return* indicate that firms with a stronger independent director's corruption culture are associated with a weak in CEO turnover-performance sensitivity. Interestingly, the independent director's corruption culture rather than the percentage of independency board and have more explanatory power for CEO turnover-performance sensitivity. For example, the turnover-performance sensitivity is weaker when the CEO has the bargaining power on board selection (Hermalin & Weisbach, 1998). The classified boards reduce the effectiveness of monitoring in the replacement CEO (Faleye, 2007). More co-opted boards create friendly boards that diminish the CEO turnover-performance sensitivity on poor performance period (Coles et al, 2014).

Second, we argue that weaker monitoring from independent director impact on firm's overinvestment by taking high-risk projects without selection the project's quality and transmitted to poor quality of corporate innovation. Our results are also economically important. For example, the average R&D expenditures in our sample are 0.0396. Thus, firms with a stronger independent director's corruption culture are increases a 64.91% of R&D expenditures per each one-standard-deviation increase in independent director's corruption culture. Furthermore, firms with stronger independent director's corruption culture are decreases a 58.35% number of patents, 49.88% number of citations, and 63.82% innovation efficiency per each one-standard-deviation increase in independent director's corruption culture. For example, co-opted boards are likely to overinvestment in ways they otherwise would not (Coles et al, 2014).

Third, we also find similar results that the independent director's corruption culture increases the compensation of the independent director and CEO. We argue that less monitoring from independent director decreases their efforts and contribution to firms and the bargaining power on decided their own package of compensation, which is preferred to receive cash more than equity compensation. Our results are also economically significant. For example, the average of excess compensation of independent director in the sample is 0.0009. Thus, firms with stronger independent director's corruption culture are increases a 73.68% of excess compensation of independent director per each one-standard-deviation increase in independent director's corruption culture. Further, firms with a stronger independent director's corruption culture are increases 65.23% of cash compensation and decrease 63.70% of equity compensation of the independent director. Also increases 73.86% of CEO excess compensation and decreases 54.78% CEO risk-taking incentives per each one-standard-deviation increase in independent director's corruption culture. For example, both CEO and director have a bargaining power to decide their own compensation package by their own interest (Ryan & Wiggins, 2004). The excess compensation could be director and management is not considered shareholder interests (Brick et al, 2006).

However, the question of whether the independent director's corruption culture influence corporate innovation and compensation of independent director and CEO has been difficult to test due that independent director's corruption culture is likely to be endogenous. To address the endogeneity issues, we use two strategies. Our first strategy is relies on a natural experiment, the Sarbanes-Oxley Act of 2002 (Coles et al, 2014; Balsmeier et al, 2017), which directly affect the independent director's corruption culture but are exogenous to firm's corporate innovation and compensation. Using a difference-in-differences methodology (namely, DiD), we setup DiD isolate the effect of independent director's corruption culture ("Clean" effect). The typical of DiD is use three dummy variables (post-SOX, noncompliant firms, and the interaction term of post-SOX and noncompliant firms) and the main focus is an interaction term of post-SOX and noncompliant firms. This interaction term captures both effects that we want to eliminate (the exogenous shock to the independent director's corruption culture and the direct effect) though of SOX. However, the interaction term between post-SOX, noncompliant firms, and independent director's corruption culture (*IDCC*) does not yield the "Clean" estimate and the clean estimate is the increase of exogenous in independent director's corruption culture that forced noncompliant firms to increase their independent director composition, is given by independent director's corruption culture (*IDCC*), interaction term between noncompliant firms and *IDCC*, and interaction term between post-SOX, noncompliant firms, and *IDCC*. Further, the effect of independent director's corruption culture on corporate innovation and compensation of independent director and CEO are stronger after SOX that indicates the independent director's corruption culture damaging the culture of the firm.

Our second strategy is a propensity-score-matching (PSM) methodology to eliminate the self-selection bias problem. We sorted the level of independent director's corruption culture into quartiles and retain only the top quartile as the firms with highest independent director's corruption culture (treatment) and the bottom quartile as the firms with lowest independent director's corruption culture (control). The PSM methodology confirms that treatment groups are decrease their corporate innovation and increases the compensation of the independent director and CEO than the control groups. Overall, both strategies confirm that the independent director's corruption culture has a negative impact on corporate innovation and positive impact on compensation of independent director and CEO.

We then identify a conditional test through which strong independent director' corruption culture leads to a decrease of corporate innovation and increase the compensation of independent director and CEO. We use DiD in subsample analysis and we expect that the independent director's corruption culture decreases corporate innovation and increases the compensation for independent director and CEO with poor external corporate governance. This is because the independent director's corruption has a substitution effect on other corporate governance mechanisms. For example, the role of institutional holdings is monitoring firm activities and enhance the quality and transparency in corporate governance (Chung & Zhang, 2011; Armstrong et al, 2014). The product market competition allows tight monitoring creditor and effective to reduce inefficiency (Köke & Renneboog, 2005; Tian & Twite, 2011). Both proxies show that poor external corporate governance mechanisms worsen the independent director's corruption culture on corporate innovation and compensation of independent and CEO. However, the poor external corporate governances are weakly supported our evidence.

Furthermore, we conduct two robustness tests to corroborate our main results. (1) We replace the independent director's corruption culture with boards' corruption culture. (2) We control the CEO's corruption culture. Overall, all of our robustness tests support our main expectations.

The main contribution of this paper is to shed light a new set of facts that the independent director's corruption culture is poorer corporate governance. Thus, our paper uncovers the consequence of independent director's corruption culture on corporate innovation and compensation of independent director and CEO (Coles et al, 2014; Kim et al, 2014). Our finding highlight the representation of independent director's corruption culture decreases the effectiveness of their effort and monitoring, thereby the boards likely to approve all project without considering the risks. The consequences are firms spend more R&D expenditures (overinvestment) with lower quality of corporate innovation. Further, the director realizes themselves to avoid the effort to monitoring and less likely to replace cash with equity for their compensation. Either the CEO or the director, their compensation package is positive correlation because it depends on their effort and a certain level of skills (Brick et al, 2006). Firms with higher social ties awards more friendly director that reduce the effectiveness and their true independence of the director (Adams &Ferreira, 2007; Hwang & Kim, 2009; Schmidt, 2015). The director will likely to choose CEO to depend on their familiarity to avoid an effort to themselves (Kumar & Sivaramakrishnan, 2008). Over workload of boards that also hold several directorships in other firms relatively less time to focus on monitoring quality and associated with lower shareholder value (Falato et al, 2014). Lastly, the co-opted boards are more tolerance to the CEO and not all independent director aware with the liability (Coles et al, 2014).

Our paper also contributes to the growing literature on corporate innovation. Several studies have focused on structure of compensation (Holthausen et al, 1995; Cheng, 2004; Lerner & Wulf, 2007; Manso, 2011; Ederer & Manso, 2013; Baranchuk et al, 2014; Chang et al, 2015), corporate governance (O'Connor & Rafferty, 2012; Balsmeier et al, 2017), bank deregulation and competition (Amore et al, 2013; Chava et al, 2013; Cornaggia et al, 2015), analyst coverage (He & Tian, 2013), stock returns (Hirshleifer et al, 2013), bribe payments (Ayyagari et al, 2014), stock liquidity (Fang et al, 2014), economic growth (Kogan et al, 2017), and transparency (Zhong, 2018). Our paper complements these studies by exploring how the independent director's corruption culture influences corporate innovation.

Another contributes to recent studies on the determinants of compensation of independent director and CEO, including board control, activity, and structure (Boyd, 1994; Chhaochharia & Grinstein, 2009; Nguyen, 2014), merger and acquisition (Bliss & Rosen, 2001), CEO's power (Grinstein & Hribar, 2004; Ryan & Wiggins, 2004), corporate risk-taking (Linn & Park, 2005; Coles et al, 2006), CEO optimism and executive overconfidence (Otto, 2014; Humphery-Jenner et al, 2016), liability protection (Aguir et al, 2014), financial accounting quality (Masulis & Mobbs, 2016), excessive compensation (Dah & Frye, 2017), board qualifications (Fedaseyeu et al, 2018). We complement these studies by showing that the independent director's corruption culture influences compensation policy.

LITERATURE REVIEW

Independent Director: Bright Side and Dark Side

A large of empirical studies have examined the influence and the role of director characteristics. For example, the detection and litigation fraud (Beasley, 1996; Brochet & Srinivasan, 2014), the regulation relevance of boards (Wintoki, 2007; Dah et al, 2014). Much of previous studies examine how independent director play their roles in monitoring firms (e.g., Coles et al, 2008; Armstrong et al, 2014; Kim et al, 2014; Masulis & Mobbs, 2016), reduce earnings management (Klein, 2002; Fogel et al, 2014), shareholders wealth (Rosenstein & Wyatt, 1990; Cotter et al, 1997; Nguyen & Nielsen, 2010; Knyazeva et al, 2013). These studies are the bright side of independent director functions.

On the other side, the independent director also may harm board functions. The director may improve the monitoring role, but another agency problem arises because they avoid the effort and relying on CEO (Kumar & Sivaramakrishnan, 2008). The social ties are associated with compensation, the sensitivity of CEO turnover, and poorer acquiring firms performance when board advises surpasses to supervision CEO (Hwang & Kim, 2009; Fracassi & Tate, 2012; Ishii & Xuan, 2014; Schmidt, 2015). CEO with co-option is less monitored by boards and received more compensation and less likely to the turnover on poor performance period (Coles et al, 2014). Over workload for the director is decrease their monitoring quality and hurt the shareholder value (Falato et al, 2014).

Independent Director's Corruption Culture and CEO Turnover-performance Sensitivity

One of the roles of the director is to evaluate CEO turnover-performance. For example, CEO turnover is related to poor firm's performance (Mace, 1971). The classified boards reduce the effectiveness of monitoring in the replacement CEO (Faleye, 2007). Also, turnover-performance sensitivity is weaker when the CEOs have powerful bargaining on board-selection and more co-opted boards (Hermalin & Weisbach, 1998; Coles et al, 2014). Thus, we expect that our first hypothesis as follows:

H1: All else equal, the forced CEO turnover-performance sensitivity decreases with the independent director's corruption culture.

Independent Director's Corruption Culture and Corporate Innovation

Corporate innovation issues become important in recent literature on corporate finance. For example, the majority of independent director is associated with the output quality of corporate innovation (Balsmeier et al, 2017; Lu & Wang, 2018). However, CEO with co-opted boards likely to influence the overinvestment behavior (Coles et al, 2014). Thus, we expect our second hypothesis as follows:

H2a: All else equal, R&D expenditures increases with the independent director's corruption culture.

H2b: All else equal, patent decreases with the independent director's corruption culture.

H2c: All else equal, citation decreases with the independent director's corruption culture.

H2d: All else equal, innovation efficiency decreases with the independent director's corruption culture.

Independent Director's Corruption Culture and Compensation of Independent Director

A large literature addresses the compensation of the board of directors. For example, the key problem of excess compensation is the friendly boards that independent director likely to bargain with managers included CEO, influence management in the pay-setting process, and expected to support rather monitor the CEO activities (Bebchuk et al, 2002). The excess compensation could be a director's culture problem because both director and management are not considered shareholder interests (Brick et al, 2006). The firm with more independent boards are awarded more compensation and both CEO and director have a bargaining power to decided their own compensation package by their own interest (Ryan & Wiggins, 2004). Further, the director of the firm with entrenched CEO and CEO duality is likely to receive cash compensation more than equity compensation. Also, the cash compensation and equity compensation is negatively relationships (Humphery-Jenner et al, 2016). Thus, we expect that our third hypothesis as follows:

H3a: All else equal, excess compensation of independent director is increases with the independent director's corruption culture.

H3b: All else equal, cash compensation increases with the independent director's corruption culture.

H3c: All else equal, equity compensation decreases with the independent director's corruption culture.

Independent Director's Corruption Culture and Compensation of CEO

The CEO's compensation structure is another concern of the board's function. For example, a CEO with friendly boards has a substantial power to influence their own package of compensation when the director's monitoring is weak (Bebchuk & Fried, 2003). The classified boards tend less likely to fire CEOs from poor performance by reducing their compensation to satisfy the shareholders (Faleye, 2007). The classified boards also require higher monitoring costs that hurt the CEO risk-incentives (Ahn & Shresthak, 2013). The excess compensation of director leads to less monitoring and decreases CEO pay-for-performance sensitivity and both excess compensation of independent director and CEO are positively related (Brick et al, 2006; Dah & Frye, 2017). The co-opted boards and CEO pay-for-performance sensitivity cannot be viewed as directly and the CEO requires higher pay when seeking greater risk, thus the co-opted boards apply liberal compensation that is likely to the CEO (Coles et al, 2014). We expect that our fourth hypothesis as follows:

H4a: All else equal, excess compensation of CEO increases with the independent director's corruption culture.

H4b: All else equal, CEO risk-incentives decreases with the independent director's corruption culture.

METHODS

Data and Variables

The main data set we built up in the empirical analysis, we joint availability of data on boards and entrenchment index from the RiskMetrics (IRRC), accounting data from Compustat, stock price data from CRSP, patent and citation data from 2016 edition of PATSTAT with supplement data from Kogan et al. (2017) and the United States Patent and Trademark Office (USPTO) database. Compensation data from ExecuComp, CEO risk-taking incentives (*CEO_Vega*) data from Coles et al. (2006) and Core and Guay (2002). Institutional ownership data is from the Thomson Reuters Institutional (13F) Holdings and Product market competition data is from Hoberg and Phillips (2010, 2016). The final sample consists of 12,332 firm-year observations exclude financial firms (SIC codes 6000–6999) during the period of 1996 – 2015.

Description of Independent Variable

We conduct the main explanatory variable is the independent director's corruption culture, based on the previous research (DeBacker et al, 2015; Liu, 2016). We calculate the average of each independent director's corruption culture index by their surname-ancestry country. Then we matching into the Transparency International release the corruption perception index (CPI) each country per year. In 2015, the CPI ranges from 0.9 (Denmark) to 9.2 (Somalia), where the U.S. index is 2.4.

Following their proxy to measure the corruption, we use more than 139,500 independent director's surname-ancestry to capture corruption behavior. Our data set for surname-ancestry is from Origins Info Ltd. For the surname-ancestry are unmatched, we hand-collect it using sources from ancestry.com.

We focus on the independent director for several reasons. (1) Independent director has an important role in the corporate governance (Armstrong et al, 2014; Masulis & Mobbs, 2014; Faleye et al, 2011; Adams et al, 2010; Nguyen & Nielsen, 2010). (2), The regulations requirement of board structure (Dah et al, 2014; Bargeron et al, 2010). (3) The independent director is also responsible for the corporate decision such as corporate innovation and compensation (Kumar & Sivaramakrishnan, 2008; Balsmeier et al, 2017).

Description of Dependent Variables

To examine whether the independent director's corruption culture can impact on corporate innovation and compensation, we examine four types of corporate innovation and five types of compensation that can be measured in previously studied.

We measure corporate innovation into four dependent variables. (1) *RD* is the total R&D expenditures scaled by total sales and the missing of R&D expenditures is equal to zero (Ciftci & Cready, 2011). (2) The *Patent* is the natural logarithm of one plus the weight-adjusted patent counts. (3) *Citation* is the natural logarithm of one plus the number of citation lag-adjusted citations per patent. We measure the *Patent* and *Citation* following previous studies (Balsmeier et al, 2017). (4) innovation efficiency (*IE*) is the natural logarithm of one plus the number of patents scaled by R&D expenditures (Hirshleifer et al, 2013).

We also measure the independent director and CEO compensation into five dependent variables. (1) excess compensation of independent director (*ID_Compensation*) is the estimated value by the residual from independent director's total compensation regression where the explanatory variables include firm size, ROA, leverage, board size, board independency, director age, female director, director tenure, and year and industry fixed effects (Masulis et al, 2009). (2) *Cash Compensation* is the total cash received by the independent director scaled by total compensation. (3) *Equity Compensation* is the total stock plus options received by independent director scaled by total compensation (Ryan & Wiggins, 2004). (4) excess compensation of CEO (*CEO_Compensation*) is the estimated value by the residual from CEO compensation regression where the explanatory variables include firm size, ROA, leverage, board size, board independency, CEO age, female CEO, CEO tenure, and year and industry fixed effects (Masulis et al, 2009). (5) *CEO Vega* is the natural logarithm of one plus stock return volatility sensitivity change in CEO wealth (Coles et al, 2006).

Control Variables

Following the previous pieces of literature, we control for a vector of firm characteristics, governance characteristics, and independent director characteristics that could disrupt the relation between independent director's corruption culture and firm's corporate innovation and compensation. We compute all variables form firm *i* in year *t-1*. *ROA* is the total net income scaled by total assets. *Leverage* is the current liabilities plus long term debt scaled by total assets. *BHR* is the buy and hold one year return. *Firm Size* is the natural logarithm of total assets. *Dividend* is the total dividend payment scaled by total assets. *E-Index* is a corporate action to boost managerial perceived value rather than to benefit firm financial based on six provisions in RiskMetrics. *HHI* is the size of firms in relation to the same industry and measurement of market concentration. *IO* is the percentage of shares held by institutional ownership. *Board Size* is the natural logarithm of the total number of directors on board. *Board Independency* is the percentage of independent board scaled by board size. *Director Age* is the natural logarithm of the average independent director's age. *Female Director* is a dummy variable that equal to one if the board has a female director and otherwise is zero. *Director Tenure* is the natural logarithm of the average independent director's tenure.

Summary Statistics

In Table 1, we provide the summary statistics of the independent director's corruption culture, corporate innovation, compensation of independent director and CEO, and other control variables. All of the variables are winsorized at 1st and 99th percentiles to minimize the influence of outliers. The average of independent director's corruption culture is 2.5199 with 0.6446 of standard deviation, which is similar to previous research on corruption culture.

In our sample, the average of corporate innovation variables is 3.96% of R&D expenditures, two patents (natural logarithm of 0.7149), two citations (natural logarithm of 0.8325), and 5.27% of innovation efficiency. The average of compensation variables is 0.09% of excess compensation of independent director, 40.22% of cash compensation of independent director, 59.78% of equity compensation of independent director, 0.50% of excess compensation of CEO, and CEO with a risk-taking incentives sensitivity of US$ 19,300 (natural logarithm of 2.9593).

Regarding other variables, the average of book value assets of US$ 0.99 billion, ROA of 5.25%, leverage of 21.91%, buy and hold one-year returns of 12.38%, dividend of 34.47%, nine board members with 74.46% of independent director.

Table 1: Summary statistics

Variable	N	Mean	Q1	Median	Q3	Std Dev
IDCC	12,332	2.5199	2.0667	2.4750	2.9222	0.6466
RD	12,329	0.0396	0.0000	0.0000	0.0410	0.0735
Patent	12,329	0.7149	0.0000	0.0000	0.6932	1.3409
Citation	12,329	0.8325	0.0000	0.0000	0.0000	1.7617
IE	12,329	0.0527	0.0000	0.0000	0.0000	0.1439
ID_Compensation	11,327	0.0009	-0.5256	0.1718	0.5546	0.9198
Cash_Compensation	11,491	0.4022	0.2447	0.4051	0.5169	0.2346
Equity_Compensation	11,491	0.5978	0.4831	0.5949	0.7553	0.2346
CEO_Compensation	10,857	0.0050	-0.2928	-0.0232	0.2439	0.5058
CEO_Vega	9,476	2.9593	1.7884	3.0778	4.2006	1.7662
ROA	12,291	0.0525	0.0250	0.0540	0.0919	0.0805
Leverage	12,253	0.2191	0.0678	0.2127	0.3303	0.1709
BHR	12,215	0.1238	-0.1290	0.0980	0.3219	0.4183
Firm Size	12,290	6.9035	5.9162	6.9192	7.9381	1.4896
Dividend	12,259	0.3447	0.0000	0.0073	0.0295	1.0072
E-Index	11,973	2.7152	2.0000	3.0000	4.0000	1.5110
IO	12,332	0.3147	0.0728	0.2492	0.5197	0.2756
HHI	12,332	0.0576	0.0313	0.0471	0.0671	0.0430
Board Size	12,328	9	7	9	11	1.33
Board Independency	12,328	0.7446	0.6667	0.7778	0.8750	0.1541
Director Age	12,328	54.82	51.50	56.50	61.82	1.42
Female Director	12,328	0.6861	0.0000	1.0000	1.0000	0.4641
Director Tenure	12,328	7.72	6.67	8.86	11.17	1.90

This table reports the summary statistics of all variables. The sample period is from 1996 – 2015 and all of the variables are winsorized at 1st and 99th percentiles to minimize the influence of outliers.

In Table 2, we report the preliminary univariate test. Most of the mean differences show a significant difference at the 1% level. Firms with strong independent director's corruption culture tend to increases R&D expenditures and decreases patents, citations, and innovation efficiency than those with weak independent director's corruption culture. We also find a similar pattern for compensation, which is firms with strong independent director's corruption culture tend to increases compensation of independent director and CEO than those weak independent director's corruption culture.

Table 2: Mean difference of strong and weak corruption culture

Variables	Strong *IDCC*	Weak *IDCC*	Difference	t-Statistics
Panel A: Dependent variables				
RD	0.0445	0.0347	0.0098***	(7.44)
Patent	0.6565	0.7735	-0.1170***	(-4.85)
Citation	0.5993	1.0668	-0.4675***	(-14.86)
IE	0.0442	0.0611	-0.0169***	(-6.54)
ID_Compensation	0.0215	-0.0229	0.0444***	(3.23)
Cash_Compensation	0.4151	0.3884	0.0267***	(6.11)
Equity_Compensation	0.5849	0.6116	-0.0267***	(-6.11)
CEO_Compensation	0.0160	-0.0167	0.0327***	(3.22)
CEO_Vega	2.9072	3.0101	-0.1029***	(-2.84)
Panel B: Firm characteristics				
ROA	0.0492	0.0445	0.0047***	(2.86)
Leverage	0.2165	0.2355	-0.0019***	(-6.01)
BHR	0.1484	0.0991	0.0493***	(-6.53)
Firm Size	6.9085	6.8986	0.0099	(0.37)
Dividend	0.3942	0.2948	0.0994***	(5.47)
Panel C: Corporate Governance characteristics				
E-Index	2.9454	2.4763	0.4691***	(17.19)
IO	0.0004	0.0002	0.0002	(1.52)
HHI	0.0601	0.055	0.0051***	(6.58)
Board Size	2.1722	2.1689	0.0033	(0.65)
Board Independency	0.7738	0.7153	0.0585***	(21.46)
Panel D: Independent Director characteristics				
Director Age	3.9770	4.0311	-0.0541***	(-8.54)
Female Director	0.7273	0.6446	0.0827***	(9.94)
Director Tenure	2.1371	1.9508	0.1863***	(16.25)

This table reports the mean difference between firms with strong and weak independent director's corruption culture. Superscripts *, ** and *** denote significance of the *t*-test for the difference in the means between the two subsamples at the 10%, 5%, and 1% levels, respectively.

RESULTS AND DISCUSSION

Independent Director's Corruption Culture and CEO Turnover-performance Sensitivity

To examine the effect of the independent director's corruption culture on CEO turnover-performance sensitivity, we follow the logistic regressions setting of Coles et al. (2014) as follow:

$$\ln\left[\frac{Prob(Forced\ Turnover)}{1-Prob(Forced\ Turnover)}\right]=\alpha_0+\alpha_1 IDCC_{i,t-1}\ X\ Performance_{i,t-1}$$
$$+\alpha_2 Performance_{i,t-1}+\alpha_3 IDCC_{i,t-1}+\boldsymbol{\theta}'\boldsymbol{Z}_{i,t-1}+\gamma_i+\mu_t+\varepsilon_{i,t}, \quad (1)$$

Table 3: *IDCC* and CEO turnover-performance sensitivity

	Value-Weighted		Equal-Weighted	
	(1)	(2)	(3)	(4)
Constant	-0.2120	50.6288***	-0.3827	50.4036***
	(-0.33)	(8.76)	(-0.59)	(8.76)
IDCC X Prior Abnormal Return	0.8486**	1.2376***	0.9350***	1.1824***
	(2.41)	(3.39)	(2.76)	(3.38)
	0.1226	*0.1494*	*0.1348*	*0.1427*
Prior Abnormal Return	-2.1696***	-3.0768***	-2.3220***	-2.9400***
	(-2.64)	(-3.60)	(-2.99)	(-3.57)
	-0.3134	*-0.3714*	*-0.3348*	*-0.3549*
IDCC	-0.0905	-0.2925**	-0.0079	-0.2037
	(-0.76)	(-2.11)	(-0.06)	(-1.36)
	-0.0131	*-0.0353*	*-0.0011*	*-0.0246*
ROA		0.0033		0.0034
		(0.09)		(0.09)
		0.0004		*0.0004*
Leverage		-0.0056		-0.0047
		(-0.70)		(-0.59)
		-0.0007		*-0.0006*
BHR		0.3805*		0.3939*
		(1.82)		(1.89)
		0.0459		*0.0475*
Firm Size		-0.0162		-0.0099
		(-0.25)		(-0.15)
		-0.0019		*-0.0012*
Dividend		-0.0618		-0.0712
		(-0.51)		(-0.59)
		-0.0075		*-0.0086*
E-Index		-0.0147		-0.0146
		(-0.19)		(-0.19)
		-0.0018		*-0.0018*
IO		6.6604***		6.6162***
		(3.13)		(3.09)
		0.8039		*0.7987*
HHI		-0.6748		-0.9398
		(-0.09)		(-0.13)
		-0.0815		*-0.1135*
Board Size		-0.6829**		-0.6487**
		(-2.11)		(-1.97)
		-0.0824		*-0.0783*
Board Independency		0.3250		0.2799
		(0.55)		(0.48)

		0.0392		*0.0338*
Director Age		-12.2657***		-12.2730***
		(-8.71)		(-8.72)
		-1.4806		*-1.4817*
Female Director		0.2120		0.2017
		(1.13)		(1.07)
		0.0256		*0.0244*
Director Tenure		-0.7032***		-0.6829***
		(-3.70)		(-3.59)
		-0.0849		*-0.0824*
Control for:				
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Obs.	1,759	1,697	1,759	1,697
Pseudo R^2	0.3494	0.4439	0.3503	0.4438

where the proxy for performance is *Prior Abnormal Return*. $IDCC_{t-1}$ represents the independent director's corruption culture index for firm *i* in year *t-1*. $Z_{i,t-1}$ is the vector of the control variables from firm, corporate governance, and independent director factor *i* in year *t-1*. γ_i , μ_t and $\varepsilon_{i,t}$ represent Fama-French 48 industries and year fixed effect and the error of the regression. In all models, the *t*-values are computed on the heteroskedasticity-robust standard errors (White, 1980). Coefficients: ***, **, and * denote significant at 1%, 5%, and 10% levels, respectively.

Table 3 reports the effect of the independent director's corruption culture (*IDCC*) on CEO turnover-performance sensitivity. In models (1) - (4), the focus variable is the interaction term of *IDCC* and *Prior Abnormal Return*. We report the first row is the estimate of coefficient, the second row is z-statistics, and the third row is the marginal effects. The marginal effect represents the probability of *Forced Turnover* percentage change with *Prior Abnormal Return* for an independent director's corruption culture.

Consistent with our first hypothesis that the coefficient on the interaction effect of *IDCC* and *Prior Abnormal Return* (α_1) is positive and significant in models (1) - (4) that indicate that firms with a stronger independent director's corruption culture are associated with a weak in CEO turnover-performance sensitivity. The marginal effects show in row three indicates how the change of the independent director's corruption culture on CEO turnover-performance sensitivity. In model (1), the CEO turnover-performance sensitivity decreases by 0.1226, from -0.3134 to -0.1908 at the mean value when the independent director's corruption culture increases by one standard deviation. Interestingly, the independent director's corruption culture rather than the percentage of independency board and have more explanatory power for CEO turnover-performance sensitivity. For example, turnover-performance sensitivity is weaker when the CEO has the bargaining power on board selection (Hermalin & Weisbach, 1998). The classified boards reduce the effectiveness of monitoring in the replacement CEO (Faleye, 2007). More co-opted boards create friendly boards that diminish the CEO turnover-performance sensitivity on poor performance period (Coles et al, 2014).

Independent Director's Corruption Culture and Corporate Innovation

We first examine whether the independent director's corruption culture impact on corporate innovation. We divided corporate innovation into R&D expenditures (*RD*), the number of patents (*Patent*), the number of citation (*Citation*), and innovation efficiency (*IE*).

Table 4: *IDCC* and corporate innovation

	RD	*Patent*	*Citation*	*IE*
	(1)	(2)	(3)	(4)
Constant	0.0974***	1.0941***	2.8431***	0.2510***
	(7.74)	(4.86)	(8.26)	(7.80)
IDCC	0.0069***	-0.0996***	-0.2564***	-0.0100***
	(7.17)	(-5.51)	(-10.76)	(-4.73)
ROA	-0.1008***	-0.1073	-0.4948**	-0.0057
	(-9.90)	(-0.72)	(-2.35)	(-0.33)
Leverage	-0.0002***	0.0124***	0.0133***	-0.0007***
	(-4.39)	(7.40)	(6.18)	(-4.86)
BHR	-0.0038***	0.0396	0.0177	0.0035
	(-2.76)	(1.47)	(0.47)	(1.14)
Firm Size	-0.0023***	0.0999***	0.0994***	-0.0045***
	(-4.35)	(8.69)	(6.81)	(-4.15)
Dividend	0.0009	0.0631***	0.0479*	-0.0038***
	(1.48)	(3.04)	(1.95)	(-2.99)
E-Index	-0.0014***	-0.1148***	-0.1572***	-0.0081***
	(-3.90)	(-14.92)	(-16.03)	(-9.52)
IO	-0.0116***	-0.3957***	-0.5635***	-0.0211***
	(-4.58)	(-8.68)	(-9.02)	(-3.66)
HHI	-0.0453**	-4.4060***	-7.1080***	-0.6585***
	(-2.03)	(-6.93)	(-6.49)	(-6.14)
Board Size	-0.0198***	0.0171	-0.0402	-0.0056
	(-8.36)	(0.37)	(-0.62)	(-1.10)
Board Independency	0.0079*	0.6320***	0.4493***	0.0575***
	(1.96)	(7.69)	(3.73)	(6.10)
Director Age	0.0009	-0.1013***	-0.0412	-0.0076**
	(1.01)	(-3.76)	(-1.58)	(-1.97)
Female Director	-0.0056***	0.0099	0.0005	-0.0070**
	(-4.07)	(0.36)	(0.01)	(-2.04)
Director Tenure	-0.0014	0.0121	-0.0507	-0.0006
	(-1.33)	(0.53)	(-1.39)	(-0.19)
Control for:				
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Obs.	11,784	11,784	11,784	11,784
Adj. R^2	0.4660	0.2658	0.2702	0.1817

Table 4 presents the results. The coefficient in the model (1) is significantly positive, implying that R&D expenditures increase with the independent director’s corruption culture. The coefficient of 0.0069 indicates that the independent director’s corruption culture is associated with an increase in R&D expenditures. The average R&D expenditures in our sample are 0.0396. Thus, firms with a stronger independent director’s corruption culture are increases a 64.91% ($e^{0.0069}$ x 0.6446) R&D expenditures per each one-standard-deviation increase in independent director’s corruption culture. The coefficient in the model (2) is significantly negative, implying that the number of patents decrease with the independent director’s corruption culture. The coefficient of -0.0996 indicates that the independent director's corruption culture is associated with a decrease in the number of patents. The average number of patents in our sample is 0.7149. Thus, firms with a stronger independent director's corruption culture are decreases a 58.35% ($e^{(-0.0996)}$ x 0.6446) number of patents per each one-standard-deviation increase in independent director's corruption culture. The coefficient in the model (3) is significantly negative, implying that the number of citations decrease with the independent director's corruption culture. The coefficient of -0.2564 indicates that the independent director's corruption is associated with a decrease in the number of citations. The average number of citations in our sample is 0.8325. Thus, firms with a stronger independent director's corruption culture are decreases a 49.88% ($e^{(-0.2564)}$ x 0.6446) number of citations per each one-standard-deviation increase in independent director's corruption culture. The coefficient in the model (4) is significantly negative, implying that innovation efficiency decrease with the independent director's corruption culture. The coefficient of -0.0100 indicates that the independent director's corruption culture is associated with a decrease in innovation efficiency. The average of innovation efficiency in our sample is 0.0527. Thus, firms with stronger independent board's corruption are decreases a 63.82% ($e^{(-0.0100)}$ x 0.6446) innovation efficiency per each one-standard-deviation increase in independent director's corruption culture.

Therefore, the unfavorable effect of the independent director's corruption culture on corporate innovation is also economically important. The results indicate that firms with strong independent director's corruption culture likely to overinvestment because taking high-risk projects. For example, co-opted boards are likely to overinvestment in ways they otherwise would not (Coles et al, 2014). Also suggesting less monitoring leads and transmitted to decreased efforts, which decreased the number of patenting of inventions, the number of citations, and innovation efficiency.

Independent Director's Corruption Culture and Compensation of Independent Director

We examine whether the independent director's corruption culture can impact on compensation of the independent director. We divided compensation into excess compensation of independent director (*ID_Compensation*), *Cash_Compensation*, and *Equity_Compensation*.

Table 5: *IDCC* and compensation of independent director			
	ID_Compensation	*Cash_Compensation*	*Equity_Compensation*
	(1)	(2)	(3)
Constant	-0.8291**	0.2734***	0.7266***
	(-2.40)	(3.54)	(9.42)
IDCC	0.1337***	0.0118***	-0.0118***
	(8.83)	(3.14)	(-3.14)
ROA	0.2347**	-0.0951***	0.0951***
	(2.24)	(-3.38)	(3.38)
Leverage	-0.0005	-0.0003	0.0003
	(-0.65)	(-1.03)	(1.03)
BHR	0.1669***	-0.0238***	0.0238***
	(6.85)	(-4.51)	(4.51)
Firm Size	-0.0257***	-0.0020	0.0020
	(-3.69)	(-1.02)	(1.02)
Dividend	-0.0512***	0.0066***	-0.0066***

	(-6.74)	(2.66)	(-2.66)
E-Index	0.0408***	0.0127***	-0.0127***
	(5.56)	(8.54)	(-8.54)
IO	-0.4236***	0.0871***	-0.0871***
	(-12.13)	(8.49)	(-8.49)
HHI	1.4722**	0.4428**	-0.4428**
	(2.29)	(2.37)	(-2.37)
Board Size	-0.0503	0.0136	-0.0136
	(-1.23)	(1.18)	(-1.18)
Board Independency	0.2592***	-0.1385***	0.1385***
	(3.42)	(-6.87)	(6.87)
Director Age	0.0074	-0.0358***	0.0358***
	(0.28)	(-4.24)	(4.24)
Female Director	0.0082	0.0083	-0.0083
	(0.42)	(1.46)	(-1.46)
Director Tenure	-0.0040	0.0268***	-0.0268***
	(-0.17)	(4.32)	(-4.32)
Control for:			
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs.	10,952	11,059	11,059
Adj. R^2	0.4245	0.1312	0.1312

Table 5 presents the results. The coefficient in the model (1) is significantly positive, implying that excess compensation of independent director increase with the independent director's corruption culture. The coefficient of 0.1337 indicates that the independent director's corruption culture is associated with an increase in excess compensation of the independent director. The average of excess compensation of independent director in our sample is 0.0009. Thus, firms with a stronger independent director's corruption culture are increases a 73.68% ($e^{0.1337}$ x 0.6446) excess compensation of independent director per each one-standard-deviation increase in independent director's corruption culture. The coefficient in the model (2) is significantly positive, implying that cash compensation of independent director increase with the independent director's corruption culture. The coefficient of 0.0118 indicates that the independent director's corruption culture is associated with an increase in cash compensation of the independent director. The average cash compensation of independent director in our sample is 0.4022. Thus, firms with a stronger independent board's corruption are increases a 65.23% ($e^{0.0118}$ x 0.6446) cash compensation of independent director per each one-standard-deviation increase in independent director's corruption culture. The coefficient in the model (3) is significantly negative, implying that equity compensation of independent boards decrease with the independent director's corruption culture. The coefficient of -0.0118 indicates that the independent director's corruption culture is associated with a decrease in equity compensation of the independent director. The average of equity compensation of independent director in our sample is 0.5978. Thus, firms with a stronger independent board's corruption are decreases a 63.70% ($e^{(-0.0118)}$ x 0.6446) equity compensation of independent director per each one-standard-deviation increase in independent director's corruption culture.

Therefore, the unfavorable effect of the independent director's corruption culture on the compensation of the independent director is also economically important. The results are suggesting the bargaining power of an independent director, which increased excess compensation of the independent director. Also likely to receive cash rather than equity compensation. For example, both CEO and director have a bargaining power to decide their own compensation package by their own interest (Ryan & Wiggins, 2004). The excess compensation could be director and management is not considered shareholder interests (Brick et al, 2006).

Independent Director's Corruption Culture and Compensation of CEO

We examine whether the independent director's corruption culture can impact on compensation of the CEO. We divided compensation into excess compensation of CEO (*CEO_Compensation*) and CEO risk-taking incentives (*CEO_Vega*).

Table 6: *IDCC* and compensation of CEO

	CEO_Compensation	*CEO_Vega*
	(1)	(2)
Constant	-0.5326**	1.6508***
	(-2.40)	(3.49)
IDCC	0.1361***	-0.1628***
	(14.36)	(-6.02)
ROA	0.1271*	1.3376***
	(1.73)	(6.73)
Leverage	-0.0014*	0.0327***
	(-1.68)	(13.52)
BHR	0.1613***	0.1116***
	(10.34)	(2.96)
Firm Size	-0.0065	0.2802***
	(-1.23)	(17.54)
Dividend	-0.0117*	0.2056***
	(-1.66)	(7.78)
E-Index	0.0090*	-0.0116
	(1.79)	(-0.91)
IO	-0.2451***	-1.1085***
	(-10.66)	(-15.96)
HHI	1.7866***	-5.8812***
	(4.90)	(-5.13)
Board Size	-0.1119***	0.4712***
	(-4.24)	(6.26)
Board Independency	0.4513***	0.2168*
	(9.64)	(1.71)
Director Age	-0.0101	-0.0304
	(-0.57)	(-0.60)
Female Director	0.0462***	0.1312***
	(3.59)	(3.30)
Director Tenure	-0.0076	0.0280
	(-0.47)	(0.87)
Control for:		
Year FE	Yes	Yes

Industry FE	Yes	Yes
Obs.	10,503	9,121
Adj. R^2	0.2346	0.2411

Table 6 presents the results. The coefficient in the model (1) is significantly positively, implying that excess compensation of CEO increase with the independent director's corruption culture. The coefficient of 0.1361 indicates that the independent director's corruption culture is associated with an increase in excess compensation of the CEO. The average of excess compensation of the CEO in our sample is 0.0050. Thus, firms with a stronger independent board's corruption are decreases a 73.86% ($e^{(0.1361)}$ x 0.6446) excess compensation of CEO per each one-standard-deviation increase in independent director's corruption culture. The coefficient in the model (2) is significantly negative, implying that CEO risk-taking incentives decrease with independent director's corruption. The coefficient of -0.1628 indicates that the independent director's corruption culture is associated with a decrease in CEO risk-taking incentives. The average of CEO risk-taking incentives in our sample is 2.9593. Thus, firms with a stronger independent board's corruption are decreases a 54.78% ($e^{(-0.1628)}$ x 0.6446) CEO risk-taking incentives per each one-standard-deviation increase in independent director's corruption culture.

Therefore, the unfavorable effect of the independent director's corruption culture on the compensation of the CEO is also economically important. The results are suggesting the bargaining power of the CEO, which increased excess compensation of the CEO. For example, both CEO and director have a bargaining power to decide their own compensation package by their own interest (Ryan & Wiggins, 2004). Independent director's corruption culture adopting liberal compensation that likely to the CEO, which decreased CEO risk-taking incentives (Coles et al, 2014).

Endogeneity Issues

One of the important concerns in corporate governance literature is endogeneity issues. It is possible that endogeneity issues influence the effect of independent director's corruption culture on corporate innovation and compensation of independent director and CEO. First, if the independent director's corruption culture correlates with an omitted variable, then the estimator is biased and inconsistent. We solve an omitted variable with difference-in-dfferences (DiD). Second, the hire of corrupt independent directors as boards could be a concern for the self-selection bias effect, we following Rosenbaum and Rubin (1983) to solve the self-selection bias with propensity-score-matching (PSM). We use two ways to process to reduce the effect of the independent board's corruption culture on corporate innovation and compensation of independent director and CEO.

<u>Difference-in-Differences (DiD) Methodology</u>

We eliminate the endogeneity of the independent director's corruption culture using a difference-in-differences (DiD) methodology. To eliminate the omitted variable effect of independent director's corruption culture, we turn to a natural experiment to address endogeneity issues. We use SOX as an exogenous shock in DiD methodology. Following by the rules of the Sarbanes-Oxley Act of 2002 (SOX), listed firms are required to increase their independent director percentage on their board more than 50%. We define post-SOX is the period after SOX is applied in NASDAQ and NYSE stock exchanges (2002 – 2015). Pre-SOX noncompliant firms are the firms that increase their independent director proportion after the regulation of SOX and these firms following the regulation to add more new independent directors (Linck et al, 2009).

We allow the possibility of SOX have a direct effect on corporate innovation and compensation of independent director and CEO. This is because SOX is a regulation that likely to affected monitoring firms and provides long-term benefits through several channels. For example, the regulation of SOX increased the role of independent boards for audit, compensation, and monitoring committees.

We setup DiD isolate the effect of independent director's corruption culture ("Clean" effect) (Coles et al, 2014). The typical of DiD is use three dummy variables (namely, *Post-SOX*, *Noncompliant*, and *Post-SOX x Noncompliant*) and the main focus is an interaction term of *Post-SOX x Noncompliant*. This interaction term captures both effects that we want to eliminate (the exogenous shock to the independent director's corruption culture and the direct effect) though of SOX. The DiD model is as follow:

$$Y = \beta_0 + \beta_1 IDCC + \beta_2 Post - SOX\ x\ IDCC + \beta_3 Noncompliant\ x\ IDCC$$
$$+\beta_4 Post - SOX\ x\ Noncompliant\ x\ IDCC + \beta_5 Post - SOX + \beta_6 Noncompliant$$
$$+\boldsymbol{\theta}'\boldsymbol{Z}_{i,t-1} + \nu_k + \mu_t + \varepsilon_{i,t}, \quad (5)$$

Table 7: Difference-in-differences (DiD)

Panel A: A "Clean" estimate of DiD method			
	Pre-SOX period (1996 - 2001)	Post-SOX period (2002 - 2015)	Difference
Compliant	β_1	$\beta_1 + \beta_2$	β_2
	(Clean + $Bias^C$)	(Clean + $Bias^C$ + SOX)	(SOX)
Noncompliant	$\beta_1 + \beta_3$	$\beta_1 + \beta_2 + \beta_3 + \beta_4$	$\beta_2 + \beta_4$
	(Clean + $Bias^{NC}$)	(Clean + SOX)	(SOX - $Bias^{NC}$)
Difference	β_3	$\beta_3 + \beta_4$	β_4
	($Bias^{NC}$ - $Bias^C$)	(- $Bias^C$)	(- $Bias^{NC}$)
Clean = (Clean + SOX) - SOX = $\beta_1 + \beta_2 + \beta_3 + \beta_4 - \beta_2 = \beta_1 + \beta_3 + \beta_4$			
Panel B: Estimates of the "Clean" effects of independent board corruption			
Coefficient estimate		Results from base case	"Clean" estimate
Table 4: Column (1)		0.0069***	-0.0005
(*RD*)		(7.17)	(-0.40)
Table 4: Column (2)		-0.0996***	-0.1394***
(*Patent*)		(-5.51)	(-4.85)
Table 4: Column (3)		-0.2564***	-0.3872***
(*Citation*)		(-10.76)	(-7.10)
Table 4: Column (4)		-0.0100***	-0.0091**
(*IE*)		(-4.73)	(-2.07)
Table 5: Column (1)		0.1337***	0.2229**
(*ID_Compensation*)		(8.83)	(2.28)
Table 5: Column (2)		0.0118***	0.0772***
(*Cash_Compensation*)		(3.14)	(4.96)
Table 5: Column (3)		-0.0118***	-0.0772***
(*Equity_Compensation*)		(-3.14)	(-4.96)

Table 6: Column (1)		0.1361***	0.2261***
(*CEO_Compensation*)		(14.36)	(4.48)
Table 6: Column (2)		-0.1628***	-0.2157***
(*CEO_Vega*)		(-6.02)	(-3.24)

where, *Post-SOX* is a dummy variable that equal to one if the year is 2002 or after and zero otherwise. *Noncompliant* is a dummy variable that equal to one if the firm was not in compliance in 2001 and zero otherwise. *Post-SOX x Noncompliant x IDCC* is typical of DiD methodology but does not yield the "Clean" estimate. *IDCC* represents the independent director's corruption culture index. *Y* represents the dependent variables of corporate innovation and compensation of independent director and CEO.

Panel A of Table 7 present the sensitivity of corporate innovation and compensation of independent director and CEO to independent director's corruption culture in subsamples of (1) compliant with the pre-SOX period, (2) noncompliant with pre-SOX period, (3) compliant with post-SOX period, and (4) noncompliant with post-SOX period. Both coefficients of (1) and (2) are bias thus endogeneity issues. The coefficient of (3) not only bias but reflects the direct effect of SOX. We denote the bias for compliant firms ($Bias^{C}$) and noncompliant firms ($Bias^{NC}$).

The interesting subsample is noncompliant with *Post-SOX* period. This subsample contains the exogenous shock to the independent director's corruption culture. The sensitivity of coefficient (4) is influenced by the combination of the independent director's corruption culture and SOX effects on corporate innovation and compensation of independent director and CEO. The typical of DiD methodology does not report the clean estimate rather than the negative of $Bias^{NC}$. According to Coles et al. (2014), the clean estimate is the increase of exogenous in independent director's corruption culture that forced noncompliant firms to increase their independent director composition, is given by $\beta1 + \beta3 + \beta4$.

Panel B of Table 7 presents the result of a clean estimate for the total impact of independent director's corruption culture on corporate innovation and compensation of independent director and CEO. We compare the results from the base case (Table 4 – 6) and a clean estimate. The *RD* is insignificant and the other of the estimates that increase of exogenous shock in independent director's corruption culture have the same sign and significant in statistically. It supports the argument of Balsmeier et al. (2017) that independent boards are unrelated to the firm's R&D expenditures. But, the independent director spends more effort that increased the monitoring of explorative innovation. However, we find the result that consistent with agency theory. If the independent director's corruption culture is stronger than the role of monitoring will decrease and impact on decreased in innovation. We also support the argument that independent boards are likely to receive cash compensation and less likely to receive equity compensation when the role of monitoring is decreases and leads to ineffective corporate governance (Ryan & Wiggins, 2004). Also the positive relationship between excess compensation of independent director and CEO (Brick et al, 2006).

Overall, we argue that corporate innovation is decreases and compensation of independent director and CEO are increases when the independent director's corruption culture increases after SOX. Therefore, a stronger independent director's corruption culture leads to weakening corporate governance.

A Propensity-Score-Matching (PSM) Methodology

In this subsection, we sorted the level of independent director's corruption culture into quartiles and retain only the top quartile as the firms with highest independent director's corruption culture (treatment) and the bottom quartile as the firms with lowest independent director's corruption culture (control). Matching starts with a probit regression based on the treatment and control firms and uses all control variables. For the robustness results, we use 6 different matching methods: (1) Nearest neighbor (n=2), (2) Mahalanobis, (3) Gaussian Kernel, (4) Epanechnikov Kernel, and (5) Radius (1.0).

Table 8: Propensity-Score-Matching (PSM)				
Panel A1 Matching estimation: differences of *RD* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.0508	0.0432	0.0076***	(3.00)
Mahalanobis	0.0508	0.0416	0.0092***	(3.17)
Kernel Gaussian	0.0508	0.0437	0.0071***	(3.42)
Kernel Epanechnikov	0.0508	0.0443	0.0065***	(3.07)
Radius (0.1)	0.0508	0.0437	0.0071***	(3.43)
Panel A2 Matching estimation: differences of *Patent* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.6465	1.0456	-0.3991***	(-8.78)
Mahalanobis	0.6465	0.9656	-0.3191***	(-5.92)
Kernel Gaussian	0.6465	0.9831	-0.3366***	(-9.36)
Kernel Epanechnikov	0.6465	1.0100	-0.3635***	(-9.78)
Radius (0.1)	0.6465	0.9855	-0.3390***	(-9.40)
Panel A3 Matching estimation: differences of *Citation* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.5426	1.2543	-0.7117***	(-8.30)
Mahalanobis	0.5426	1.2702	-0.7276***	(-7.35)
Kernel Gaussian	0.5426	1.3006	-0.7580***	(-11.06)
Kernel Epanechnikov	0.5426	1.2779	-0.7353***	(-10.17)
Radius (0.1)	0.5426	1.3010	-0.7584***	(-11.02)
Panel A4: Matching estimation: differences of *IE* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.0466	0.0843	-0.0377***	(-7.09)
Mahalanobis	0.0466	0.0819	-0.0353***	(-5.46)
Kernel Gaussian	0.0466	0.0774	-0.0308***	(-7.34)
Kernel Epanechnikov	0.0466	0.0781	-0.0315***	(-7.25)
Radius (0.1)	0.0466	0.0775	-0.0309***	(-7.35)
Panel B1 Mathing estimation: differences of *ID_Compensation* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.0278	-0.2356	0.2634**	(2.33)
Mahalanobis	0.0278	-0.1189	0.1467*	(1.66)
Kernel Gaussian	0.0278	-0.2660	0.2938***	(4.01)
Kernel Epanechnikov	0.0278	-0.2418	0.2696***	(2.76)
Radius (0.1)	0.0278	-0.2762	0.3040***	(4.05)
Panel B2 Mathing estimation: differences of *Cash_Compensation* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.4167	0.3691	0.0476***	(3.94)
Mahalanobis	0.4167	0.3787	0.0380***	(2.85)

Kernel Gaussian	0.4167	0.3766	0.0401***	(4.09)
Kernel Epanechnikov	0.4167	0.3779	0.0388***	(3.74)
Radius (0.1)	0.4167	0.3770	0.0397***	(4.03)
Panel B3 Mathing estimation: differences of Equity_Compensation between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.5832	0.6308	-0.0476***	(-3.94)
Mahalanobis	0.5832	0.6212	-0.0380***	(-2.85)
Kernel Gaussian	0.5832	0.6233	-0.0401***	(-4.09)
Kernel Epanechnikov	0.5832	0.6220	-0.0388***	(-3.74)
Radius (0.1)	0.5832	0.6229	-0.0397***	(-4.03)
Panel C1 Mathing estimation: differences of *CEO_Compensation* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	0.0297	-0.2148	0.2445***	(3.67)
Mahalanobis	0.0297	-0.1695	0.1991***	(3.77)
Kernel Gaussian	0.0297	-0.1900	0.2197***	(5.13)
Kernel Epanechnikov	0.0297	-0.2033	0.2330***	(4.19)
Radius (0.1)	0.0297	-0.1852	0.2149***	(4.95)
Panel C2: Mathing estimation: differences of *CEO_Vega* between treatment and control firms				
Matching Methods	Treatment	Control	Difference	t-Statistic
Near neighbor (n=2)	2.8322	3.2710	-0.4388***	(-7.23)
Mahalanobis	2.8322	3.1828	-0.3506***	(-5.09)
Kernel Gaussian	2.8322	3.2217	-0.3895***	(-7.60)
Kernel Epanechnikov	2.8322	3.2544	-0.4222***	(-7.98)
Radius (0.1)	2.8322	3.2245	-0.3923***	(-7.63)

In all models, the *t*-values are computed on the heteroskedasticity-robust standard errors (White, 1980). Coefficients: ***, **, and * denote significant at 1%, 5%, and 10% levels, respectively.

Panel A1 – A4 of Table 8 presents the average corporate innovation of treatment and control groups and the difference in corporate innovation between groups. The differences are to have the same sign and significant. Those results mean that the independent director's corruption culture leads to weakening corporate governance, where less monitoring impact on overinvestment and lower quality of corporate innovation. Panel B1 – B3 and C1 – C2 present the same results. Those results mean that the independent director's corruption culture increases their excess compensation of independent director and CEO.

The Effect of External Corporate Governance Mechanisms

Institutional holdings concern on the composition of the board of directors and quality of corporate governance (Chung and Zhang, 2011). One of the roles of institutional holdings is used the public information disclosure in monitoring firm activities and enhance transparency in corporate governance (Armstrong et al, 2014). The agency problem from an independent director's corruption culture could be the reason why corporate innovation is decreases and compensation of independent director and CEO are increases. Also, that product market competition increases productivity growth, tight monitoring creditor, and also a substitute of block-holder control mechanisms (Köke & Renneboog, 2005). Product market competition is a substitution for internal corporate governance and effective to reduce the managerial inefficiency (Tian & Twite, 2011).

In this section, we explore a conditional test for how increased independent director's corruption culture causes decreased of corporate innovation and increased compensation of independent director and CEO.

Table 9: External corporate governance mechanisms

Panel A: Institutional ownership and product market similarity

Coefficient estimate	*IO < median*	*IO ≥ median*	*Similarity < median*	*Similarity ≥ median*
RD	0.0085	0.0031	0.0129***	0.0016
	(0.83)	(0.49)	(2.64)	(0.22)
Patent	-0.2506***	-0.0033	-0.1618**	0.0281
	(-5.89)	(-0.10)	(-2.35)	(0.52)
Citation	-0.6074***	-0.1304	-0.1289**	0.5468
	(-8.04)	(-1.54)	(-2.44)	(0.62)
IE	-0.0159***	-0.0019	-0.0174**	0.0420
	(-3.01)	(-0.24)	(-2.35)	(0.62)
ID_Compensation	0.2348*	0.1994	0.4954***	0.0165
	(1.70)	(1.57)	(3.75)	(0.14)
Cash_Compensation	0.0697***	0.0451	0.0561***	0.0364*
	(5.87)	(1.47)	(3.95)	(1.75)
Equity_Compensation	-0.0697***	-0.0451	-0.0561***	-0.0364*
	(-5.87)	(-1.47)	(-3.95)	(-1.75)
CEO_Compensation	0.1856**	0.1608*	0.2599***	0.1888**
	(2.51)	(1.82)	(3.67)	(2.05)
CEO_Vega	-0.3734***	-0.0914	-0.1822**	-0.0997*
	(-4.01)	(-0.92)	(-2.11)	(-1.79)

Panel B: Analyst coverage, takeover index, and HHI

Coefficient estimate	*AC < median*	*AC ≥ median*	*TO < median*	*TO ≥ median*	*HHI > median*	*HHI ≤ median*
RD	0.0008	-0.0013	-0.0032	0.0009	-0.0003	0.0012
	(0.39)	(-0.60)	(-1.38)	(0.71)	(-0.24)	(0.46)
Patent	-0.1150**	-0.0447	-0.0638	-0.1697***	-0.2043***	-0.1127***
	(-2.22)	(-0.35)	(-1.28)	(-3.88)	(-5.22)	(-2.62)
Citation	-0.2780***	-0.2304	-0.3730***	-0.3733***	-0.5348***	-0.3172***
	(-3.13)	(-1.30)	(-3.85)	(-2.61)	(-6.95)	(-4.01)
IE	0.0004	-0.0147**	-0.0043	-0.0070	-0.0141**	-0.0094
	(0.01)	(-2.11)	(-0.55)	(-1.09)	(-2.38)	(-1.34)
ID_Compensation	0.2291	0.2138	0.2372	0.2499**	0.2651**	0.2138*
	(1.57)	(1.20)	(1.50)	(2.05)	(2.28)	(1.71)
Cash_Compensation	0.0524***	0.0401**	0.0785***	0.0612***	0.0282	-0.0046
	(3.06)	(2.28)	(5.06)	(2.63)	(1.16)	(-0.17)

Equity_Compensation	-0.0524***	-0.0401**	-0.0785***	-0.0612***	-0.0282	0.0046
	(-3.06)	(-2.28)	(-5.06)	(-2.63)	(-1.16)	(0.17)
CEO_Compensation	0.1222	0.2750***	0.1416*	0.2255***	0.1485*	0.2864***
	(1.37)	(3.68)	(1.66)	(3.47)	(1.91)	(4.10)
CEO_Vega	-0.2924***	-0.1962**	-0.1866*	-0.1818**	-0.0750	-0.3269***
	(-2.71)	(-2.24)	(-1.73)	(-1.98)	(-0.84)	(-3.29)

In table 9 presents the conditional test of external corporate governance on the independent director’s corruption culture. We find that institutional ownership and product market similarity plays like weaker monitoring that decreases the corporate innovation and increases the compensation of independent director and CEO. Overall, the external corporate governance mechanisms weakly support our evidence.

Robustness Checks

<u>Controlling for CEO’s Corruption Culture and Alternative Measures of Boards’ Corruption Culture</u>

In this subsection, we replace the independent director’s corruption culture with the board of director’s corruption culture and control the CEO’s corruption culture in all models. In table 10, we find the independent director’s corruption culture and board of director’s corruption culture still remains. Further, independent director’s corruption culture leading weakens of corporate governance. Surprisingly, our findings confirm that the independent director’s corruption culture is stronger than the CEO’s corruption culture.

CONCLUSIONS

This study examines the effect of independent director’s corruption culture on corporate innovation and compensation of independent director and CEO. We use 12,332 firm-year observations during the period of 1996 – 2015 and measure the independent director’s corruption culture by determining their surname-ancestry with Corruption Perception Index (CPI).

We first document a strong negative relation between the independent director’s corruption culture and corporate innovation and a strong positive relation on the compensation of independent director and CEO. In addition, we conduct two methodologies to eliminate endogeneity issues, included (1) a difference-in-differences (DiD) with SOX as an exogenous shock, and (2) a propensity-score-matching (PSM) methodology. All of the results support our main expectation. Such evidence indicates (1) that independent director’s corruption culture impact on firm’s overinvestment by taking high risk projects without selection the project’s quality and transmitted to poor quality of corporate innovation, (2) that independent director’s corruption culture likely to less monitoring to applied friendly boards and use bargaining power to set own package of compensation.

We next explore a conditional test of external corporate governance. We find weakly evidence that poor corporate governance plays like weaker monitoring that decreases the corporate innovation and increases the compensation of independent director and CEO.

We continue to perform more robustness tests with (1) an alternative of corruption measurement, and (2) control for CEO corruption. All of the robustness tests are support our main expectation.

Table 10: Controlling for CEO's corruption culture and alternative measures of board's corruption culture

Panel A: Corporate innovation

	RD		*Patent*		*Citation*		*IE*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.1013***	0.1034***	1.1492***	1.1280***	2.9510***	3.0381***	0.2692***	0.2700***
	(8.14)	(8.20)	(4.95)	(4.80)	(8.40)	(8.57)	(8.12)	(7.99)
IDCC	0.0053***		-0.0874***		-0.2334***		-0.0092***	
	(5.34)		(-4.60)		(-9.41)		(-4.14)	
Board_Corruption		0.0057***		-0.0775***		-0.2672***		-0.0092***
		(4.93)		(-3.25)		(-8.59)		(-3.27)
CEO_Corruption	0.0020***	0.0015***	-0.0278***	-0.0217**	-0.0473***	-0.0208*	-0.0020**	-0.0012
	(4.63)	(3.18)	(-3.31)	(-2.29)	(-4.52)	(-1.75)	(-2.05)	(-1.08)
ROA	-0.0974***	-0.0961***	-0.1046	-0.0340	-0.4677**	-0.3805*	-0.0054	-0.0007
	(-9.38)	(-9.15)	(-0.69)	(-0.23)	(-2.22)	(-1.80)	(-0.30)	(-0.04)
Leverage	-0.0002***	-0.0002***	0.0129***	0.0132***	0.0138***	0.0141***	-0.0007***	-0.0006***
	(-3.84)	(-4.00)	(7.48)	(7.45)	(6.22)	(6.22)	(-4.64)	(-4.23)
BHR	-0.0037***	-0.0044***	0.0467*	0.0460	0.0304	0.0351	0.0039	0.0042
	(-2.66)	(-3.03)	(1.69)	(1.62)	(0.79)	(0.89)	(1.23)	(1.29)
Firm Size	-0.0019***	-0.0019***	0.1077***	0.1074***	0.1063***	0.1074***	-0.0045***	-0.0042***
	(-3.49)	(-3.48)	(9.05)	(8.81)	(7.06)	(7.02)	(-3.98)	(-3.72)
Dividend	0.0013**	0.0015**	0.0764***	0.0732***	0.0623**	0.0586**	-0.0033**	-0.0036***
	(2.13)	(2.44)	(3.62)	(3.40)	(2.49)	(2.29)	(-2.50)	(-2.69)
E-Index	-0.0015***	-0.0015***	-0.1129***	-0.1126***	-0.1553***	-0.1539***	-0.0082***	-0.0083***
	(-3.97)	(-3.89)	(-14.34)	(-14.03)	(-15.44)	(-15.06)	(-9.33)	(-9.26)
IO	-0.0110***	-0.0105***	-0.3853***	-0.3931***	-0.5571***	-0.5825***	-0.0213***	-0.0224***
	(-4.25)	(-3.98)	(-8.26)	(-8.29)	(-8.78)	(-9.03)	(-3.60)	(-3.72)
HHI	-0.0419*	-0.0413*	-4.2206***	-4.3125***	-6.7157***	-6.8273***	-0.6536***	-0.6617***
	(-1.90)	(-1.85)	(-6.51)	(-6.58)	(-6.03)	(-6.10)	(-5.92)	(-5.96)
Board Size	-0.0212***	-0.0218***	-0.0492	-0.0525	-0.1246*	-0.1585**	-0.0129**	-0.0148**
	(-8.46)	(-8.33)	(-0.96)	(-0.99)	(-1.78)	(-2.18)	(-2.25)	(-2.50)
Board Independency	0.0077*	0.0080*	0.6262***	0.5989***	0.4194***	0.3644***	0.0612***	0.0604***
	(1.89)	(1.92)	(7.39)	(6.86)	(3.40)	(2.87)	(6.16)	(5.86)
Director Age	0.0012	0.0011	-0.1028***	-0.0941***	-0.0428	-0.0332	-0.0076*	-0.0072*

	(1.28)	(1.21)	(-3.73)	(-3.39)	(-1.59)	(-1.21)	(-1.95)	(-1.85)
Female Director	-0.0057***	-0.0056***	0.0100	0.0123	0.0060	0.0132	-0.0068*	-0.0066*
	(-4.13)	(-3.97)	(0.35)	(0.43)	(0.16)	(0.35)	(-1.91)	(-1.81)
Director Tenure	-0.0017	-0.0023**	0.0109	0.0084	-0.0539	-0.0562	-0.0010	-0.0005
	(-1.59)	(-2.10)	(0.47)	(0.35)	(-1.45)	(-1.47)	(-0.31)	(-0.15)
Control for:								
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	11,405	11,084	11,405	11,084	11,405	11,084	11,405	11,084
Adj. R^2	0.4664	0.4652	0.2708	0.2701	0.2741	0.2735	0.1842	0.1843

Panel B: Independent director's compensation						
	ID_Compensation		*Cash_Compensation*		*Equity_Compensation*	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.3436**	1.0567*	0.2471***	0.2131***	0.7529***	0.7869***
	(2.11)	(1.79)	(3.20)	(2.75)	(9.74)	(10.16)
IDCC	0.1446***		0.0104***		-0.0104***	
	(9.33)		(2.73)		(-2.73)	
Board_Corruption		0.1548***		0.0168***		-0.0168***
		(8.79)		(3.73)		(-3.73)
CEO_Corruption	0.0055	-0.0091	0.0025*	0.0005	-0.0025*	-0.0005
	(0.99)	(-1.48)	(1.67)	(0.31)	(-1.67)	(-0.31)
ROA	0.3276***	0.4047***	-0.1011***	-0.1051***	0.1011***	0.1051***
	(3.02)	(3.66)	(-3.55)	(-3.62)	(3.55)	(3.62)
Leverage	-0.0001	0.0007	-0.0003	-0.0004	0.0003	0.0004
	(-0.10)	(0.90)	(-1.32)	(-1.53)	(1.32)	(1.53)
BHR	0.1602***	0.1580***	-0.0248***	-0.0238***	0.0248***	0.0238***
	(6.23)	(6.13)	(-4.66)	(-4.44)	(4.66)	(4.44)
Firm Size	-0.0250***	-0.0188***	-0.0025	-0.0034*	0.0025	0.0034*
	(-3.62)	(-2.76)	(-1.28)	(-1.67)	(1.28)	(1.67)
Dividend	-0.0555***	-0.0538***	0.0061**	0.0049*	-0.0061**	-0.0049*
	(-7.16)	(-6.90)	(2.40)	(1.92)	(-2.40)	(-1.92)
E-Index	0.0395***	0.0422***	0.0128***	0.0123***	-0.0128***	-0.0123***
	(5.32)	(5.63)	(8.45)	(8.03)	(-8.45)	(-8.03)
IO	-0.3659***	-0.3416***	0.0849***	0.0813***	-0.0849***	-0.0813***
	(-10.25)	(-9.55)	(8.17)	(7.78)	(-8.17)	(-7.78)
HHI	-8.0718**	-6.1007*	0.4145**	0.4356**	-0.4145**	-0.4356**
	(-2.23)	(-1.83)	(2.22)	(2.32)	(-2.22)	(-2.32)
Board Size	-0.0471	-0.0905**	0.0289**	0.0415***	-0.0289**	-0.0415***
	(-1.01)	(-2.28)	(2.34)	(3.50)	(-2.34)	(-3.50)
Board Independency	0.3235***	0.3486***	-0.1392***	-0.1375***	0.1392***	0.1375***
	(4.10)	(4.45)	(-6.73)	(-6.58)	(6.73)	(6.58)
Director Age	0.0145	0.0091	-0.0360***	-0.0362***	0.0360***	0.0362***
	(0.51)	(0.31)	(-4.20)	(-4.16)	(4.20)	(4.16)
Female Director	0.0164	0.0195	0.0075	0.0074	-0.0075	-0.0074
	(0.81)	(0.95)	(1.28)	(1.25)	(-1.28)	(-1.25)

Director Tenure	-0.0112	-0.0265	0.0287***	0.0308***	-0.0287***	-0.0308***
	(-0.45)	(-1.04)	(4.58)	(4.82)	(-4.58)	(-4.82)
Control for:						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	10,616	10,320	10,719	10,420	10,719	10,420
Adj. R^2	0.4641	0.4672	0.1352	0.1384	0.1352	0.1384

Panel C: CEO's compensation				
	CEO_Compensation		*CEO_Vega*	
	(1)	(2)	(3)	(4)
Constant	-0.5417**	-1.0700**	1.6587***	1.6822***
	(-2.43)	(-2.42)	(3.47)	(3.50)
IDCC	0.1359***		-0.1501***	
	(14.05)		(-5.33)	
Board_Corruption		0.1303***		-0.2084***
		(10.04)		(-6.11)
CEO_Corruption	0.0046	-0.0067	-0.0278**	-0.0023
	(1.20)	(-1.59)	(-2.42)	(-0.18)
ROA	0.1294*	0.1237	1.3085***	1.2913***
	(1.75)	(1.64)	(6.46)	(6.25)
Leverage	-0.0014*	-0.0013	0.0324***	0.0328***
	(-1.69)	(-1.63)	(13.09)	(12.96)
BHR	0.1611***	0.1667***	0.1267***	0.1309***
	(10.30)	(9.73)	(3.31)	(3.36)
Firm Size	-0.0071	-0.0080	0.2718***	0.2670***
	(-1.33)	(-1.52)	(16.66)	(16.02)
Dividend	-0.0120*	-0.0156**	0.1972***	0.1872***
	(-1.69)	(-2.21)	(7.44)	(6.95)
E-Index	0.0093*	0.0051	-0.0112	-0.0090
	(1.85)	(0.98)	(-0.86)	(-0.69)
IO	-0.2432***	-0.2128***	-1.1071***	-1.1151***
	(-10.54)	(-9.05)	(-15.72)	(-15.62)
HHI	1.8060***	-0.3833	-6.1414***	-6.1258***
	(4.96)	(-0.15)	(-5.27)	(-5.26)
Board Size	-0.1111***	-0.0803***	0.5786***	0.6239***
	(-4.08)	(-2.89)	(7.08)	(7.40)
Board Independency	0.4546***	0.4794***	0.1736	0.1799
	(9.67)	(9.91)	(1.35)	(1.36)
Director Age	-0.0104	-0.0132	-0.0440	-0.0493
	(-0.57)	(-0.70)	(-0.85)	(-0.93)
Female Director	0.0485***	0.0497***	0.1477***	0.1533***
	(3.76)	(3.76)	(3.65)	(3.73)
Director Tenure	-0.0072	-0.0137	0.0213	0.0188
	(-0.44)	(-0.84)	(0.65)	(0.56)
Control for:				

Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Obs.	10,435	10,434	8,861	8,638
Adj. R^2	0.2348	0.2528	0.2438	0.2449

APPENDIX

Variable	Definition	Data source
Panel A: Independent variables		
IDCC	The average corruption values for all independent directors by their surname-ancestry country from Transparency International CPI	Risk Metrics
CEO_Corruption	The average corruption values for CEOs by their surname-ancestry country from Transparency International CPI	Risk Metrics
Board_Corruption	The average corruption values for all boards by their surname-ancestry country from Transparency International CPI	Risk Metrics
Panel B: Dependent variables		
Patent	The natural logarithm of the weight-adjusted patent counts	PATSTAT and USPTO
Citation	The natural logarithm of the citation lag-adjusted citations per patent	PATSTAT and USPTO
IE	The natural logarithm of the patents or citations scaled by R&D expenditures	PATSTAT and USPTO
ID_Compensation	The estimated value by the residual from independent director's total compensation regression	ExecuComp
Cash_Compensation	Total cash scaled by total compensation	ExecuComp
Equity_Compensation	Total stock plus options scaled by total compensation	ExecuComp
CEO_Compensation	The estimated value by the residual from CEO compensation regression	ExecuComp
CEO_Vega	The natural logarithm of one plus stock return volatility sensitivity change in CEO wealth	Core & Guay, (2002)
Panel C: Firm characteristics		
ROA	Total net income scaled by total assets	Compustat
Leverage	Current liabilities plus long term debt scaled by total assets	Compustat
BHR	Buy and hold one year return	CRSP
Firm Size	The natural logarithm of total assets	Compustat
Dividend	Total dividend payment scaled by total assets	Compustat
Panel D: Corporate governance		
E-Index	A corporate governance quality based on six provisions following Bebchuk et al. (2009)	Risk Metrics
HHI	The market concentration of an industry	Compustat
IO	The percentage of shares held by institutional ownership	Thomson Reuters 13F
Board Size	The natural of logarithm of total number of directors on board	Risk Metrics

Board Independency	The percentage of independent board scaled by board size	Risk Metrics
Panel E: Independent director		
Director Age	The natural of logarithm of average independent board's age	Risk Metrics
Female Director	Dummy variable equal to one if board has a female director and otherwise is zero	Risk Metrics
Director Tenure	The natural of logarithm of average independent board's tenure	Risk Metrics
Panel F: Product market competition		
Similarity	A text-based analysis of product similarity descriptions	Hoberg and Phillips (2010, 2016)

REFERENCES

Adams, R. B., & Ferreira, D. (2007). A theory of friendly boards. *Journal of Finance*, *62*(1), 217-250.

Adams, R. B., Hermalin, B. E., & Weisbach, M. S. (2010). The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature, 48* (1), 58-107.

Aguir, I., Burns, N., Mansi, S. A., & Wald, J. K. (2014). Liability protection, director compensation, and incentives. *Journal of Financial Intermediation*, *23*(4), 570-589.

Ahn, S., & Shrestha, K. (2013). The differential effects of classified boards on firm value. *Journal of Banking &Finance*, *37*(11), 3993-4013.

Amore, M. D., Schneider, C., & Žaldokas, A. (2013). Credit supply and corporate innovation. *Journal of Financial Economics*, *109*(3), 835-855.

Armstrong, C. S., Core, J. E., & Guay, W. R. (2014). Do independent directors cause improvements in firm transparency?. *Journal of Financial Economics*, *113*(3), 383-403.

Ayyagari, M., Demirgüç-Kunt, A., & Maksimovic, V. (2014). Bribe payments and innovation in developing countries: Are innovating firms disproportionately affected?. *Journal of Financial and Quantitative Analysis*, *49*(1), 51-75.

Balsmeier, B., Fleming, L., & Manso, G. (2017). Independent boards and innovation. *Journal of Financial Economics*, *123*(3), 536-557.

Baranchuk, N., Kieschnick, R., & Moussawi, R. (2014). Motivating innovation in newly public firms. *Journal of Financial Economics*, *111*(3), 578-588.

Bargeron, L. L., Lehn, K. M., & Zutter, C. J. (2010). Sarbanes-Oxley and corporate risk-taking. *Journal of Accounting and Economics*, *49*(1-2), 34-52.

Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *Accounting Review*, 443-465.

Bebchuk, L., Cohen, A., & Ferrell, A. (2008). What matters in corporate governance?. *Review of Financial Studies*, *22*(2), 783-827.

Bebchuk, L. A., & Fried, J. M. (2003). Executive compensation as an agency problem. *Journal of Economic Perspectives*, *17*(3), 71-92.

Bebchuk, L. A., Fried, J. M., & Walker, D. I. (2002). Managerial Power and Rent Extraction in the Design of Executive Compensation. *The University of Chicago Law Review*, *69*(3), 751-846.

Bliss, R. T., & Rosen, R. J. (2001). CEO compensation and bank mergers. *Journal of Financial Economics*, *61*(1), 107-138.

Boyd, B. K. (1994). Board control and CEO compensation. *Strategic Management Journal, 15* (5), 335-344.
Brick, I. E., Palmon, O., & Wald, J. K. (2006). CEO compensation, director compensation, and firm performance: Evidence of cronyism?. *Journal of Corporate Finance*, *12*(3), 403-423.
Brochet, F., & Srinivasan, S. (2014). Accountability of independent directors: Evidence from firms subject to securities litigation. *Journal of Financial Economics*, *111*(2), 430-449.
Chang, X., Fu, K., Low, A., & Zhang, W. (2015). Non-executive employee stock options and corporate innovation. *Journal of Financial Economics*, *115*(1), 168-188.
Chava, S., Oettl, A., Subramanian, A., & Subramanian, K. V. (2013). Banking deregulation and innovation. *Journal of Financial Economics*, *109*(3), 759-774.
Cheng, S. (2004). R&D expenditures and CEO compensation. *Accounting Review*, *79*(2), 305-328.
Chhaochharia, V., & Grinstein, Y. (2009). CEO compensation and board structure. *Journal of Finance*, *64*(1), 231-261.
Chung, K. H., & Zhang, H. (2011). Corporate governance and institutional ownership. *Journal of Financial and Quantitative Analysis*, *46*(1), 247-273.
Ciftci, M., & Cready, W. M. (2011). Scale effects of R&D as reflected in earnings and returns. *Journal of Accounting and Economics*, *52*(1), 62-80.
Coates, I. V., & John, C. (2007). The goals and promise of the Sarbanes-Oxley Act. *Journal of Economic Perspectives*, *21*(1), 91-116.
Coles, J. L., Daniel, N. D., & Naveen, L. (2014). Co-opted boards. *Review of Financial Studies, 27*(6), 1751-1796.
Coles, J. L., Daniel, N. D., & Naveen, L. (2008). Boards: Does one size fit all?. *Journal of Financial Economics*, *87*(2), 329-356.
Coles, J. L., Daniel, N. D., & Naveen, L. (2006). Managerial incentives and risk-taking. *Journal of Financial Economics*, *79*(2), 431-468.
Core, J., & Guay, W. (2002). Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research*, *40*(3), 613-630.
Cornaggia, J., Mao, Y., Tian, X., & Wolfe, B. (2015). Does banking competition affect innovation? *Journal of Financial Economics*, *115*(1), 189-209.
Cotter, J. F., Shivdasani, A., & Zenner, M. (1997). Do independent directors enhance target shareholder wealth during tender offers?. *Journal of Financial Economics*, *43*(2), 195-218.
Dah, M. A., & Frye, M. B. (2017). Is board compensation excessive?. *Journal of Corporate Finance*, *45*, 566-585.
Dah, M. A., Frye, M. B., & Hurst, M. (2014). Board changes and CEO turnover: The unanticipated effects of the Sarbanes–Oxley Act. *Journal of Banking & Finance*, *41*, 97-108.
DeBacker, J., Heim, B. T., & Tran, A. (2015). Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics, 117* (1), 122-138.
Ederer, F., & Manso, G. (2013). Is pay for performance detrimental to innovation?. *Management Science*, *59*(7), 1496-1513.
Falato, A., Kadyrzhanova, D., & Lel, U. (2014). Distracted directors: Does board busyness hurt shareholder value?. *Journal of Financial Economics*, *113*(3), 404-426.
Faleye, O. (2007). Classified boards, firm value, and managerial entrenchment. *Journal of Financial Economics*, *83*(2), 501-529.
Faleye, O., Hoitash, R., & Hoitash, U. (2011). The costs of intense board monitoring. *Journal of Financial Economics*, *101*(1), 160-181.
Fang, V. W., Tian, X., & Tice, S. (2014). Does stock liquidity enhance or impede firm innovation? *Journal of Finance*, *69*(5), 2085-2125.

Fedaseyeu, V., Linck, J. S., & Wagner, H. F. (2018). Do qualifications matter? New evidence on board functions and director compensation. *Journal of Corporate Finance*, *48*, 816-839.
Fogel, K., Ma, L., & Morck, R. (2014). Powerful Independent Directors. Working Paper No. 19809). National Bureau of Economic Research.
Fracassi, C., & Tate, G. (2012). External networking and internal firm governance. *Journal of Finance*, *67*(1), 153-194.
Grinstein, Y., & Hribar, P. (2004). CEO compensation and incentives: Evidence from M&A bonuses. *Journal of Financial Economics*, *73*(1), 119-143.
He, J. J., & Tian, X. (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, *109*(3), 856-878.
Hermalin, B. E., & Weisbach, M. S. (1998). Endogenously chosen boards of directors and their monitoring of the CEO. *American Economic Review*, *88*(1), 96-118.
Hirshleifer, D., Hsu, P. H., & Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, *107*(3), 632-654.
Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423-1465.
Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, *23*(10), 3773-3811.
Holthausen, R. W., Larcker, D. F., & Sloan, R. G. (1995). Business unit innovation and the structure of executive compensation. *Journal of Accounting and Economics*, *19*(2-3), 279-313.
Humphery-Jenner, M., Lisic, L. L., Nanda, V., & Silveri, S. D. (2016). Executive overconfidence and compensation structure. *Journal of Financial Economics*, *119*(3), 533-558.
Hwang, B. H., & Kim, S. (2009). It pays to have friends. *Journal of Financial Economics*, *93*(1), 138-158.
Ishii, J., & Xuan, Y. (2014). Acquirer-target social ties and merger outcomes. *Journal of Financial Economics*, *112*(3), 344-363.
Kim, K., Mauldin, E., & Patro, S. (2014). Outside directors and board advising and monitoring performance. *Journal of Accounting and Economics*, *57*(2-3), 110-131.
Klein, A. (2002). Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics*, *33*(3), 375-400.
Knyazeva, A., Knyazeva, D., & Masulis, R. W. (2013). The supply of corporate directors and board independence. *Review of Financial Studies*, *26*(6), 1561-1605.
Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, *132*(2), 665-712.
Köke, J., & Renneboog, L. (2005). Do corporate control and product market competition lead to stronger productivity growth? Evidence from market-oriented and blockholder-based governance regimes. *Journal of Law and Economics*, *48*(2), 475-516.
Kumar, P., & Sivaramakrishnan, K. (2008). Who monitors the monitor? The effect of board independence on executive compensation and firm value. *Review of Financial Studies*, *21*(3), 1371-1401.
Lerner, J., & Wulf, J. (2007). Innovation and incentives: Evidence from corporate R&D. *Review of Economics and Statistics*, *89*(4), 634-644.
Linck, J. S., Netter, J. M., & Yang, T. (2008). The effects and unintended consequences of the Sarbanes-Oxley Act on the supply and demand for directors. *Review of Financial Studies*, *22*(8), 3287-3328.
Linn, S. C., & Park, D. (2005). Outside director compensation policy and the investment opportunity set. *Journal of Corporate Finance*, *11*(4), 680-715.
Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Economics, 122*(2), 307-327.

Lu, J., & Wang, W. (2018). Managerial conservatism, board independence and corporate innovation. *Journal of Corporate Finance*, *48*, 1-16.
Mace, M. L. (1971). *Directors: Myth and reality*. Boston: Harvard Business School Press.
Manso, G. (2011). Motivating innovation. *Journal of Finance*, *66*(5), 1823-1860.
Masulis, R. W., & Mobbs, S. (2016). Independent Director Reputation Incentives: CEO Compensation Contracting and Financial Accounting Quality. Working Paper, University of New South Wales.
Masulis, R. W., & Mobbs, S. (2014). Independent director incentives: Where do talented directors spend their limited time and energy?. *Journal of Financial Economics*, *111*(2), 406-429.
Masulis, R. W., Wang, C., & Xie, F. (2009). Agency problems at dual-class companies. *Journal of Finance*, *64*(4), 1697-1727.
Nguyen, B. D., & Nielsen, K. M. (2010). The value of independent directors: Evidence from sudden deaths. *Journal of Financial Economics*, *98*(3), 550-567.
Nguyen, N. Q. (2014). On the compensation and activity of corporate boards. *Journal of Corporate Finance*, *29*, 1-19.
O'Connor, M., & Rafferty, M. (2012). Corporate governance and innovation. *Journal of Financial and Quantitative Analysis*, *47*(2), 397-413.
Otto, C. A. (2014). CEO optimism and incentive compensation. *Journal of Financial Economics, 114*(2), 366-404.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
Rosenstein, S., & Wyatt, J. G. (1990). Outside directors, board independence, and shareholder wealth. *Journal of Financial Economics*, *26*(2), 175-191.
Ryan Jr, H. E., & Wiggins III, R. A. (2004). Who is in whose pocket? Director compensation, board independence, and barriers to effective monitoring. *Journal of Financial Economics*, *73*(3), 497-524.
Schmidt, B. (2015). Costs and benefits of friendly boards during mergers and acquisitions. *Journal of Financial Economics*, *117*(2), 424-447.
Shivdasani, A., & Yermack, D. (1999). CEO involvement in the selection of new board members: An empirical analysis. *Journal of Finance*, *54*(5), 1829-1853.
Tian, G. Y., & Twite, G. (2011). Corporate governance, external market discipline and firm productivity. *Journal of Corporate Finance*, *17*(3), 403-417.
Weisbach, M. S. (1988). Outside directors and CEO turnover. *Journal of Financial Economics, 20*, 431-460.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838.
Wintoki, M. B. (2007). Corporate boards and regulation: The effect of the Sarbanes–Oxley Act and the exchange listing requirements on firm value. *Journal of Corporate Finance*, *13*(2-3), 229-250.
Zhong, R. I. (2018). Transparency and firm innovation. *Journal of Accounting and Economics, 66*(1), 67-93.

DECISION SCIENCES INSTITUTE

The Strategic Impact of Human Resource Information Systems and the Moderating Role of Organizational Culture: A Preliminary Empirical Study

Joonghee Lee
Auburn University
Email: jzl0165@auburn.edu

ABSTRACT

Human capital is important, and the strategic role of HRM has been emphasized over recent years. There is limited empirical research regarding the application of Human Resource Information Systems for strategic HRM. This study examines the complementarity among human resource information systems (HRIS) and organizational culture by assessing their impact on the strategic involvement of HR. Our findings suggest that organizational culture positively moderates the relationship between HRIS and strategic involvement of HR. We also find that HRIS contributes to the strategic role of HR.

KEYWORDS: Information systems, Strategic human resource management, Organizational culture.

INTRODUCTION

Demonstrating the value of information technology (IT) is a central issue in the information systems (IS) discipline, and a significant body of research has examined the organizational performance impacts of IT (e.g. Chae et al., 2014; Mithas & Rust, 2016; Mithas et al., 2017; Nevo & Wade, 2010; Wade & Hulland, 2004; Weill, 1992). IT in literature is frequently used as aggregate variables (Devaraj & Kohli, 2003), such as IT investments (Mithas et al., 2012) and IT portfolio (Oh & Pinsonneault, 2007). These variables represent to what extent a firm allocates its budget into IT resources, and different types of IT resource. However, scholars in the IS field have argued that we need to examine the IT-performance link at a disaggregated level, such as at the level of specific information systems (Aral & Weill, 2007; DeSanctis,1986; Mithas et al., 2006). This study investigates the strategic benefits of human resource information system (HRIS). By examining the link between specialized information system within a traditional functional area of an organization and the business value, this study extends our understanding regarding the IT-related value creation at an IT system level.

The human resource (HR) function is ubiquitous and human capital is a critical for organizational performance (Becker & Gerhart, 1996; Liu et al., 2017; Sun et al., 2007). To achieve administrative and strategic benefits, firms have increasingly adopted human resource information systems (HRIS) (Bondarouk et al., 2017) that allow HR activities and processes to occur electronically. While many prior studies suggested positive potential of using technologies to increase the effectiveness of the HR function, researchers increasingly call for more empirical studies regarding consequences of HRIS (Bondarouk et al., 2017). Moreover, a critical trend over recent years in HR management (HRM) is the emphasis on the strategic roles of HRM (SHRM) rather than its administrative role. In other words, HR practitioners are under pressure to be strategic partners and to contribute to a firm's competitive advantage (Paauwe, 2004). HRIS can help HR departments play that strategic role. Literature has broadly suggested two main benefits of HRIS: the improvement of efficiency and reduction of costs associated with HR activities involving day-to-day transactions (e.g. payroll and hiring), and facilitation of a more strategic role for the HR function (Parry, 2011). A large proportion of administrative roles of HRM are now delivered through information technology (IT), which enables HR people to focus more on strategic roles. Thus, we will examine two questions in this study:

(1) *Does HRIS enable HR functions to make more contributions at a strategic level?*
(2) *Under what conditions does HRIS contribute to increase the strategic values of HR?*

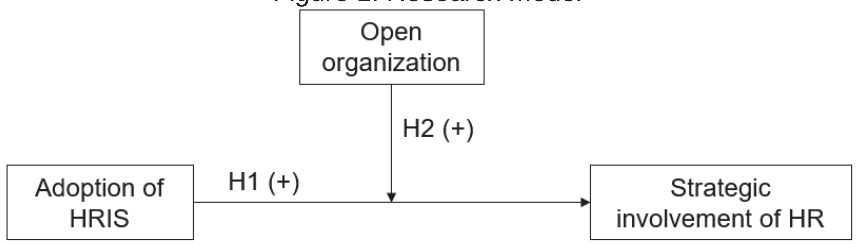
There is a substantial body of research on HRIS with different foci, such as implementation (Lippert & Michael Swiercz, 2005), administrative efficiency (Bondarouk et al., 2009), employee perceptions towards HRIS (Wickramasinghe, 2010), and job satisfaction and turnover intention (Maier et al., 2013). However, little literature exists regarding HRIS. In addition, the link between HRIS and its strategic value is not well established in empirical research. Although scholars have argued the significant relationship between HRIS and its strategic outcomes at a theoretical level (Lippert & Swiercz, 2005; Martin & Reddington, 2010), empirical evidence is inconsistent. Some studies show that HRIS is related to SHRM (Hussain et al., 2007; Marler & Parry, 2016; Parry, 2011), but others did not find evidence showing that HRIS results in direct strategic benefits (Bondarouk & Ruël, 2013; Dery et al., 2013; Parry & Tyson, 2011). Thus, the debate on HRIS and its strategic value is inconclusive (Bondarouk & Ruël, 2013), indicating a need to produce more empirical studies on HRIS and SHRM outcomes (Marler & Fisher, 2013). Moreover, the resource-based view suggests that IT resources may be more effective when they are paired with other resources (Mithas et al., 2011; Melville et al., 2004; Nevo & Wade, 2010; Wade & Hulland, 2004; Bresnahan et al., 2002), in that the organizational influences of IT can be affected by other organizational factors. In addition, the ‘sociotechnical framework’ suggests that technologies should be merged with human dimensions (e.g. organizational culture and leadership) for successful technological performance (Miller & Rice, 1967). In particular, Leidner and Kayworth (2006) suggested that organizational culture is significantly associated with IT value creation. This study investigates the boundary conditions where the link between HRIS and SHRM can vary by including a cultural variable. In particular, open organization, which refers to the culture of trusting and open relationships with minimal bureaucracy, free communications within and across business units, is critical cultural factor linked with IT performance (Powell & Dent-Micallef, 1997; Zuboff, 1998).

This study contributes to the literature in three ways. First, it deepens our understanding of the business value of IT by investigating a specialized information system within a traditional functional area of an organization. Second, it begins to resolve contradictory results in previous research of the relationship between HRIS and SHRM by proposing a moderation model. The inconsistency may be explained by exclusion of other organizational factors or limitation of data sets. Previous studies focused on the direct link between HRIS and SHRM have relied on cross-sectional data from surveys (Bondarouk & Ruël, 2013; Hussain et al., 2007; Parry, 2011) and case studies (Parry & Tyson, 2011; Bondarouk & Ruël, 2009). This study includes organizational factors and use panel data from firms in South Korea. Using longitudinal data and panel models mitigates the problem of endogeneity by reducing omitted variable bias, which results in more precise estimates of coefficients in our model. Third, this study contributes to our understanding of HRIS strategic values by testing the suggested model in a different context, namely South Korean organizations. Previous studies have been conducted in the U.S. and in Europe (e.g. Bondarouk et al., 2009; Bondarouk & Ruël, 2013; Hussain et al., 2007).

RESEARCH MODEL AND HYPOTHESIS

This study investigates the effects of HRIS on the strategic roles of HR functions in organizations – as well as the moderating effect of organizational culture. Specifically, the proposed model is shown in Figure 1.

Figure 1: Research model



The Strategic Benefits of HRIS

Literature in HRM and IS suggests that HRIS reduce administrative burdens of HR staff and allow them to focus on more strategic support for their organizations (Bondarouk & Ruël, 2013; Ngai & Wat, 2006). Specifically, HR departments can automate routine HR tasks and replace 'filing cabinets' by adopting HRIS (Parry et al., 2007), expend more effort on the improvement of HR strategies and policies for the strategic management of employees (Martin et al., 2008; Ruël et al., 2004; Shrivastava & Shaw, 2003) and support strategic decision-making (Bondarouk & Ruël, 2009; Hussain et al., 2007; Parry & Tyson, 2011). Thus, HRIS help transform HR professionals from administrative workers to strategic partners (Bell et al., 2006; Burbach & Dundon, 2005; Haines & Lafleur, 2008; Voermans & Van Veldhoven, 2007). The adoption of HRIS may allow the HR functions to increase its value in strategic levels.

Hypothesis 1: The adoption of HRIS has a positive association with the strategic involvement of HR functions in organizations.

The Moderating Role of Open Organization

According to the resource-based view (RVB), resources are solely responsible for organizational performance (Wade & Hulland, 2004). In particular, IT resources can be more effective for improving organizational performance in conjunction with other firm resources, such as structure, culture, and skills (Chae et al., 2014). Thus, many scholars have viewed the effects of IT as both contingent and complementary (Holland et al., 1992; Kettinger et al., 1994; Kettinger et al., 1995; Powell & Dent-Micallef, 1997). Complementarity refers the effect that one resource in organizations may have on another. Resources may influence performance when combined with another resource (Teece, 1986); the effect of HRIS likely depends on other constructs. Bondarouk and Ruël (2013) suggested HRIS alone is not sufficient to enable HR functions to contribute strategically. In a study of a large federal government organization, they found that HRIS do not automatically result in direct benefits, and so do not necessarily increase the HR functions' capabilities.

In this section, we discuss a boundary condition by which the relationship between HRIS and the strategic role of HR functions are affected. The perspective that firms should merge technology with cultural dimensions is rooted in the 'sociotechnical framework'. This framework suggests that maximized technological performance requires an optimization of social and technological subsystems in an organization (Miller & Rice, 1967). In addition, IS researchers have considered culture to be a critical variable when explaining the effects of IT. Leidner and Kayworth (2006) suggests that organizational culture is related to information systems development, IT adoption, diffusion, use, outcomes, and more. Also, several IS literature suggest that corporate culture may influence the relationship between IT and organizational performance (Melville et al., 2004; Wade & Hulland, 2004; Piccoli & Ives, 2005). An important cultural variable linked with IT performance is openness of the organization (Powell & Dent-Micallef, 1997). The benefits of IT may be negated by structural, relational, and communicational constrictions, such as limited access to information, top-down communication, and autocratic command and control (Zuboff, 1988). A critical benefit of information systems is to increase the accessibility of information. So, firms can increase the value by allowing employees access to information traditionally controlled by upper management, and avoiding traditional hierarchies, and top-down communications (Powell & Dent-Micallf, 1997).

Hypothesis 2: Organizational openness positively moderates the relationship between HRIS and the strategic involvement of HR functions.

METHOD

Sample and Procedure

The data come from the Human Capital Corporate Panel (HCCP) Survey from the Korean Research Institute for Vocational Education and Training (KRIVET, 2015). The survey data of the sample companies are combined with the corporate financial data held by the Korea Investors Service (KIS) using a corresponding firm code (KIS, 2015).

KRIVET has conducted the HCCP survey every two years since 2005; the survey includes a battery of questions regarding HR practices. We focused on the survey data and financial data between 2011 and 2015 for two reasons. First, the HCCP survey sample was changed in 2009. The 9th Korean standard industrial classification was placed into effect on February 1, 2008. Second, to eliminate the effects of the Great Recession, we focused on the data after that time.

For the survey process, KRIVET contacted HR managers at targeted firms, and administered the HCCP survey using on-site interviews. The referent for the survey items was the firm, and HR managers responded to the items. The population for HCCP sample extraction consisted of companies listed in the company overview information of the KIS database with at least 100 employees across six industries: (1) Manufacturing, (2) Finance, banking and insurance, (3) Publishing/video/broadcasting and telecommunications, and information service, (4) Specialist/scientific and technical service (5), education service, and (6) Arts/sports and leisure service. Between 2011 and 2015, companies for the survey sample numbered between 467 and 500.

Variable Operationalization and Descriptive Statistics

Strategic involvement of HR (SHR). This dependent variable consists of responses to four survey questions. The first three questions were asked to HR managers of the firms. The first indicates how much the HR function reflects the company’s mid and long-term strategies with four possible responses between the lowest value of 1 (rarely reflected) and the highest value of 4 (much reflected). The second and third questions indicate whether the firm engages in HR planning each year (1: yes, 0: no), and whether that planning reflects management strategy (1:

yes, 0: no). The fourth question was asked to employees in a firm to indicate how much the HR department contributes to establishing business management strategies (5: from the outset to 0: no contribution). Responses to these items were summed to create a scale from 2 (rarely strategically involved) to 11 (strongly strategically involved).

Table 1. Strategic Involvement of HR and Open Organization Measures

Variables	Items	
Strategic involvement of HR	SHR1	How much the HR function reflects the company's mid and long-term strategies
	SHR2	Whether the firm engages in HR planning each year
	SHR3	Whether the HR planning reflects management strategy
	SHR4	How much the HR department contributes to establishing business management strategies
Open organization	OPENORG1	My company communicates its situation in detail to employees
	OPENORG2	Employees speak freely to their managers
	OPENORG3	My company has good communication between teams
	OPENORG4	Employees trust each other
	OPENORG5	My company has an atmosphere wherein hierarchy is stressed
	OPENORG6	My company communicates thoughts and information top down

HRIS adoption (HRIS). This is a binary variable that is based on the response to a question for HR managers in the firm that indicates whether they use HRIS. Response choices were 0 indicating the use of HRIS, and 1 meaning no use of HRIS.

Open organization (OPENORG). This is a cultural variable in our study, which is conceptualized as a culture of trusting and open relationships with minimal bureaucracy, free communications within and across business units (Powell & Dent-Micallef, 1997). Six items with responses ranging from 1 ("strongly disagree") to 5 ("strongly agree") are used to measure the perception of employees about the openness of their firms. The four items are (1) My company communicates its situation in detail to employees, (2) Employees speak freely to their managers, (3) My company has good communication between teams, (4) Employees trust each other, (5) My company has an atmosphere wherein hierarchy is stressed, and (6) My company communicates thoughts and information top down. The item (5) and (6) are reverse coded.

Controls. The research model in this study controls for firm age (AGE), firm size measured as the number of employees (SIZE1: ≤ 299, SIZE2: ≤ 999, SIZE3: ≤ 1999, SIZE4: ≥ 2000), management systems [fully managed by the owner (OWNER), managed by a management specialist with significant owner intervention (MAS_OWNER1), managed by a management specialist with slight owner intervention (MAS_OWNER2), fully managed by a management specialist], the top priority in the company management [improvement the quality of product/service(QUALITY), cost reduction (COST), new product development], having an organization exclusively responsible for HR (HRDPT), the number of HR employees (HREMP), number of HR managers (HRMG), the number of HR senior managers (HRSMG).

Table 2. Variable Operationalizations

Quantitative variable	Operationalization	Mean	Std. dev
SHR	The sum of ratings on four items (SHR1 – SHR4) in a firm *i* in year *t*	7.475	1.691
OPENORG	The average ratings on six items (OPENORG1 – OPENORG6) in a firm *i* in year *t*	3.369	.355
AGE	A firm *i*'s age	33.914	17.620
HREMP	The number of HR employees in a firm *i* in year *t*	2.841	4.941
HRMG	The number of HR managers in a firm *i* in year *t*	2.566	4.979
HRSMG	The number of HR senior managers in a firm *i* in year *t*	.421	.668
Categorical variable	Operationalization	Proportion	Std. dev
HRIS	1 if a firm *i* adopts and operates HRIS, otherwise, coded as 0 in year *t*	.700	.459
SIZE1	1 if the number of employees in a firm *i* is less than or equal to 299, otherwise, coded as 0 in year *t*	.508	.500
SIZE2	1 if the number of employees in a firm *i* is between 300 and 999, otherwise, coded as 0 in year *t*	.350	.477
SIZE3	1 if the number of employees in a firm *i* is between 1000 and 1999, otherwise, coded as 0 in year *t*	.079	.270
OWNER	1 if a firm *i* is fully managed by the owner, otherwise, coded as 0 in year *t*	.469	.499
MAS_OWNER1	1 if a firm *i* is managed by a management specialist with significant owner intervention, otherwise, coded as 0 in year *t*	.165	.371
MAS_OWNER2	1 if a firm *i* is managed by a management specialist with slight owner intervention in, otherwise, coded as 0 in year *t*	.166	.372
SPECIAL	0 if a firm *i* is fully managed by a management specialist in year *t*, otherwise, coded as 0 in year *t*	.200	.400
QUALITY	1 if a firm *i*'s top priority in the company management is improvement the quality of product/service, otherwise, coded as 0 in year *t*	.523	.500
COST	1 if a firm *i*'s top priority in the company management is cost reduction, otherwise, coded as 0 in year *t*	.296	.457
HRDPT	1 if a firm *i* has an organization exclusively responsible for HR, otherwise, coded as 0 in year *t*	.685	.465

Preliminary analysis

Because of the panel nature of our data set, we specify the following equation for the panel models:

$$Y_{it} = X_{it}\beta + u_i + \varepsilon_{it} \quad (1)$$

where *Y* represents the endogenous variable (i.e. SHR); *X* represents a vector of firm characteristics, such as HRIS, OPORG and other control variables; *β*s are the parameters to be estimated; subscript *i* indicates firms and subscript *t* indicates time; u_i represents unobserved time invariant fixed factors associated with a firm *i*; and *ε* is the error term associated with each observation. Our main independent variable (HRIS) is not a time-constant variable, and the Hausman test rejected the null hypothesis ($X^2(16) = 46.76$, $p < 0.01$)(see Table 3) that the key random effect assumption (i.e., that the unobserved effects is uncorrelated with each explanatory variable) is false so that a fixed effect model is used (Wooldridge, 2013).

Table 3: Hausman Test Results

	Fixed Effects (a)	Random Effects (b)	Difference (a-b)	S.E.
HRIS	-2.540	-1.754	-.786	.356
OPENORG	.375	.780	-.405	.097
HRIS*OPENORG	.812	.568	.244	.103
SIZE1	-.657	-.949	.291	.428
SIZE2	-.706	-.655	-.051	.393
SIZE3	-.358	-.200	-.158	.282
AGE	-.044	-.004	-.040	.022
OWNER	-.080	-.269	.189	.145
MAS_OWNER1	-.206	-.304	.098	.118
MAS_OWNER2	.246	.019	.227	.105
QUALITY	-.044	-.075	.030	.086
COST	-.250	-.276	.026	.088
HRDPT	.724	.974	-.250	.070
HREMP	-.015	-.001	-.015	.009
HRMG	.006	.001	.005	.013
HRSMG	.199	.180	.019	.050
Hausman Test	$X^2(16) = 46.76^{**}$			

RESULTS

A summary of the results of the fixed effects regression models are shown in Table 4. The coefficient of HRIS in model1 is positive and significant, supporting Hypothesis 1. In addition, the interaction term HRIS*OPENORG is positive and significant, supporting Hypothesis 2.

Table 4: Fixed Effects Regressions

Independent Variables	Dependent variable: SHR		
	Model1	Model2	Model3
HRIS	.270** (.127)	.237** (.123)	-2.540*** (1.133)
OPENORG		.887*** (.168)	.375 (.227)
HRIS*OPENORG			.812*** (.325)
SIZE1	-.429 (.363)	-.666 (.363)	-.657 (.368)
SIZE2	-.419 (.310)	-.671* (.308)	-.706** (.317)
SIZE3	-.265 (.202)	-.324 (.203)	-.358 (.209)
AGE	-.046 (.024)	-.044 (.024)	-.044 (.023)
OWNER	-.144 (.258)	-.080 (.242)	-.080 (.238)
MAS_OWNER1	-.184 (.237)	-.180 (.230)	-.206 (.227)
MAS_OWNER2	.222 (.187)	.257 (.181)	.246 (.181)
QUALITY	-.049 (.140)	-.037 (.141)	-.044 (.139)
COST	-.331** (.166)	-.255 (.162)	-.250 (.162)
HRDPT	.640*** (.148)	.729*** (.148)	.724*** (.147)
HREMP	-.023 (.013)	-.017 (.023)	-.015 (.013)
HRMG	.014 (.011)	.009 (.010)	.006 (.010)
HRSMG	.190 (.106)	.182 (.106)	.199 (.107)
*Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, *p<0.1 (one-tailed test for the main variables)*			

CONCLUSION AND FUTURE RESARCH

The objective of this study was to analyze and investigate the complementary relationship between HRIS and organizational culture. Based on our preliminary analyses, we find that an open organizational culture positively moderates the impact of HRIS on strategic involvement of HR. We also find that HRIS is positively associated with the strategic involvement of HR. This paper is a preliminary analysis. In the future, we will do robust checks by using different econometric models, such as 2SLS, to handle any endogeneity that might exist in the data.

REFERENCES

Aral, S., & Weill, P. (2007). IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation. *Organization Science*, *18*(5), 763-780.

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of management journal*, *39*(4), 779-801.

Bell, B. S., Lee, S. W., & Yeung, S. K. (2006). The impact of e-HRM on professional competence in HRM: Implications for the development of HR professionals. *Human Resource Management*, 45(3), 295–308.

Bondarouk, T., & Ruël, H. (2009), Electronic human resource management: Challenges in the digital Era. *The International Journal of Human Resource Management*, 20(3), 505–514.

Bondarouk, T., & Ruël, H. (2013). The strategic value of e-HRM: Results from an exploratory study in a governmental organization. *The International Journal of Human Resource Management*, *24*(2), 391-414.

Bondarouk, T., Ruël, H., & van der Heijden, B. (2009). e-HRM effectiveness in a public sector organization: A multi-stakeholder perspective. *The International Journal of Human Resource Management*, *20*(3), 578-590.

Bondarouk, T., Parry, E., & Furtmueller, E. (2017). Electronic HRM: Four decades of research on adoption and consequences. *The International Journal of Human Resource Management*, *28*(1), 98-131.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, *117*(1), 339-376.

Burbach, R., & Dundon, T. (2005). The strategic potential of human resource information systems: Evidence from the Republic of Ireland. *International Employment Relations Review*, *11*(1/2), 97.

Chae, H. C., Koh, C. E., & Prybutok, V. R. (2014). Information technology capability and firm performance: Contradictory findings and their possible causes. *MIS Quarterly*, *38*(1), 305-326.

Dery, K., Hall, R., Wailes, N., & Wiblen, S. (2013). Lost in translation? An actor-network approach to HRIS implementation. *The Journal of Strategic Information Systems*, *22*(3), 225-237.

DeSanctis, G. (1986). Human resource information systems: a current assessment. *MIS Quarterly*, 10(1), 15-27.

Devaraj, S., & Kohli, R. (2003). Performance impacts of information technology: Is actual usage the missing link? *Management Science*, *49*(3), 273-289.

Haines, V. Y., & Lafleur, G. (2008). Information technology usage and human resource roles and effectiveness. *Human Resource Management, 47*(3), 525-540.

Holland, C., G. Lockett & I. Blackman (1992). Planning for electronic data interchange. *Strategic Management Journal,* 13(7), 539–550.

Hussain, Z., Wallace, J., & Cornelius, N. E. (2007). The use and impact of human resource information systems on human resource management professionals. *Information & Management*, *44*(1), 74-89.

Kettinger, W., V. Grover, S. Guha & A. Segars (1994). Strategic information systems revisited: A study in sustainability and performance. *MIS Quarterly*, 18(1), 31–58.

Kettinger, W. J., Grover, V., & Segars, A. H. (1995). Do strategic systems really pay off? An analysis of classic strategic IT cases. *Information System Management*, *12*(1), 35-43.

KIS 2015. KISData. Available at: http://www.niceinfo.co.kr/business/KISData.nice

Korea Research Institute for Vocational Education & Training (KRIVET) (2015). About HCCP. Available at: http://www.krivet.re.kr/eng/eu/eh/euDDALs.jsp

Leidner, D. E., & Kayworth, T. (2006). A review of culture in information systems research: Toward a theory of information technology culture conflict. *MIS Quarterly*, *30*(2), 357-399.

Lippert, S. K., & Michael Swiercz, P. (2005). Human resource information systems (HRIS) and technology trust. *Journal of Information Science*, *31*(5), 340-353.

Liu, D., Gong, Y., Zhou, J., & Huang, J. C. (2017). Human resource systems, employee creativity, and firm innovation: The moderating role of firm ownership. *Academy of Management Journal*, *60*(3), 1164-1188.

Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2013). Analyzing the impact of HRIS implementations on HR personnel's job satisfaction and turnover intention. *The Journal of Strategic Information Systems*, *22*(3), 193-207.

Marler, J. H., & Fisher, S. L. (2013). An evidence-based review of e-HRM and strategic human resource management. *Human Resource Management Review*, *23*(1), 18-36.

Marler, J. H., & Parry, E. (2016). Human resource management, strategic involvement and e-HRM technology. *The International Journal of Human Resource Management*, *27*(19), 2233-2253.

Martin, G., & Reddington, M. (2010). Theorizing the links between e-HR and strategic HRM: a model, case illustration and reflections. *The International Journal of Human Resource Management*, *21*(10), 1553-1574.

Martin, G., Reddington, M., & Alexander, H. (2008), Technology, Outsourcing and Transforming HR, Amsterdam: Elsevier.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, *28*(2), 283-322.

Miller, E. & A. Rice (1967). *Systems of Organization: The Control of Task and Sentient Boundaries*. Tavistock, London.

Mithas, S., Almirall, D., & Krishnan, M. S. (2006). Do CRM systems cause one-to-one marketing effectiveness?" *Statistical Science,* 21(2), 223-233.

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly*, *35*(1), 237-256.

Mithas, S., & Rust, R. T. (2016). How information technology strategy and investments influence firm performance: Conjecture and empirical evidence. *MIS Quarterly*, *40*(1), 223-245.

Mithas, S., Tafti, A., Bardhan, I., & Goh, J. M. (2012). Information technology and firm profitability: mechanisms and empirical evidence. *MIS Quarterly*, 36(1), 205-224.

Mithas, S., Whitaker, J., & Tafti, A. (2017). Information technology, revenues, and profits: Exploring the role of foreign and domestic operations. *Information Systems Research*, *28*(2), 430-444.

Ngai, E. W. T., & Wat, F. K. T. (2006). Human resource information systems: A review and empirical analysis. *Personnel Review*, *35*(3), 297-314.

Nevo, S., & Wade, M. R. (2010). The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships. *MIS Quarterly*, 34(1), 163-183.

Oh, W., & Pinsonneault, A. (2007). On the assessment of the strategic value of information technologies: conceptual and analytical approaches. *MIS Quarterly*, 31(2), 239-265

Paauwe, J. (2004), *HRM and Performance: Achieving Long Term Viability*, Oxford: Open University Press.

Parry, E. (2011). An examination of e-HRM as a means to increase the value of the HR function. *The International Journal of Human Resource Management*, *22*(5), 1146-1162.

Parry, E., & Tyson, S. (2011). Desired goals and actual outcomes of e-HRM. *Human Resource Management Journal*, *21*(3), 335-354.

Parry, E., Tyson, S., Selbie, D., & Leighton, R. (2007), *HR and Technology: Impact and Advantages*, London: Charted Institute of Personnel and Development.

Piccoli, G., & Ives, B. (2005). IT-dependent strategic initiatives and sustained competitive advantage: a review and synthesis of the literature. *MIS Quarterly*, *29*(4), 747-776.

Powell, T. C., & Dent-Micallef, A. (1997). Information technology as competitive advantage: The role of human, business, and technology resources. *Strategic Management Journal*, *18*(5), 375-405.

Ruël, H.J.M., Bondarouk, T.V., & Looise, J.C. (2004), *E-HRM: Innovation or Irritation. An Exploration of Web-Based Human Resource Management in Large Companies*, Utrecht: Lemma Publishers.

Shrivastava, S., & Shaw, J. B. (2003). Liberating HR through technology. *Human Resource Management*, *42*(3), 201-222.

Sun, L. Y., Aryee, S., & Law, K. S. (2007). High-performance human resource practices, citizenship behavior, and organizational performance: A relational perspective. *Academy of management Journal*, *50*(3), 558-577.

Teece, D. J. (1986). Firm Boundaries, Technological Innovation and Strategic Planning,. in *The Economics of Strategic Planning*, G. L. Thomas (ed.), D. C. Heath, Lexington, MA, 187-199.

Voermans, M., & Van Veldhoven, M. (2007), Attitude Towards e-HRM: An Empirical Study at Philips. *Personnel Review*, 36(6), 887–902.

Wade, M., & Hulland, J. (2004). The resource-based view and information systems research: Review, extension, and suggestions for future research. *MIS Quarterly*, *28*(1), 107-142.

Weill, P. (1992). The relationship between investment in information technology and firm performance: A study of the valve manufacturing sector. *Information Systems Research*, *3*(4), 307-333.

Wickramasinghe, V. (2010). Employee perceptions towards web-based human resource management systems in Sri Lanka. *The International Journal of Human Resource Management*, *21*(10), 1617-1630.

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach 5th*, Mason, OH: South-Western Cengage Learning

Zuboff, S. (1988). *In the Age of the Smart Machine* Basic Books, New York.

DECISION SCIENCES INSTITUTE
A Last Mile Delivery Fleet Model with Drones

Jay R. Brown
Loyola University Maryland
Email: jbrown11@loyola.edu

Maxim A. Bushuev
Morgan State University
Email: maxim.bushuev@morgan.edu

ABSTRACT

The development and potential adoption of drones as delivery vehicles creates incredible opportunities and unique challenges for last mile delivery. This research presents a last mile delivery fleet model with drones that can be modified and expanded over time as the theoretical morphs into the practical. The model shows the optimal number of drones needed based on demand and how additional batteries can mitigate the total number of drones. The model is presented with a central depot and a delivery radius dependent on the range of the drones.

KEYWORDS: Last mile problem, Drones, UAVs, Supply chain delivery, Logistics

INTRODUCTION

The last mile problem (LMP), which is defined as the optimizing of the last leg of the business-to-consumer delivery service, is of particular concern to supply chain managers who manage road freight transportation. The LMP is considered to be one of the most costly and highest polluting segments of the supply chain and is often the least efficient part of the supply chain due to the high degree of empty running (Gevaers et al., 2011). While costs vary with population density, product type, package size, and package weight, last mile delivery incurs the highest transportation costs in the supply chain (Chopra, 2003). Drones or unmanned aerial vehicles (UAVs) offer a potential resource for companies to reduce both cost and carbon emissions associated with last mile delivery. This paper develops a last mile delivery fleet model with drones delivering one package per trip from a central depot. The model shows the optimal number of drones needed based on demand and how additional batteries can mitigate the total number of drones. The model is presented with a delivery radius dependent on the range of the drones.

As e-commerce continues to grow, a greater burden is placed on last mile delivery. Logistics systems which were traditionally designed to accommodate a single lot delivery of multiple products from one business to another are now under pressure to deliver high volumes of single parcels from a transportation/warehouse hub to multiple individual customers. With the global e-commerce market growing 18% to $2.86 trillion in 2018, solving last mile delivery issues is a major customer service concern for firms (Young, 2019).

In a climate of enhanced awareness of environmentally sustainable business practices, greater pressure is being placed on organizations to advance sustainable logistics within the operation of their supply chains (Abbasi and Nilsson, 2016; Björklund et al., 2016; Golicic et al., 2010). Comprehensive reviews of sustainable logistics have been conducted for various modes of freight transportation including roadways (Demir et al., 2014), railways (Aditjandra et al., 2016), maritime (Davarzani et al., 2015), air (Teoh, and Khoo, 2016), and intermodal (Roso, 2013). The carbon footprint associated with freight transport across all modes is rapidly becoming a key managerial concern and is especially prevalent in road transportation by transport vehicle in North America and Europe. Drones provide a delivery option that can result in a lower overall carbon footprint, but it can depend on the delivery density and the efficiency of the drones and trucks (Goodchild & Toy, 2018).

Determining an optimal drone delivery fleet for dealing with the LMP including the number of vehicles, available delivery hours, delivery range, and total freight capacity is a problem with a solution that varies depending on the customer demand for that delivery cycle. If demand is high, more drones, more delivery time, or more freight capacity may be needed. On a day with less demand, the opposite is true. Demand for last mile delivery is typically not constant and as a result changes to the fleet operating characteristics may change from week to week or even from day to day. In this paper we present a last mile delivery model with drones that can be used for planning and to aid decision makers in managing the operating characteristics of their drone delivery fleet in support of the LMP.

The contributions of this research add to the literature by providing a base model that can be modified and adapted over time as the theory of last mile delivery with drones becomes reality. The model provides a method for calculating the size of the necessary drone fleet based on either deterministic demand or stochastic demand with a set service level.

LITERATURE REVIEW

As a gateway to the general literature on the LMP the reader is referred to the review papers of Ranierie et al. (2018) and Datta (2018). These review papers provide a general overview of the scope of problems (logistical and environmental) associated with last mile delivery in supply chain management and provide a foundation for areas of future research on the LMP.

Research on drones and how they may impact last mile delivery in the coming years has grown steadily as companies like Amazon and Google are investigating drones and how they may be utilized in the future (Gharibi et al, 2016). Perhaps the biggest obstacle holding up the adoption of drones for last mile delivery involves commercial drone regulations. See Jones (2017) for details on these regulations by country.

Goodchild & Toy (2018) compare the CO_2 emissions of a traditional truck vehicle fleet with that of a hypothetical fleet of drones. Their findings suggest that the density of deliveries and the efficiency of the trucks and drones are large factors in determining the overall carbon footprint of the last mile delivery.

Murray & Chu (2015) introduce mathematical programming models aimed at optimal routing and scheduling of drones in conjunction with delivery trucks. They find this type of delivery system can provide faster receipt of customer orders with a reduced environmental impact. Boysen et al (2018) find that a truck fleet can considerably be reduced if drones with decentralized depots support the delivery process. Agatz et al (2018) show that substantial savings are possible when combining drones with trucks in comparison to truck-only delivery.

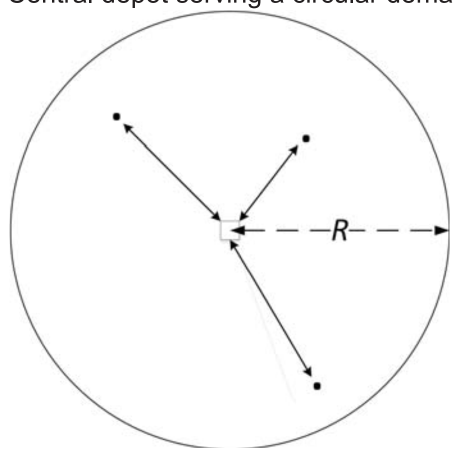
Chiang et al (2019) show that the use of drones can help save fixed logistics costs by reducing the total delivery time and the number of vehicles required. Shifting smaller package delivery from trucks to drones, when adequately deployed, would cut down energy use and CO_2 emissions. To realize the environmental benefits of drone delivery, firms are recommended to carefully plan and control the routing and coordination of the delivery vehicles and drones.

The review of research studies on drones and the LMP identifies that drones have the potential to reduce cost, time, and carbon emissions if applied appropriately to the last mile problem. Much of the research is limited by the knowledge of the current ability of known drones and makes conclusions based on these current technologies. By the time drone flight regulations are in place to allow last mile delivery with drones, the technology may be more advanced and change what was once thought possible. This research develops a base last mile delivery fleet model with drones delivering one package per trip from a central depot. The model is adaptable to future advancements and provides a foundation for further research.

MODEL

The basis for the last mile delivery fleet model with drones is a circular region with a central depot or warehouse. The demand is assumed to be uniformly and randomly distributed throughout the region. Figure 1 shows this general model with drones leaving the depot, making a delivery within the service region, and returning to the depot to either begin a new delivery, have its battery charged, or swap its battery and begin a new delivery.

Figure 1: Central depot serving a circular demand region



The radius, R, of a circular demand region around a central depot depends on the comfortable roundtrip range of the drones loaded with a maximum payload. Since the drone will be empty on the return trip, the range of the drone is determined by the average of the range of a fully loaded drone and an empty drone. In this case the total flight distance or drone range, r_m, is defined as

$$r_m = \frac{r_e + r_f}{2} \tag{1}$$

where

r_e = maximum range of an empty drone

r_f = maximum range of a drone at maximum capacity

The radius of the demand region can be defined as

$$R = \frac{r_m p}{2} \tag{2}$$

where

p = safety factor between 0 and 1 (i.e. 90%)

The expected or average distance from a customer to the depot if demand is uniformly and randomly distributed can be defined mathematically (Brown & Guiffrida, 2014) as

$$\bar{L} = \frac{2}{3}R \tag{3}$$

Thus, the mean roundtrip distance is

$$2\bar{L} = \frac{4}{3}R \tag{4}$$

The average time duration of a delivery can be defined as

$$T_M = \frac{4R}{3V} \quad (5)$$

where

V = average travel velocity of a drone

A drone fleet can be set up either with additional batteries available to swap out or with just one battery per drone. The number of drones needed with battery swapping can be defined as (6). The ceiling brackets indicate the result should always be rounded up. The average setup time includes the time needed to change the battery, load the new parcel, set the delivery location, and launching the drone. Troudi et al (2018) map out the expected activities of how a drone operator manages the parcel delivery process between flights.

$$N_D = \left\lceil \frac{D\left(\frac{4R}{3V} + A\right)}{H} \right\rceil \quad (6)$$

where

D = demand per period

A = average setup time (spent between deliveries) including battery swap

H = operational time per period (i.e. hours per day)

The average number of deliveries made by each drone per period depends on demand and can be defined as

$$D_D = \frac{D}{N_D} \quad (7)$$

The minimum number of batteries depends on a combination of charge time, demand, available time for deliveries, and the number of drones. An important factor in deriving the minimum number of batteries is the average time each drone has between deliveries (average cycle time that includes setup and delivery). This can be defined as

$$T_B = \frac{HN_D}{D} \quad (8)$$

In order to arrive at the minimum number of batteries, we divide the time it takes to charge a battery by (8), multiply that by the number of drones, and add the number of drones. After some term simplification, the minimum number of batteries needed can be defined as

$$B = \left\lceil \frac{C_T D}{H} + N_D \right\rceil \quad (9)$$

where

C_T = charge time for a battery

Charge time is variable depending on how much power is left after the previous trip. This model assumes linear battery usage and charging times. For instance, if 50% of the battery has been used, we assume the charge time is $0.5C_T$. See Dorling et al (2016) for information on the linear energy consumption of drones based on total weight. For simplicity, this model assumes no battery capacity fade over time.

Equation (10) can be reworded as the optimal number of spare batteries, which would assume that the drones would each contain a battery so this would be the number of batteries that are separate from the drones. The optimal number of spare batteries can be defined as

$$B_s = \left\lceil \frac{C_T D}{H} \right\rceil \quad (10)$$

The number of drones needed without additional batteries to swap out changes since a drone will be out of service while charging. The average setup time here includes the time needed to plug the returning drone into its charger, unplug the departing drone, load the new parcel, set the delivery location, and launching the departing drone. The number of drones needed without battery swapping can be defined as

$$N_N = \left\lceil \frac{D\left(\frac{4R}{3V} + A_N + C_T\right)}{H} \right\rceil \quad (11)$$

where

A_N = average setup time (spent between deliveries) without battery swap

NUMERICAL EXAMPLES

In this section, the model is illustrated with a numerical example to demonstrate the differences in the makeup of a fleet of last mile delivery drones with battery swapping and without battery swapping. The following parameter values have been assigned to support the model illustration: r_f = 14.0 miles; r_e = 16.0 miles; p = 90%; D = 120 customers per period; V = 30 miles per hour; A = 0.1 hours; A_N = 0.1 hours; H = 8 hours; C_T = 1 hour.

Based on these parameters, the model was solved for the battery swapping scenario. Table 1 contains the results with battery swapping.

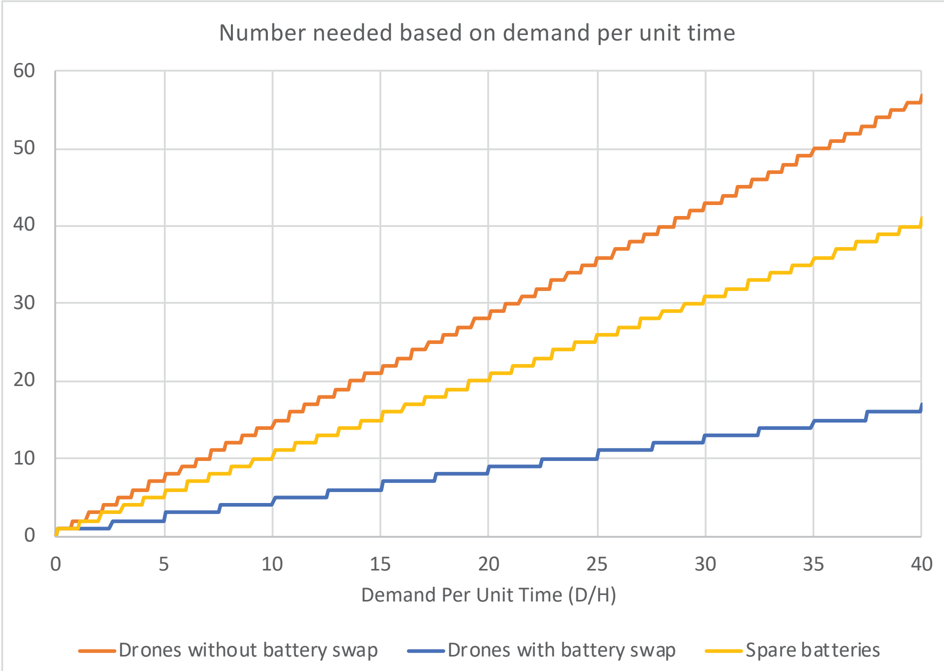
Table 1: Results for deterministic demand example with battery swapping

VARIABLE	RESULT
N_D	6 drones
D_D	20 deliveries per drone per period
T_M	0.3 hours average trip duration
B_s	15 spare batteries needed

Without battery swapping, the number of drones needed would be 21 drones based on these parameters. The only upside to not employing battery swapping is less wear and tear on each drone since each drone would only need to do 5.7 deliveries per period instead of 20 deliveries per period. The additional up-front cost, however, would be substantial. In addition, more storage space would be required for a fleet of 21 drones.

While most of the parameters will remain constant once they are first determined, demand may not. As demand changes, the number of drones needed also changes. The key ratio is D/H or demand per unit time. As the ratio moves, the number of drones and batteries moves at a steady rate. Figure 2 shows how the number of drones and batteries changes as demand per unit time changes.

Figure 2: Drones and batteries needed based on demand per unit time



This information can be used with a demand distribution to determine the optimal number of drones and batteries needed with a desired service level, *s*. This can be illustrated with a new example. Let *f(D)* denote the daily demand function. For this numerical example, *f(D)* is normally distributed ~N(μ,σ). The following parameter values have been added to the previous example's parameters to support the model illustration:
μ = 120; σ = 20; *s* = 99%.

Based on this demand distribution, the last mile delivery fleet would have to be constructed to serve 166.5 customers per period or 20.8 customers per unit time. Table 2 contains the results with this stochastic demand example.

Table 2: Stochastic demand results with battery swapping

VARIABLE	RESULT
N_D	9 drones
B_s	21 spare batteries needed
N_N	30 drones

A fleet of 9 drones with 21 spare batteries (30 total batteries) would be able to serve up to 179 customers based on the parameters. This equates to a service level of 99.84% based on the demand distribution. A fleet of 30 drones without battery swapping would be able to serve up to 171 customers, which equates to a service level of 99.46% based on the demand distribution.

CONCLUSIONS

This paper developed a last mile delivery fleet model with drones delivering from a central depot to find the minimum number of drones needed based on either deterministic or stochastic demand. The model is adaptable to future advancements and provides a foundation for further research. Findings suggest that additional batteries can mitigate the total number of drones and presumably lower the total cost. The model also allows a logistics manager to determine the probability of meeting changing demand with a current fleet of drones.

The contributions of this research add to the literature by providing a base model that can be modified and adapted over time as the theory of last mile delivery with drones becomes reality. The model provides a way to calculate the size of the necessary drone fleet based on either deterministic demand or stochastic demand with a set service level.

The proposed base model has several limitations that can be addresses in the future research. First, it does not take into account that the weight of the package carried by drone will affect its range and battery usage. The model uses the maximum drone capacity to calculate range. Second, the model does not consider the likelihood of no-fly zones and how they would impact flight distance and scheduling within the region . That situation can be addressed in the future.

Future research could also look at ways to extend the delivery range of drones either through more efficient drones or through innovative methods such as distributed landing locations with available batteries and charging facilities. Another extension could be the possibility of a drone making more than one delivery per trip if the cargo capacity and trip duration allowed. This could save time and money and reduce the total travel distance. At this point, there are a myriad of possibilities and applications for drones and last mile delivery and research in this area should explore these possibilities and extend the boundaries as much as possible.

REFERENCES

Abbasi, M., & Nilsson, F. (2016). Developing environmentally sustainable logistics: Exploring themes and challenges from a logistics service providers' perspective. *Transportation Research Part D: Transport and Environment*, *46*, 273-283.

Aditjandra, P. T., Zunder, T. H., Islam, D. M. Z., & Palacin, R. (2016). Green rail transportation: Improving rail freight to support green corridors. In *Green Transportation Logistics* (pp. 413-454). Springer, Cham.

Agatz, N., Bouman, P., & Schmidt, M. (2018). Optimization approaches for the traveling salesman problem with drone. *Transportation Science*, *52*(4), 965-981.

Björklund, M., Forslund, H., & Persdotter Isaksson, M. (2016). Exploring logistics-related environmental sustainability in large retailers. *International Journal of Retail & Distribution Management*, *44*(1), 38-57.

Boysen, N., Schwerdfeger, S., & Weidinger, F. (2018). Scheduling last-mile deliveries with truck-based autonomous robots. *European Journal of Operational Research*, *271*(3), 1085-1099.

Brown, J. R., & Guiffrida, A. L. (2014). Carbon emissions comparison of last mile delivery versus customer pickup. *International Journal of Logistics Research and Applications*, *17*(6), 503-521.

Chiang, W. C., Li, Y., Shang, J., & Urban, T. L. (2019). Impact of drone delivery on sustainability and cost: Realizing the UAV potential through vehicle routing optimization. *Applied Energy*, *242*, 1164-1175.

Chopra, S. (2003). Designing the distribution network in a supply chain. *Transportation Research Part E: Logistics and Transportation Review*, *39*(2), 123-140.

Datta, P. P. (2018). Decision framework for selecting last mile delivery performance in Indian e-commerce companies. *International Journal of Information and Decision Sciences*, *10*(1), 77-93.

Davarzani, H., Fahimnia, B., Bell, M. G., & Sarkis, J. (2015). A Review of the Literature of Green Ports and Maritime Logistics. In *Green Logistics and Transportation* (pp. 149-158). Springer, Cham.

Demir, E., Bektaş, T., & Laporte, G. (2014). A review of recent research on green road freight transportation. *European Journal of Operational Research*, *237*(3), 775-793.

Dorling, K., Heinrichs, J., Messier, G. G., & Magierowski, S. (2017). Vehicle routing problems for drone delivery. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(1), 70-85.

Gevaers, R., Van de Voorde, E., & Vanelslander, T. (2011). Characteristics and typology of last-mile logistics from an innovation perspective in an urban context. *City Distribution and Urban Freight Transport: Multiple Perspectives, Edward Elgar Publishing*, 56-71.

Gharibi, M., Boutaba, R., & Waslander, S. L. (2016). Internet of drones. *IEEE Access*, *4*, 1148-1162.

Golicic, S., Boerstler, C., & Ellram, L. (2010). 'Greening'the Transportation in Your Supply Chain. *MIT Sloan Management Review*, *51*(2), 47.

Goodchild, A., & Toy, J. (2018). Delivery by drone: An evaluation of unmanned aerial vehicle technology in reducing CO2 emissions in the delivery service industry. *Transportation Research Part D: Transport and Environment*, *61*, 58-67.

Jones, T. (2017). *International Commercial Drone Regulation and Drone Delivery Services* (No. RR-1718/3-RC). RR-1718/3-RC.

Murray, C. C., & Chu, A. G. (2015). The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies*, *54*, 86-109.

Ranieri, L., Digiesi, S., Silvestri, B., & Roccotelli, M. (2018). A review of last mile logistics innovations in an externalities cost reduction vision. *Sustainability*, *10*(3), 782.

Roso, V. (2013). Sustainable intermodal transport via dry ports–importance of directional development. *World Review of Intermodal Transportation Research*, *4*(2-3), 140-156.

Teoh, L. E., & Khoo, H. L. (2016). Green air transport system: An overview of issues, strategies and challenges. *KSCE journal of civil engineering*, *20*(3), 1040-1052.

Troudi, A., Addouche, S. A., Dellagi, S., & Mhamedi, A. (2018). Sizing of the Drone Delivery Fleet Considering Energy Autonomy. *Sustainability*, *10*(9), 3344.

Young, J. (2019). *Global e-commerce sales grow 18% in 2018*. Retrieved from https://www.digitalcommerce360.com/article/global-ecommerce-sales/, May 7.

DECISION SCIENCES INSTITUTE

Exploring the Mediating Effect of Price on the Link between CSR-Driven Perception, Moral Obligation, and Purchase Intention

Yeh, Quey-Jen
National Cheng Kung University
Email: yehqj@mail.ncku.edu.tw

Nguyen Thi Hong Nhung
National Cheng Kung University
Email: rosienguyenhongnhung@gmail.com

ABSTRACT

Ethical consumers have a dilemma in making a purchase decision: “buy good or buy well.” This study modified the theory of reasoned action to explore driven personal motives of behavioural intention to purchase CSR-related products from customers’ perspectives in the context of ethical consumerism. This is the first study exploring the full mediating effect of price sensitivity on the association between CSR-driven perceptions of personal values and purchase intention. SEM analytical techniques were used to analyze 311 valid questionnaires collected in a Taiwanese dairy food marketplace. Theoretical and practical implications are discussed.

KEYWORDS: Ethical consumption, CSR perception, Price sensitivity, Moral Obligation, Mediating effect, Purchase Intention

INTRODUCTION

Corporate Social Responsibility (CSR) is becoming a global phenomenon and a great concern for society. Firms are using it as one of the profitable investments necessary to maintain a healthy relationship with salient stakeholders and to convey a solid brand image in the public eye. It seems that all sectors of the economy are influenced by the increases in societal expectations and interests in CSR initiatives. Mostly, CSR in the food industry has raised much attention from overseas scholars with many studies from different approaches (Devin & Richards, 2016; Kim, 2017; Maloni & Brown, 2006; Tian, Wang, & Yang, 2011; Wei, Kim, Miao, Behnke, & Almanza, 2018). CSR is highly relevant to the food industry because this sector has a high dependence on natural resources (with multiple effects on the environment) and the food industry is integrated with several societal and ethical concerns including animal welfare, labor’s right protection, food safety, fair-trade and so on.

Consumers are often cautious about food safety and brand credibility based on purchase decision experiences. Health safety becomes much more crucial than ever before because of the suffering of society due to information asymmetry relevant to food quality and whether companies manufactured the food in responsible ways or not. Being fearful of seller’s opportunism, consumers tend to use credence attributes such as CSR initiatives like a safeguard against risks. A recent study pointed out that if firms’ stimulate CSR initiatives, the information asymmetry will reduce (Cui, Jo, & Na, 2018). These CSR activities can affect consumer purchasing behaviors directly since these reflect consumers' perceptions. Researchers have shown that today’s consumers increasingly believe firms engage in CSR campaigns and are eager to pay higher prices for similar quality products perceived as socially responsibility oriented (Becker-Olsen, Cudmore, & Hill, 2006; Carvalho, Sen, de Oliveira Mota, & de Lima, 2010; Creyer, 1997; Mohr, Webb, & Harris, 2001; Rodrigues & Borges, 2015; Sen, Bhattacharya, & Korschun, 2006). Nevertheless, the customer's perception of CSR is diverse and still mostly unexplored. To fill the gap, this study reinvestigated customer's CSR-driven perceptions and its influences on patronage intention toward CSR-related products.

(Cooper-Martin & Holbrook, 1993) defined ethical consumption as "consumption experiences that are affected by the consumer's ethical concerns." Meanwhile, ethical concerns or moral obligations are understood as intrinsic ethical motives that play essential roles in shaping standards of behavior. Several consumption behavior studies found that customers perceived a strong sense of moral obligations when making decisions inherently in relation to ethical issues (Bommer, Gratto, Gravander, & Tuttle, 1987; Shaw & Shiu, 2003). For instance, Shaw carried out a study within the context of fair trade grocery and concluded that the role of moral obligations was imperative in predicting ethical purchase intentions (Shaw & Shiu, 2002). (Arvola et al., 2008) proved a vital link between moral attitudes and patronage intention of organic food. However, a wide variety of literature reviews suggested that the use of consumer's internal ethics, such as moral obligation in the CSR context, had not yet been cultivated.
On the other hand, (De Pelsmacker, Driesen, & Rayp, 2005) found that consumer responses, in some cases, did not align with their positive attitude toward ethical products. A survey (Deng & Jiang, 2011) found that only 12% of respondents would purchase an ethical product among 44% of respondents who made positive responses to cause-marketing practices. Additionally, CSR is not the most critical factor that affects consumers in purchasing decisions because traditional standards, such as price, quality, and brand familiarity are still the dominant criteria in the buying process for consumers (Boulstridge & Carrigan, 2000). A question raised in this research asked if "customers' patronage intention is consistent with their moral values and CSR perceptions under the effect of price sensitivity?". Notably, the role of price sensitivity in the context of social responsibility is still limited and controversial. Therefore, we aimed to explore this research gap.

We investigated the three following primary objectives in this paper. Firstly, we identified significant antecedences of consumers' purchase intentions to CSR-related products in ethical consumption. Secondly, we explored the mediating effect of price sensitivity in the association between purchase intention to CSR-related products and its antecedences. Thirdly, we demonstrated the trade-offs between personal ethical values and price perception in the purchase decision-making process from consumer perspectives. The present study concluded that several significant findings contributed to both theoretical development and managerial practices. It especially recognized the mediating effect of price sensitivity in the context of CSR. We chose milk as an example of experience good since milk also had long been associated with human basic physical needs in daily consumption. That allowed less restriction in data collection. The study followed a quantitative approach structured. We begin reviewing relevant literature of primary constructs and develop hypotheses to test. Next, the research methodology is described and is accompanied by empirical results. Finally, the authors summarize research findings and conclude implications.

LITERATURE REVIEW & HYPOTHESES DEVELOPMENT

Corporate Social Responsibility Initiatives

Corporate social responsibility developed out of the early 1950s and became popular in the 1970s (Carroll, 1999). Nowadays, in the highly competitive market, CSR is a global phenomenon and widely recognized as for strategic importance to a firm. This strategic importance includes deriving consumer-friendly images, reinforcing relationship with significant stakeholders, enhancing competitive advantages, and boosting financial performance to outperform rivals (Becker-Olsen et al., 2006; McGuire, Sundgren, & Schneeweis, 1988; McWilliams, Siegel, & Wright, 2006; Porter & Kramer, 2006; Sen et al., 2006; Shrivastava, 1995; Valentine & Fleischman, 2008). CSR is a broad and complex concept to the extent of what firms should do to keep commitments and respect the community's interests for long-run sustainability. No longer a new concept, CSR has numerous explanations (Perrini, 2006). Nevertheless, many arguments challenged scholars and industrialists as to how CSR should be correctly defined in both the academic and corporate worlds (Dahlsrud, 2008). This probably occurs because CSR's definitions are all biased towards specific interests for distinguished stakeholders (Freeman, Harrison, Wicks, Parmar, & De Colle, 2010).

Every firm operates in a social environment with various stakeholders who have both direct and indirect impacts on its survival. Therefore, the scope of CSR initiatives emphasizes organizational commitments for stakeholders' well-being (Mohr et al., 2001). Indeed, CSR initiatives refer to both ethical and legal obligations that ultimately provide tremendous benefits and diminish the negative influences of firms' actions on the business environment in accordance with moral standards, social requirements, and stakeholder expectations (Berens, Riel, & Bruggen, 2005; Mohr et al., 2001; Pérez, del Mar García de los Salmones, & Rodríguez del Bosque, 2013). In the present research, we adopted the definition of CSR developed by (Dahlsrud, 2008) at one of the primary purposes to examine CSR perceptions concerned by customers who were primarily salient stakeholders of all-size corporates. He drew a relatively general conclusion about the conceptualization of CSR initiatives by systemizing 37 published CSR definitions and viewed CSR as a multi-dimension social construction that encompassed five aspects that included environmental, social, economic, stakeholder, and voluntariness dimensions. It emphasized that there were not many previous studies that examined multiple dimensions of CSR, most concentrated on the role of each dimension of CSR in consumer evaluation (Choi & Ng, 2011)

The theoretical approach in predicting customer response

In recent years, the growth of ethics studies, including CSR, business ethics, and corporate citizenship evoked great insights for academic researchers to produce various theoretical developments from perspectives of both business and customer ethics. A Japanese scholar, Fukukawa, reviewed these studies using two approaches; namely, a normative or descriptive approach to explore customer's responses to relevantly ethical issues in the marketplace (Fukukawa, 2003). In that research, she posited that the normative approach rooted in philosophy, defined ethical conducts of "what a person should do," while the descriptive approach emphasized psychological perspectives as personal values, behavioral psychology, and pro-social psychology referring to "what a person would do." Descriptive approaches can be extremely beneficial in exploring the nature of human beings in ethical contexts, and it develops from applications of cognitive and attitude-behavioral theories such as the Theory of Reasoned Action (TRA) (M. Fishbein & Ajzen, 1975) or the Theory of Planned Behavior (TPB) model as extended from the TRA model (Ajzen, 1991). Ethics scholars applying these descriptive oriented theories investigate underlying personal motives of ethical consumption to predict purchase intention, pre-purchase alternative evaluation, and actual purchase behaviors.

Over the past decade, Ajzen and Fishbein's TRA, one of the most developed theoretical models, has been applied to explore antecedents of consumer intention and behaviors. In the middle and late 1970s, the theory initially introduced by (M. E. Fishbein, 1967) pointed out that both actual behaviors and intentions were the likely outcomes that reflected the information-based processing beliefs and attitudes; an assumption that human actions react under their volitional control. Many researchers attempted to prove the predictive utility of TRA model, among that the findings concluded by (Sheppard, Hartwick, & Warshaw, 1988) seemed to be overwhelming proofs that strongly support the overall reliability of the TRA model. Furthermore, TRA is recognized as an appropriate theoretical approach applied in customer-based research as better than other models (Wilson, Mathews, & Harvey, 1975). In this study, the descriptive approach was adopted, and we modified the TRA model by examining the interrelationship of consumer perception, personal values, and patronage intention toward products made in responsible ways, especially in the food industry. The mediating effects of price sensitivity were also included and examined in this modified framework.

Purchase Intention

Purchase intention is a kind of prior-purchase representing motivations to make a deal with a seller to buy a specific product or service. In other words, purchase intention has a power of foretelling purchasing and consuming process (Ghosh, Dasgupta, & Ghosh, 1990). The concept of consumer purchase intention initially originated from several psychological and behavioral studies. In this study, purchase intention merely was understood as customers' willingness to buy CSR-related products. Besides, a wide range of marketing research proved that purchase intention highly correlated with customer’s beliefs, affective attitudes, and actual purchase. Hence, it quietly makes sense in the present study that we applied TRA to examine customers purchase intention upon the integration of moral obligation (ethical values), CSR-initiatives, and customer’s price sensitivity in the dairy food market. Nevertheless, this work found that the current study was not carried out to measure customer’s purchase behavior explicitly, but concentrated on customer’s purchase intention instead.

Consumer’s CSR-driven perception

Consumers altogether want companies to do good things that are responsible and beneficial to the whole community. Moreover, consumer perceptions of CSR initiatives critically influence involvement and responsiveness to CSR programs (Barone, Miyazaki, & Taylor, 2000; Becker-Olsen et al., 2006). These CSR activities can even directly affect consumer purchasing decisions when these initiatives reflect the CSR beliefs of the consumer. However, in the literature, the customer’s perception of CSR is diverse and still mostly unexplored. To fill the gap, we included this construct into our study to reinvestigate customer’s CSR perception and its influences on ethical consumption toward CSR-related products. Therefore, in this paper, we defined consumer’s CSR-driven perception as the degree to which consumers perceived and believed a firm engaged in its activities interrelated to social causes. Therefore, the construct of CSR-driven beliefs in the current research proposed to consist of two components: CSR expectation (prosocial psychology beliefs) and perceived CSR (cognitive beliefs).

<u>CSR expectation</u>

Research has shown that today’s consumers increasingly believe firms engage in CSR campaigns (Becker-Olsen et al., 2006; Harrison, 2003). Customer expectations can be understood as whatever they desire, want, or hope for what a firm should do, rather than what a firm would do (Zeithaml, Berry, & Parasuraman, 1993). It is likely that customers’ expectations towards CSR initiatives incorporate the way customers compensate the firms (Creyer, 1997;

Nebenzahl, Jaffe, & Kavak, 2001). Therefore, consumer's CSR expectations are indispensable in determining purchasing decisions. Customers bear in mind that meeting the needs of society is a responsibility for firms, and evidence suggests a willingness to buy products perceived as socially responsible is a way to support personal beliefs further (Mohr et al., 2001). Therefore, the definition of CSR expectations in this paper identifies merely some extents that consumer's desire and want every firm to behave ethically in complying with social responsibilities. We proposed the presence of a positive relationship between CSR expectation and purchase intention. Hence, the following hypothesizes was developed.

Hypothesis 1a: Customer's CSR expectations positively impact on their intention to purchase CSR-related product.

Perceived CSR

CSR is portrayed commonly as activities related to society's perception, as far as its obligations towards society (Sen and Bhattacharya, 2001). Previous studies implied that firms that embraced social initiatives could create positive cognitive perceptions; the expectation for customers (Becker-Olsen et al., 2006; Choi & Ng, 2011; Creyer, 1997; Tian et al., 2011). A consumer who has a positive perception of the firm's CSR initiatives tend to have favorable expectations and beliefs that it would not ever act unethically and vice versa (Öberseder, Schlegelmilch, & Murphy, 2013). Several studies provided insights that customer's perceived CSR not only affected their expectation but also influenced their behavioral consequences. Previous research showed a positive relationship between perceived CSR and purchase intention (Carvalho et al., 2010; Grimmer & Bingham, 2013; Tian et al., 2011). Therefore, this study describes the concept of perceived CSR as the extent to which customers acknowledge and recognize that a company supports activities related to social responsibilities. Hence, we developed the second hypothesis to test the impact of customer's perceived CSR on expectation toward CSR and purchase intention.

Hypothesis 2a: Customer's perceived CSR positively impacts on their CSR expectation.
Hypothesis 2b: Customer's perceived CSR positively impacts on their intention to purchase CSR-related products.

Moral obligation

In the literature, the concept of moral obligation is commonly interchanged by either ethical concerns or personal norms. It refers to intrinsic ethical motives driven by each person, reflecting an individual's self-concept about doing what is good/right and what is bad/wrong (S. H. Schwartz, 1977; Shaw, Shiu, & Clarke, 2000). There are two likely basics for moral obligation depending on the morality of the act itself and the goodness of the action consequences (Prichard, 2002). These moral values belong to the personality system and orientate ethical beliefs and attitude that embed into actions (Crowe & Simon, 2000; Fritzsche, 1995). Thus, consumers in the different self-concept of moral obligations have differences in the perception of ethical behaviors.

Several ethical consumption behavior studies found that customers perceived a strong sense of moral norms when making decisions inherently to ethical issues (Bommer et al., 1987; Shaw & Shiu, 2003). Consumers were interested in CSR-related products or responsible firms since they believed they ought to be ethical consumers and "do the right thing." Additionally, moral obligation embraced a significant predictive power of consumer's expectation and patronage behavior (Dickson, 2000; Gorsuch & Ortberg, 1983; Sparks, Shepherd, & Frewer, 1995). Therefore, within the present modified TRA model examining the relationship between CSR initiatives and ethical consumer's patronage intention, studying the impact of moral obligation was imperative, because CSR initiatives were highly embroiled with ethical issues (M. S. Schwartz & Carroll, 2008). Within the context of social responsibility, adapted from prior studies (Arvola et al., 2008; Beck & Ajzen, 1991) the concept of moral obligation in this paper is understood as a consumer's feel of a sense of responsibility to act morally. Thereby, we developed the following hypotheses.

Hypothesis 3a: Moral obligation positively impacts on customer's intention to purchase CSR-related product.
Hypothesis 3b: Moral obligation positively impacts on customer's CSR expectation.

Price Sensitivity and its mediating effects

From perspectives of business-decision makers, pricing is not only one of the most critical strategic management concepts since it affects market demand; generates sources of optimal profits, but it is also a competitive weapon gaining customer's insights. From another perspective, price throughout the past decades is a vital preference most frequently considered by customers in the purchase decision-making process (Kotler & Armstrong, 2010). Different market segmentation (i.e., price-conscious segment) observe various levels of price sensitivity. It is noticeable that economic bargain hunters will consume less as the price increases while those with low price sensitivity will not have a concern as the prices increase. How consumers react according to the perception of price levels or changes can be explained by the term of price sensitivity (Goldsmith & Newell, 1997). Another conclusion about price sensitivity was drawn as the degree of buyers' cognitive reactions when they dealt with the increase in prices of a specific product or service (Monroe, 1973). When purchasing decisions place more on the price attribute, consumers will be sensitive with small pricing fluctuations (Tellis, 1988). In other word, price sensitivity is inconsistent as it can change a consumer's purchase intention (Munnukka, 2008).

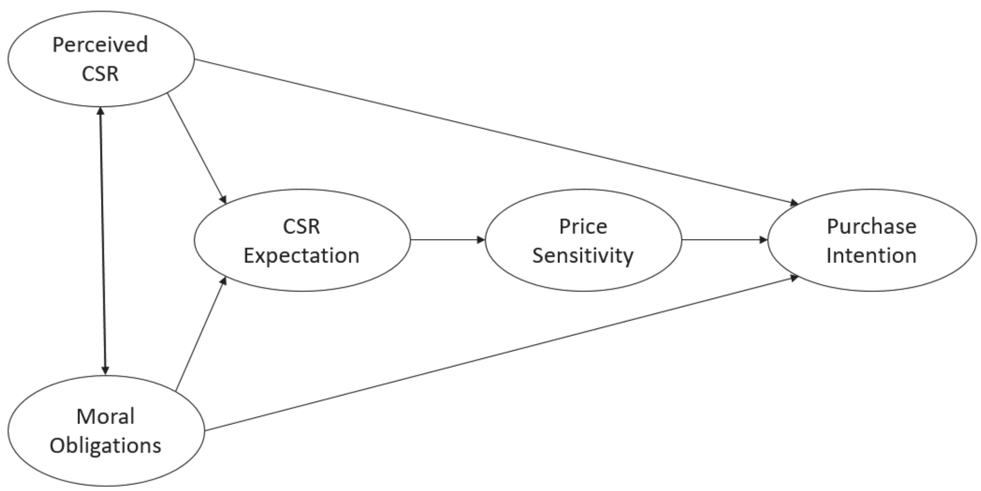
In the literature, evidence demonstrated that consumers were willing to pay higher prices for similar quality products perceived as social responsibility oriented (Becker-Olsen et al., 2006; Creyer, 1997; Öberseder et al., 2013; Rodrigues & Borges, 2015), but only when the products were ethically assured (McGoldrick & Freestone, 2008). For example, the presence of purchase intention when the CSR associated sneaker priced 10% higher was proven by a group of Brazilian scholars. They further confirmed that price fairness mediated the link between CSR perception and purchase intention significantly (Carvalho et al., 2010). As for consumers who are highly sensitive to prices, they might not buy socially responsible foods because they can immediately find cheaper ones at retail stores. Thus, we expect consumers' price sensitivities to vary directly with the extent to which they perceive and desire a company to be socially responsible. Thus, we suppose the likelihood of purchasing a more expensive product from a CSR company will be mediated by the consumer's degree of price consciousness.
Price sensitivity in this paper is defined as price differences that consumers willing (or accept) to pay at a given time. However, many studies investigating customer response to price escalation are still limited. Therefore, we measured the customer's price sensitivity based on a price climb as the same approach developed in previous studies (Gabor & Granger, 1979; Lichtenstein, Bloch, & Black, 1988; Monroe, 1973). We anticipated the well-established association between price sensitivity and purchase intention in the context of social responsibility. Particularly, the mediating effect of price sensitivity was firstly explored in the interrelationship among

consumers' CSR driven perception, moral obligations (ethical values, and patronage intention to CSR-related products in the marketplace. We have not been able to find any published studies that explored the potential mediating role of price sensitivity in the context of ethical consumption. Hence, the following hypotheses were constructed based on the above discussion.

Hypothesis 4: Price sensitivity significantly impacts on purchase intention toward CSR-related products.
Hypothesis 5a: Price sensitivity plays a mediating role in the positive relationship between Perceived CSR and Purchase Intention.
Hypothesis 5b: Price sensitivity plays a mediating role in the positive relationship between Moral Obligation and Purchase Intention.
Hypothesis 5c: Price sensitivity plays a mediating role in the positive relationship between CSR Expectation and Purchase Intention.

CONCEPTUAL MODEL

Figure 1: Research conceptual model



METHODOLOGY

Data collection and characteristics of respondents

This research approached respondents via paper-based surveys. We collected data from consumers shopping in supermarkets and shopping malls in Taiwan. Respondents were initially asked to look through an advertising poster introducing one fictitious dairy-food brand (its information, products, selling price, and its CSR profile). Then they were asked to fill in a questionnaire as a pool of 20 items (which was included detail in Table 3 showing the construct measurements of this study and four items of personal information related to age, gender, education, and monthly disposable income. Participants were asked to scale each item on a seven-point Likert scale (ranging from 1= strongly disagree to 7= strongly disagree). A total of 311 valid samples used for the data analysis process. The sampling process conducted without missing data.

The demographics of the population summarized in Table 1. Respondents were all Taiwanese and slightly more likely to be male (51.4%). Regarding age, the majority of the respondents were younger generations aged 21 to 30 years old (59.5%), followed by 31 to 40 years old (24.7%), over 40 (9.3%), and under 20 years old (5.8%). The statistics for the level of education found most have a Master's degree (45.7%) and 30.5% had a Bachelor's degree. We also surveyed disposable personal income of how much respondents spent per month, and the results showed the most significant percentage (~32%) of respondents had a spending range from over NTD 10,000 to 20,000 per month with the other most common range being from NTD 5,000 to NTD 10,000 (29.6%).

Table 1: Characteristics of Respondents

Demographical variables		Frequency	Percentage
Gender			
1	Male	160	51.4%
2	Female	151	48.6%
Age			
1	Under 20	18	5.8%
2	21-30	187	59.5%
3	31-40	56	24.7%
4	Over 40	29	9.3%
Level of education			
1	High school or below	73	23.5%
2	Bachelor's degree	95	30.5%
3	Master's degree	142	45.7%
4	PhD's degree and above	1	0.3%
Monthly disposable income (NT$)			
1	Below 5,000	29	9.3%
2	5,001 – 10,000	92	29.6%
3	10,001 – 20,000	99	31.8%
4	20,001 – 30,000	34	10.9%
5	30,001 – 40,000	30	9.6%
6	40,001-50,000	18	5.8%
7	Above 50,000	9	2.9%

Measurement of constructs

At the aim of testing the hypotheses, the study used the following five primary constructs: perceived CSR, CSR expectation, moral obligation, price sensitivity, and purchase intention toward CSR-related products. All of the items were tested by other scholars and employed in this study in an existing or slightly modified manner by a 7-point Likert scale (i.e., 1= strongly disagree; 7= strongly agree). *Perceived CSR* was defined as the extent to which customers acknowledged and recognized that a company supported activities related to social responsibilities. It was measured using four scale items adapted from (Brown & Dacin, 1997). *CSR expectation* demonstrated that consumers desired and wanted every firm to behave ethically in complying with social responsibilities. It was evaluated by four scale items originated from (Creyer, 1997). The concept of *Moral obligation* in this paper was understood as a consumer feel of a sense of responsibility to act morally. Three questions were adapted from

(Arvola et al., 2008). The *purchase intention* construct was measured by four questions adapted from (Jamieson & Bass, 1989) and (Carvalho et al., 2010). Merely purchase intention was customer willingness to buy CSR-related products. By identifying *price sensitivity* as price differences that consumers were willing (or accepting) to pay at a given time was an important aspect to study. We measured it by a 5-level pricing climb with price increases $NT 5 higher compared to the fixed price (present market price of sample product) as the same approach developed in previous studies (Gabor & Granger, 1979; Lichtenstein et al., 1988; Monroe, 1973). Participants were asked to answer whether they accepted to pay for sample product at different price levels or not, based on the 7-point Likert scale (i.e., 1= totally not accept; 7= totally accept).

DATA ANALYSIS & RESULTS

Reliability and validity test

A Confirmatory factor analysis (CFA) using AMOS 21 software was adopted to verify the underlying structure of constructs in the proposed hypotheses. Reflective measurement model tests were conducted to evaluate the reliability, convergent validity, and discriminant validity of the latent constructs between perceived CSR, moral obligation, CSR expectation, purchase intention, and price sensitivity.

Reliability of the constructs was estimated by composite reliability (CR) and Cronbach's alpha. All latent constructs achieved acceptable levels of reliability based on Cronbach's alpha that was higher than the recommended .70 level (Hair et al., 2010): perceived CSR (.92), moral obligation (.89), CSR expectation (.73), purchase intention (.88), and price sensitivity (.94) (see Table 3). Similarly, the composite reliability of the latent constructs ranged from .727 to .938 (see Table 2). These findings indicated that all constructs included in the proposed model were reliable.

Convergent validity indicated the extent to which measuring items of a scale that were theoretically related in reality and could be assessed by the standardized factor loading exceeded .400, was significant at .001, *and* the average extracted variances (AVE) were above .50 (Anderson & Gerbing, 1988). In two constructs, namely CSR expectation and moral obligation, to increase the model fit due to the complexity of the whole model, we deleted two items because of their unsatisfied factor loadings. As provided in Table 3, all of the factor loadings that remained in the research model ranged from .60 to .98, and all were statistically significant at p-value <.001 on its underlying construct. The average variance extracted AVE across the constructs mostly exceeded 0.5. The findings suggested these items explained a large portion of the variation and convergent validity was satisfied.

Discriminant validity (divergent validity tested whether constructs that were not supposed to be related were unrelated. The discriminant tests were satisfied if the correction coefficient among latent constructs was lower than the square root of AVE (Chin, 1998). The square root of AVE of the latent constructs ranged from .689 and .894, higher than correlation among latent variables (see Table 2). The findings indicated adequate discriminant validity.

Table 2: Validity, Reliability and Correlation Matrix

	CR	AVE	MSV	MaxR(H)	(1)	(2)	(3)	(4)	(5)
(1)CSR expectation	0.727	**0.745**	0.279	0.758	**0.689**				
(2)Perceived CSR	0.920	**0.742**	0.278	0.937	0.528	**0.862**			
(3)Price Sensitivity	0.938	**0.753**	0.754	0.977	-0.438	-0.429	**0.868**		
(4)Moral Obligation	0.888	**0.799**	0.104	0.981	0.313	0.102	-0.322	**0.894**	
(5)Purchase Intention	0.882	**0.661**	0.752	0.989	0.369	0.393	-0.687	0.288	**0.873**

Table 3: Constructs and Factor Analysis

Construct	Reference	Item	Description	Factor loading	Cronbach's alpha
Perceived CSR	(Brown & Dacin, 1997)	CSR1	Brand X demonstrates that it is concerned with the environment	0.85	**0.92**
		CSR2	Brand X is involved in the community and invests in worthwhile causes	0.85	
		CSR3	Brand X focus on employee benefit and provides an excellent working environment	0.86	
		CSR4	Brand X implements good corporate governance and information transparency	0.89	
Moral Obligation	(Arvola et al., 2008)	MO1	Buying milk from Brand X, I would feel like making a personal responsibility to something better	0.92	**0.89**
		MO2	Buying milk from Brand X, I would feel like the morally right thing	0.87	
		MO3	Buying milk from Brand X, it would not make me feel guilty	delete	
CSR Expectation	(Creyer, 1997)	EXP1	I expect firms have a responsibility not ever to act unethically	0.65	**0.73**
		EXP2	I hope firms have a responsibility to always act with the highest of ethical standards	0.81	
		EXP3	I expect the firms that I buy to act ethically at all times	0.60	
		EXP4	All firms will be unethically sometimes: it is normal	delete	

Purchase Intention	(Jamieson & Bass, 1989)	PI1	Buying a product of brand X will make you feel good about yourself?	0.61	**0.88**
		PI2	Buying a product of brand X will make you feel that you are doing the right thing?	0.65	
		PI3	Does Brand X have a high probability of you purchasing it?	0.98	
		PI4	Are you more willing to buy a product of brand X?	0.94	
Price Sensitivity		P1	I will buy a CSR product if its price at a fixed price	0.91	**0.94**
		P2	I will buy a CSR product if its price at $NT 5 more	0.95	
		P3	I will buy a CSR product if its price at $NT 10 more	0.95	
		P4	I will buy a CSR product if its price at $NT15 more	0.79	
		P5	I will buy a CSR product if its price at $NT 20 more	0.71	

Consequently, the goodness-of-fit result for our final measurement found good improvement in the overall model fit with some most important criteria: Chi-square/d.f. equals to 2.907 smaller than 3, RMSEA= 0.078, CFI = 0.935. The GFI in the model equals to 0.899, which is lower than value 0.90. However, according to (L.-T. Hu & P. M. Bentler, 1995) reported emerging evidence that 0.90 might not always be a reasonable cut-off for all adjunct fit indexes under all modelling circumstances. The choice of a cut-off value lower than 0.90, may be justified: "The rule of thumb to consider models acceptable if a fit index exceeds 0.90 is an inadequate rule" (L. Hu & P. Bentler, 1995, p. 95). We might delete more items to archive the better model fit. Nevertheless, we decided to keep these items to make a broader view of the whole model. However, the most important was that Chi-square/d.f., the CFI and RMSEA provided sufficient and unique information to evaluate the model (Hair, Black, Babin, Anderson, & Tatham, 2010, p. 654).

Testing the proposed research hypotheses

Following the CFA test to attain the best-fitting model, we applied the result in the CFA model to the hypothesized structural equation model. According to (Rowley & Berman, 2000) and (Galbreath & Shum, 2012), the structural equation model(SEM) is much more appropriate than traditional regression analysis in CSR studies. To examine the theoretical model indicated that the t-value of all completely standardized coefficients was statistically significant from 0.001 to 0.05 levels. As can be seen in the below figure (table 4), the results of estimating the basic framework (total effect model) revealed that consumer's perceived CSR, CSR expectation, and moral obligation had a significantly positive relationship with purchase intention of CSR-related products. These results are essential for testing the mediating effect of price sensitivity on the links between purchase intention and its antecedences. The fitness indicators found acceptable model fits with GFI=.946, CFI=.976, and RMSEA=.061.

The hypothesized positive relationship between CSR expectation and purchase intention (**Hypothesis 1a**) was supported (β=.135, t =1.965, $p < .05$). **Hypothesis H2a**, which predicted a positive relationship between perceived CSR and purchase intention was also supported (β= .304, t = 4.542, $p < .001$). Perceived CSR also found a statistically significant positive influence on CSR expectation (β= .488, t = 7.936, p < .001) providing support for **Hypothesis 2b**. As predicted by hypotheses **Hypotheses 3a** and **3b**, moral obligation significantly impacted both purchase intention (β= .217, t =3.717, p < .001) and CSR expectation (β= .258, t =4.352, p < .001).

Table 4: Structural Equation Model Results

Hypotheses		Standardized coefficient	(S.E)	t-value	Model fit measurement
Direct effect model (without mediator)					
H2a	Perceived CSR→ Purchase Intention	.304***	0.145	4.542	Ch-square/df = 2.169; p = .000; GFI=.946; AGFI=.914; NFI=.957; CFI=.976; RMR=.081; RMSEA=.061
H2b	Perceived CSR→CSR Expectation	.488***	0.052	7.936	
H3a	Moral Obligation → Purchase Intention	.217***	0.126	3.717	
H3b	Moral Obligation → CSR Expectation	.258***	0.05	4.352	
H1a	CSR Expectation → Purchase Intention	.135*	0.195	1.765	
Indirect effect model					
H2b	Perceived CSR → CSR Expectation	.502***	0.052	7.836	Ch-square/df = 2.846; p = .000; GFI=.899; AGFI=.862; NFI=.935; CFI=.957; RMR=.104; RMSEA=.077
H2c	Perceived CSR → Price Sensitivity	-.295***	0.151	-4.486	
H3b	Moral Obligation → CSR Expectation	.261***	0.049	4.161	
H3c	Moral Obligation → Price Sensitivity	-.227***	0.126	-3.944	
H1b	CSR Expectation → Price Sensitivity	-.211**	0.216	-2.743	
H4	Price Sensitivity → Purchase Intention	-.867***	0.033	-24.794	
Total effect model					
H2a	Perceived CSR → Purchase Intention	.040*	0.089	0.983	Ch-square/df = 2.907; p = .000; GFI=.899; AGFI=.859; NFI=.935; CFI=.956; RMR=.103; RMSEA=.078
H2b	Perceived CSR → CSR Expectation	.502***	0.052	7.847	
H2c	Perceived CSR → Price Sensitivity	-.294***	0.151	-4.48	
H3a	Moral Obligation → Purchase Intention	.017*	0.074	0.491	
H3b	Moral Obligation → CSR Expectation	.262***	0.049	4.181	
H3c	Moral Obligation → Price Sensitivity	-.225***	0.126	-3.938	
H1a	CSR Expectation → Purchase Intention	-0.033	0.125	-0.707	
H1b	CSR Expectation → Price Sensitivity	-.212**	0.216	-2.757	
H4	Price Sensitivity → Purchase Intention	-.859***	0.039	-20.923	
Sobel's test		***z-test [a]***			***p-value***
H5a	Perceived CSR → Price Sensitivity → Purchase Intention	.253***(a=-.294 x b=-.859)			<0.0001
H5b	Moral Obligation → Price Sensitivity → Purchase Intention	.193***(a=-.225 x b=-.859)			<0.0001
H5c	CSR Expectation → Price Sensitivity → Purchase Intention	.028*** (a=-.033 x b=-.859)			<0.0001

*Note: ***p < .001, **p < .01, *p <0.05 and significant level at |t-value| >1.96*

[a] : The significance of mediator was examined by a Z-test which calculated using Sobel's approach. The z-test results were adopted by an interactive calculation tool via this link: http://quantpsy.org/sobel/sobel.htm

Mediating effect model testing

There are plenty of statistical methods to investigate the mediation effects, namely hierarchical regression and SEM approaches. In this study, to examine the mediation effect of price sensitivity, we first adopted the statistical method developed by (Preacher & Hayes, 2004). We used the PROCESS macros version 3.3 for SPSS provided by (Hayes, 2017) to test the significance of indirect effects of perceived CSR; moral obligation; CSR expectation through price sensitivity, on purchase intention. In all analyses, we employed a bias-corrected bootstrap approach with 5000 samples, with a level of 95% confidence interval (CI). The indirect effects of three predictors of purchase intention through price sensitivity can be seen in Table 5 as below.

Table 5: Testing the Significance of Indirect Effects through Price Sensitivity

Independent Variables	**Purchase Intention**			
	Indirect effect	SE	LLCI	ULCI
Perceived CSR	.4404	.0695	.3141	.5872
Moral Obligation	.3349	.0599	.2207	.4573
CSR Expectation	.5467	.0809	.3972	.7108

Note: All items are statistically significant with p < .05, as the lower limit of the confidence interval (LLCI) and upper limit of the confidence interval (ULCI), do not cross zero

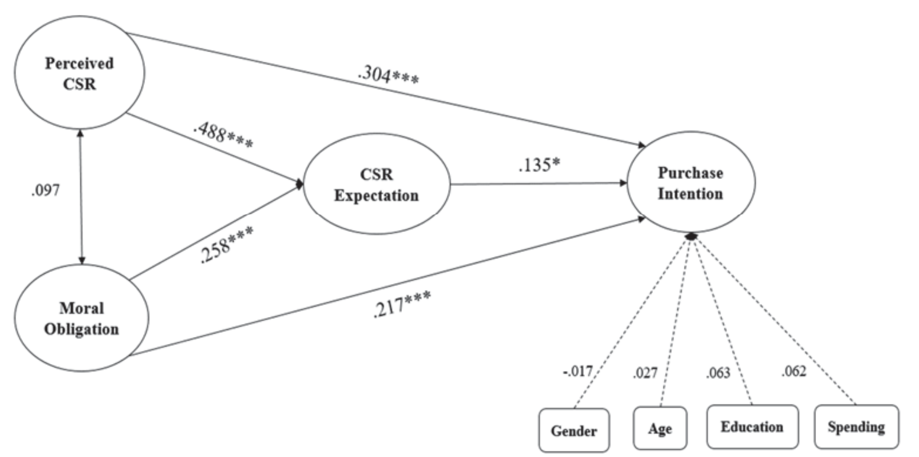
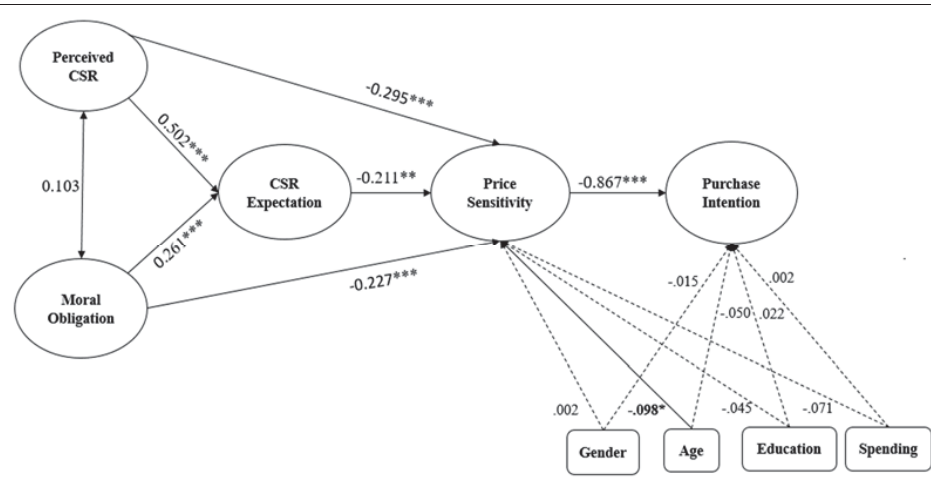
Besides, we also applied the SEM method following the procedure of (Baron & Kenny, 1986) and Sobel test (Sobel, 1982) to further confirm the mediation effects of price sensitivity in the interrelationship among variables. We ran three models: (1) direct effect for examining the significant relationship between purchase intention and its antecedences when mediator's effect was not regarded; (2) indirect effect model for testing the considerable association of dependent variable and its predictors with mediator (price sensitivity); and (3) total model for testing both direct and indirect effects.

As shown in Table 4, the z-test for Perceived CSR→ price sensitivity→ purchase intention (β= .253, z-test= 4.381, significant level p< .0001) showed that the mediating effect of price sensitivity on the relationship between perceived CSR and purchase intention toward CSR related product was confirmed in **Hypothesis 5a**. Moral obligation→ price sensitivity→ purchase intention (β= .193, z-test= 3.870, significant level p< .001), which indicated that the mediating role of price sensitivity on the relationship between moral obligation and purchase intention toward CSR related product was significant. Therefore, **Hypothesis 5b** was supported. Furthermore, the z-score for CSR expectation→ price sensitivity→ purchase intention (β= .028, z-test= 2.730, significant level p< .01), which indicated that the mediating effects of price sensitivity on the relationship between CSR expectation and purchase intention toward CSR related product were significant. Therefore, **Hypothesis 5c** was supported.

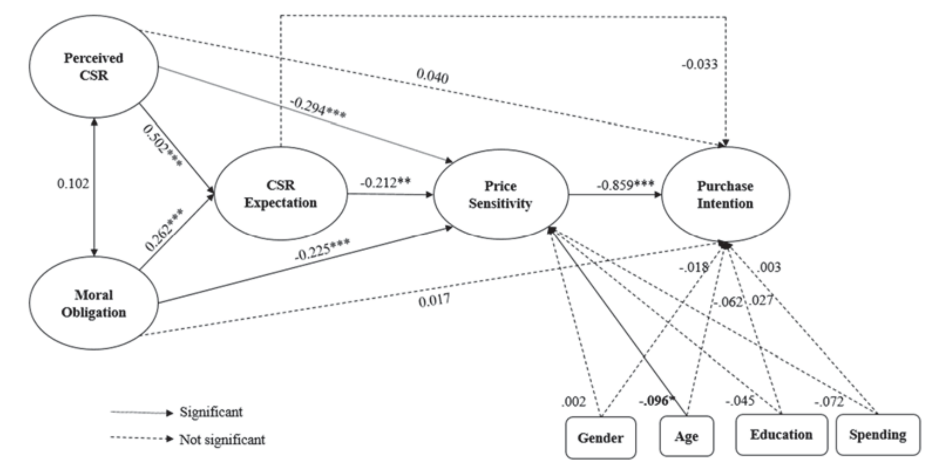
Additionally, as noted from the total effect model's results; with the intervention caused by price sensitivity, independent variables (perceived CSR, moral obligation and CSR expectation) no longer had significant impacts on purchase intention (all standardized coefficients turn to insignificant with t-value < 1.96). Based on suggestions by (Baron & Kenny, 1986), it was concluded from results that price sensitivity was a complete mediator when it fully mediated the relationship between purchase intention and its antecedences in the model (see figure 2).

Control variables

Regarding the control variables, we included gender, age, level of education, and disposable personal income to examine its effects on purchase intention and price sensitivity. It was shown that none had a significant effect on purchase intention, and these results were similar to previous studies in which the non-linear relationship between consumer demographics and CSR responses was proven (Tian et al., 2011). However, the population age had a significantly negative impact on price sensitivity (β= -.098; t= -2.126, $p < 0.05$).

Figure 2: Graphical figure of tested models	
Direct effect	
Indirect effect	

Total effect



DISCUSSION & CONCLUSION

As a primary stakeholder, consumer-based responses to CSR initiatives are prudent to provide both the academic world and businesses with contributions from the current study. An enterprise that has a tremendous performance of CSR will build a good corporate image in the minds of consumers to obtain trust and ultimately, consumer willingness to buy the product; thus enhancing the competitiveness of enterprises. Firstly, this study modified the TRA model based on the descriptive approach that concentrated on impacts of CSR-driven perceptions (perceived CSR; CSR expectation) and moral obligations (internal ethics) as personal motives driving consumer patronage intention towards CSR-related products. The effects of CSR initiatives practiced by firms on consumers' purchase intentions was reconfirmed in this paper as the same conclusions were previously found in several studies (Mohr et al., 2001; Öberseder et al., 2013; Rodrigues & Borges, 2015; Tian et al., 2011). Also, it should be emphasized that consumer's perceptions to CSR initiatives were examined comprehensively on multiple dimensions including environmental dimension, social dimension, economic dimension, stakeholder dimension, and voluntariness dimension (Dahlsrud, 2008). Therefore, in business practices, CSR should be more widely implemented as strategic importance to every enterprise for sustainable growth. Consumer social concerns should be integrated into business operations, particularly in the food industry.

Secondly, previous research on corporate social responsibilities tended to ignore the significant role of individuals' intrinsic ethical motives. This study further contributed theoretical implications by identifying the role of consumer's self-concept of moral obligations in predicting consumers' behavioral consequences (i.e., purchase intention). Thus, authors suggest that when developing CSR policies for firms' marketing campaigns, marketing managers should regard CSR initiatives based on studying customer level of moral obligations. Further, consumers can perceive positive perceptions toward firms and simultaneously expect firms to engage in CSR activities, as well as make favorable patronage intention for their products and services.

Finally, yet importantly, this study also contributed to the fast consuming sectors (food industry) background by recognizing the effects of price sensitivity on the linkage between consumers' perceived CSR, CSR expectation, moral obligations, and intentions to buy CSR-related products. We found that price sensitivity fully mediated these associations referring to CSR responses from the perspectives of consumers. In literature, it seems hard to find any published studies that explore the potential mediating role of price sensitivity in the context of ethical consumption. We found that ethical consumers had a dilemma in making a purchase decision. They tend to be price sensitive in the current economic situation, particularly when it came to ethical products; namely, buyers could be ethical up to a certain level of price-consciousness.

Therefore, we expect that evidence of the complete mediation effect of price sensitivity explored in the current paper will contribute crucial references, not only for theoretical studies on consumer-based behavior but also for managerial practices. From perspectives of business-decision makers, pricing is one of the most critical strategic management activities since it affects market demand; generated sources of optimal profits, but also a competitive weapon gaining customers' insights. Meanwhile, price is ultimately a vital preference considered by customers. We also found that consumer's price sensitivity had considerable influence on altering behavioral intention in ethical consumption. They tended to be in an ethical dilemma and had to a trade-off between ethical concerns with price perceptions. Hence, the pricing strategy implemented by firms should be more concerned about customer's segmentation in term of price-consciousness. We believe that favorable pricing strategies and other promotional techniques can make consumers less price-focused and more concerned with core product benefits that align with ethics.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin, 103*(3), 411.

Arvola, A., Vassallo, M., Dean, M., Lampila, P., Saba, A., Lähteenmäki, L., & Shepherd, R. (2008). Predicting intentions to purchase organic food: The role of affective and moral attitudes in the Theory of Planned Behaviour. *Appetite, 50*(2-3), 443-454.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173.

Barone, M. J., Miyazaki, A. D., & Taylor, K. A. (2000). The influence of cause-related marketing on consumer choice: does one good turn deserve another? *Journal of the Academy of Marketing Science, 28*(2), 248-262.

Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of research in personality, 25*(3), 285-301.

Becker-Olsen, K. L., Cudmore, B. A., & Hill, R. P. (2006). The impact of perceived corporate social responsibility on consumer behavior. *Journal of Business Research, 59*(1), 46-53.

Berens, G., Riel, C. B. v., & Bruggen, G. H. v. (2005). Corporate associations and consumer product responses: The moderating role of corporate brand dominance. *Journal of Marketing, 69*(3), 35-48.

Bommer, M., Gratto, C., Gravander, J., & Tuttle, M. (1987). A behavioral model of ethical and unethical decision making. *Journal of Business Ethics, 6*(4), 265-280.

Brown, T. J., & Dacin, P. A. (1997). The company and the product: Corporate associations and consumer product responses. *Journal of Marketing, 61*(1), 68-84.

Carroll, A. B. (1999). Corporate social responsibility: Evolution of a definitional construct. *Business & society, 38*(3), 268-295.

Carvalho, S. W., Sen, S., de Oliveira Mota, M., & de Lima, R. C. (2010). Consumer reactions to CSR: A Brazilian perspective. *Journal of Business Ethics, 91*(2), 291-310.

Choi, S., & Ng, A. (2011). Environmental and economic dimensions of sustainability and price effects on consumer responses. *Journal of Business Ethics, 104*(2), 269-282.

Cooper-Martin, E., & Holbrook, M. B. (1993). Ethical consumption experiences and ethical space. *ACR North American Advances*.

Creyer, E. H. (1997). The influence of firm behavior on purchase intention: do consumers really care about business ethics? *Journal of consumer Marketing, 14*(6), 421-432.

Crowe, R., & Simon, W. (2000). *Who are the ethical consumers?* : Citeseer.

Cui, J., Jo, H., & Na, H. (2018). Does corporate social responsibility affect information asymmetry? *Journal of Business Ethics, 148*(3), 549-572.

Dahlsrud, A. (2008). How corporate social responsibility is defined: an analysis of 37 definitions. *Corporate social responsibility and environmental management, 15*(1), 1-13.

De Pelsmacker, P., Driesen, L., & Rayp, G. (2005). Do consumers care about ethics? Willingness to pay for fair-trade coffee. *Journal of Consumer affairs, 39*(2), 363-385.

Deng, D., & Jiang, K. (2011). Connotation of corporate social responsibility: A view from consumers. *China Soft Science, 10*, 0-11.

Devin, B., & Richards, C. (2016). Food waste, power, and corporate social responsibility in the Australian food supply chain. *Journal of Business Ethics*, 1-12.

Dickson, M. A. (2000). Personal values, beliefs, knowledge, and attitudes relating to intentions to purchase apparel from socially responsible businesses. *Clothing and Textiles Research Journal, 18*(1), 19-30.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*.

Fishbein, M. E. (1967). Readings in attitude theory and measurement.

Freeman, R. E., Harrison, J. S., Wicks, A. C., Parmar, B. L., & De Colle, S. (2010). *Stakeholder theory: The state of the art*: Cambridge University Press.

Fritzsche, D. J. (1995). Personal values: Potential keys to ethical decision making. *Journal of Business Ethics, 14*(11), 909-922.

Fukukawa, K. (2003). A theoretical review of business and consumer ethics research: Normative and descriptive approaches. *The Marketing Review, 3*(4), 381-401.

Gabor, A., & Granger, C. (1979). Price sensitivity of the consumer. *Management Decision, 17*(8), 569-575.

Galbreath, J., & Shum, P. (2012). Do customer satisfaction and reputation mediate the CSR–FP link? Evidence from Australia. *Australian Journal of Management, 37*(2), 211-229.

Ghosh, P., Dasgupta, D., & Ghosh, D. (1990). Retail management.

Goldsmith, R. E., & Newell, S. J. (1997). Innovativeness and price sensitivity: managerial, theoretical and methodological issues. *Journal of Product & Brand Management, 6*(3), 163-174.

Gorsuch, R. L., & Ortberg, J. (1983). Moral obligation and attitudes: Their relation to behavioral intentions. *Journal of Personality and Social Psychology, 44*(5), 1025.

Grimmer, M., & Bingham, T. (2013). Company environmental performance and consumer purchase intentions. *Journal of Business Research, 66*(10), 1945-1953.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis* (7th ed.): Pearson Prentice Hall, New Jersey.

Harrison, R. (2003). Corporate social responsibility and the consumer movement. *Consumer Policy Review, 13*(4), 127-127.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*: Guilford Publications.

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit.

Hu, L., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 76-99). London Sage.

Jamieson, L. F., & Bass, F. M. (1989). Adjusting stated intention measures to predict trial purchase of new products: A comparison of models and methods. *Journal of Marketing Research, 26*(3), 336-345.

Kim, Y. (2017). Consumer responses to the food industry's proactive and passive environmental CSR, factoring in price as CSR tradeoff. *Journal of Business Ethics, 140*(2), 307-321.

Kotler, P., & Armstrong, G. (2010). *Principles of marketing*: Pearson education.

Lichtenstein, D. R., Bloch, P. H., & Black, W. C. (1988). Correlates of price acceptability. *Journal of Consumer Research, 15*(2), 243-252.

Maloni, M. J., & Brown, M. E. (2006). Corporate social responsibility in the supply chain: an application in the food industry. *Journal of Business Ethics, 68*(1), 35-52.

McGoldrick, P. J., & Freestone, O. M. (2008). Ethical product premiums: antecedents and extent of consumers' willingness to pay. *The International Review of Retail, Distribution and Consumer Research, 18*(2), 185-201.

McGuire, J. B., Sundgren, A., & Schneeweis, T. (1988). Corporate social responsibility and firm financial performance. *Academy of Management Journal, 31*(4), 854-872.

McWilliams, A., Siegel, D. S., & Wright, P. M. (2006). Corporate social responsibility: Strategic implications. *Journal of Management Studies, 43*(1), 1-18.

Mohr, L. A., Webb, D. J., & Harris, K. E. (2001). Do consumers expect companies to be socially responsible? The impact of corporate social responsibility on buying behavior. *Journal of Consumer affairs, 35*(1), 45-72.

Monroe, K. B. (1973). Buyers' subjective perceptions of price. *Journal of Marketing Research*, 70-80.

Munnukka, J. (2008). Customers' purchase intentions as a reflection of price perception. *Journal of Product & Brand Management, 17*(3), 188-196.

Nebenzahl, I. D., Jaffe, E. D., & Kavak, B. (2001). Consumers' punishment and rewarding process via purchasing behavior. *Teaching Business Ethics, 5*(3), 283-305.

Öberseder, M., Schlegelmilch, B. B., & Murphy, P. E. (2013). CSR practices and consumer perceptions. *Journal of Business Research, 66*(10), 1839-1851.

Pérez, A., del Mar García de los Salmones, M., & Rodríguez del Bosque, I. (2013). The effect of corporate associations on consumer behaviour. *European Journal of Marketing, 47*(1/2), 218-238.

Perrini, F. (2006). SMEs and CSR theory: Evidence and implications from an Italian perspective. *Journal of Business Ethics, 67*(3), 305-316.

Porter, M., & Kramer, M. R. (2006). Strategy & Society. The link between competitive advantage and corporate social responsibility. Harward Business Review, December 2006, 1–14. In.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers, 36*(4), 717-731.

Prichard, H. A. (2002). What Is The Basis of Moral Obligation? In J. MacAdam (Ed.), *Moral Writings*: Clarendon Press.

Rodrigues, P., & Borges, A. P. (2015). Corporate social responsibility and its impact in consumer decision-making. *Social Responsibility Journal, 11*(4), 690-701.

Rowley, T., & Berman, S. (2000). A brand new brand of corporate social performance. *Business & society, 39*(4), 397-418.

Schwartz, M. S., & Carroll, A. B. (2008). Integrating and unifying competing and complementary frameworks: The search for a common core in the business and society field. *Business & society, 47*(2), 148-186.

Schwartz, S. H. (1977). Normative influences on altruism. In *Advances in experimental social psychology* (Vol. 10, pp. 221-279): Elsevier.

Sen, S., Bhattacharya, C. B., & Korschun, D. (2006). The role of corporate social responsibility in strengthening multiple stakeholder relationships: A field experiment. *Journal of the Academy of Marketing Science, 34*(2), 158-166.

Shaw, D., & Shiu, E. (2002). The role of ethical obligation and self-identity in ethical consumer choice. *International Journal of Consumer Studies, 26*(2), 109-116.

Shaw, D., & Shiu, E. (2003). Ethics in consumer choice: a multivariate modelling approach. *European Journal of Marketing, 37*(10), 1485-1498.

Shaw, D., Shiu, E., & Clarke, I. (2000). The contribution of ethical obligation and self-identity to the theory of planned behaviour: An exploration of ethical consumers. *Journal of marketing management, 16*(8), 879-894.

Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*(3), 325-343.

Shrivastava, P. (1995). Environmental technologies and competitive advantage. *Strategic Management Journal, 16*(S1), 183-200.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology, 13*, 290-312.

Sparks, P., Shepherd, R., & Frewer, L. J. (1995). Assessing and structuring attitudes toward the use of gene technology in food production: The role of perceived ethical obligation. *Basic and applied social psychology, 16*(3), 267-285.

Tellis, G. J. (1988). The price elasticity of selective demand: A meta-analysis of econometric models of sales. *Journal of Marketing Research, 25*(4), 331-341.

Tian, Z., Wang, R., & Yang, W. (2011). Consumer responses to corporate social responsibility (CSR) in China. *Journal of Business Ethics, 101*(2), 197-212.

Valentine, S., & Fleischman, G. (2008). Ethics programs, perceived corporate social responsibility and job satisfaction. *Journal of Business Ethics, 77*(2), 159-172.

Wei, W., Kim, G., Miao, L., Behnke, C., & Almanza, B. (2018). Consumer inferences of corporate social responsibility (CSR) claims on packaged foods. *Journal of Business Research, 83*, 186-201.

Wilson, D. T., Mathews, H. L., & Harvey, J. W. (1975). An empirical test of the Fishbein behavioral intention model. *Journal of Consumer Research, 1*(4), 39-48.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science, 21*(1), 1-12.

DECISION SCIENCE INSTITUTE

Scheduling Impact on Student Online Learning Performance: A Case Study

San Yun (Nancy) Tsai
California State University, Sacramento
Email: sytsai@csus.edu

ABSTRACT

This research study investigates course schedule scheme impacts on student learning effectiveness in a high enrollment undergraduate computer literacy course over three consecutive semesters. Learning outcome and learning behavior are two elements included to evaluate learning effectiveness. Average course grade (exam) measures learning outcome of each schedule. Passing rate and failing rate evaluate learning behavior of each semester. Descriptive discussion and statistical hypothesis testing reveal that the course schedule does have significant impacts on student learning effectiveness under a teacher-pushed online learning approach.

KEYWORDS: Online learning, Hybrid-blended online learning, Course scheduling, Learning effectiveness

INTRODUCTION

The popularity of daily Internet usage paired with inexpensive information and communication technologies have created a new enhanced virtual environment for a variety of online teaching/learning approaches in the current digital age. Every approach regardless of its format can offer two major benefits to higher education institutions. First, it provides a tool to communicate and educate unlimited number of students without time and location constraints in the universe. Second, it is cost effective by reducing instructional, physical, and overhead expenses. Consequently, almost every higher education institution has embraced these virtual approaches and permanently incorporated them into its existing traditional teacher-pushed classroom learning.

Three major teaching/learning approaches have been adopted at the college level education systems over the past century. These include the traditional teacher-pushed classroom learning, student-pulled online learning and hybrid-blended (teacher-guided) online learning. The traditional teacher-pushed classroom learning can be described as a teacher utilizing the textbook, reading assignment, lecturing, homework exercise and test to guide and evaluate student learning with a specific schedule in a classroom. The student-pulled online learning can be defined as students having total control of their studying pace, when to study, and where to study via an electronic courseware and Internet without any learning push and/or guidance from a teacher. The hybrid-blended (teacher-guided) online learning is a combination of the first two approaches where a teacher adopts some learning rules to direct, stimulate, and influence student online learning process via an electronic courseware and Internet (Eom, Ashill, & Wen, 2006; Smith, 2001).

Without exception, College of Business Administration (CBA) located in northern California has to utilize information and communication technologies for developing various learning effective and cost saving online courses to ease pressures from its funding reduction and limited educational resources. This prompts to move a high enrollment (around 1,400 students per semester) undergraduate computer literacy course (MIS 123) from a teacher-pushed classroom learning to an online student-pulled learning one. Unfortunately, low student learning performance is a major pitfall of this course teaching/learning transformation. It becomes an important and on-going mission for course instructor to renovate MIS 123 teaching strategy and learning approach that will benefit students and facilitate their learning performance.

MIS 123 teaching/learning approach has then been adopted a teacher-guided (hybrid-blended) online learning from student-pulled one (Guay, Morin, Litalien, Valois, & Vallerand, 2015; Klein, Noe, & Wang, 2006; Ratelle, Guay, Vallerand, Larose, & Senecal, 2007; Rienties, Tempelaar, Van den Bossche, Gijselaers & Segers, 2009; Simmering, Posey, & Piccoli, 2009; Wyss, Lee, Domina, & MacGillivray, 2014).

This empirical study investigates impacts of different schedule schemes on MIS 123 student learning performance including learning outcome and learning behavior in three consecutive semesters. Average course grade measures student learning outcome of a specific learning schedule. Passing rate and failing rate measure student learning behavior for each learning schedule. Descriptive discussion and statistical hypothesis testing reveal that course schedule does have a significant impact on student learning performance under the teacher-pushed online teaching/learning approach.

COMPUTER LITERACY COURSE

MIS 123 is open to every student enrolled in the university. It is a prerequisite requirement for every undergraduate student before declaring a concentration of CBA. Major objective of this course is providing students with knowledge, skill and ability to apply microcomputer technology in other courses and their professional career. Only one instructor is responsible for establishing course material, managing course activity, and directing student learning process.

The course is separated into three one unit, credit/no credit sub courses: MIS 1 Word Processing, MIS 2 Spreadsheets, and MIS 3 Presentation Graphics. The upper enrollment limit is 500 students for each sub course in a semester. A student survey has identified several sub course characteristics. MIS 1 and MIS 3 are the two easier sub courses and the most useful in taking other courses. MIS 2 is the hardest one and most useful in professional career.

No classroom meeting is scheduled for MIS 123. A detailed course syllabus is posted on Canvas sub course web site for students to obtain a set of step by step “how to” information. That includes (1) logging on course web site; (2) getting courseware technical support for any hardware and software issue; (3) learning and mastering lessons; (4) taking exams; (5) obtaining lesson and exam schedule; (6) viewing performance records; (7) finding course grade; and (8) communicating with instructor. Students receives an email with a copy of course syllabus before the first day of a sub course. In addition, a posted announcement will be given to students via an email too.

There are two approaches for students to solve any learning problem during a semester. First, students can have a face to face discussion with instructor during 3 weekly office hours or by a special appointment. Second, students can seek a solution from instructor via two email systems including CBA and Canvas without time and place restrictions. Most students (99%) prefers second approach to acquire supports and helps.

Each sub course adopts different portion of a courseware developed and supported by a reputable publisher. Both MIS 1 and MIS 3 have 8 lessons. There are 12 lessons included in MIS 2. Sub course materials are posted on Canvas sub course web site 24x7 that provides students with unlimited access during a semester. Courseware offers four lesson learning methods to students that includes e-text (reading), audio PowerPoint presentation (listening), videos (watching), and simulation training (practicing with learning aids). Each lesson comes with an exam that students must complete it for earning sub course credit. Students can pick any learning method that fits into his/her learning preference to master lessons. MIS 123 students have total control of their learning process in terms of when, where, how fast and how long to take lessons and exams within a sub course schedule. The instructor does not provide any learning guidance to students except information as stated in sub course syllabus. Therefore, it is a student-pulled online learning approach.

Unfortunately, most MIS 123 students lack self-motivation and high-discipline required to complete online lessons and submit exams on time for each sub course. This fact produces a poor performance record with low passing rate and high failing rate for every sub course. One major responsibility of instructor is to renovate each sub course for improving student learning effectiveness. It is a necessary and urgent task for instructor to incorporate some educational motivation strategies and/or implement an appropriate schedule scheme for the renovation. Consequently, this active instructor involvement in student online learning has advanced MIS 123 teaching/learning approach from a student-pulled to a teacher-guided one.

STUDY

MIS 123 implemented three different course schedule schemes (1, 2 and 3) to define allowable learning period for students over three consecutive semesters. Students took same courseware lessons and exams for acquiring knowledge and skill of a sub course in every semester. Each scheme had its own lesson studying schedule and exam submitting schedule. Students had unlimited attempts to complete and submit an exam at any time before specified deadline. The courseware will record the highest score for that exam. Students had another chance to make up and submit any missing exam during the reopen period that was last three days before the end of a sub course. The intention of second chance was to eliminate any make up request and to increase learning effectiveness. The passing course grade was 70 or higher of average exam score. These exam rules were clearly stated in sub course syllabus. Weekly Canvas announcement with an email to students served as a reminder to inform them about coming exam(s) due date. Differences among these three schedule schemes were the number of sub courses offered, offering period of each sub course, and length of offering period as described in followings.

Schedule scheme 1 allowed students to take three sub courses (MIS 1, MIS 2, and MIS 3) simultaneously from week 1 to week 15 of first semester. Students were required to submit a MIS 1 exam and a MIS 3 exam every 13 day, and submit a MIS 2 exam every 9 days. Students had the most studying time to complete two or three exams of three sub courses each week during the entire semester.

Schedule scheme 2 allowed students to take two sub courses simultaneously during the second semester. MIS 1 was scheduled from week 1 to week 8; MIS 2 was scheduled from week 3 to week 14; and MIS 3 was scheduled from week 9 to week 16 (including final exam week). MIS 2 offering time overlapped with MIS 1 or MIS 3. Students required to submit 2 exams (MIS 1 and MIS 2, or MIS 2 and MIS 3) between week 3 and week 14. Students had either a week to master one lesson and submit one exam of one sub course, or one lesson and one exam each of two sub courses.

Schedule scheme 3 allowed students to take one sub course every 5 weeks of the third semester. MIS 1 was offered during the first five weeks between week 1 and week 5. MIS 2 took the middle five weeks from week 6 and week 10. MIS 3 had the last five weeks between week 11 and week 15. Students had to submit a MIS 1 exam or MIS 3 exam every 5 days and submit a MIS 2 exam every 3 days. Students had the least studying time, but could concentrate on only one sub course material within its offering period.

The average course grade (exam), number of students receiving "credit" (passing), and number of students receiving "no credit" (failing) were gathered for every schedule scheme of each sub course. These data were the basis to measure a specific schedule scheme impact on student learning effectiveness.

ASSESSMENT OF STUDENT LEARNING PERFORMANCE

Student learning effectiveness can be evaluated by student learning outcome and student learning behavior. The average course grade measures student learning outcome. A two tails hypothesis testing with 99% (2.58 Z value) and 95% (1.96 Z value) confidence levels evaluates student learning outcome change between a pair of two schedule schemes for every sub course. The number of students receiving "credit" and "no credit" describes and assesses student learning behavior change among three schedule schemes of each sub course.

MIS 1 Word Processing: Easy Subject and Useful Skill for Other Courses

Tables 1 presents statistical summary of MIS 1 learning performance. Schedule scheme 3 has the highest average course grade with the least standard deviation. Students can achieve 17 to 7 points more of average course grade and 6 to 8 points less of standard deviation as comparing with schedule schemes 2 and 1. These results show that students will improve learning performance when there is only one sub course material to study with the shortest learning period of 5 weeks long. Schedule scheme 2 produces the worst learning performance. Its average course grade is 10 to 17 points less than other two schedule schemes 1 and 3. Students cannot handle to master a lesson and submit an exam each of two sub courses within a week. On the other hand, students can do better to comprehend a lesson and submit an exam each for three sub course if the learning period is longer than a week such as one and half or two weeks defined in schedule scheme 1.

Table 1: MIS 1 Learning Performance

Schedule Scheme	Enrollment	No. of Credit	No. of No Credit	Average Course Grade	Standard Deviation
Schedule 1	426	353	73	84.97	23.22
Schedule 2	476	433	43	74.99	20.93
Schedule 3	470	426	44	91.64	15.00

Hypothesis 1: There is no statistical significant difference of MIS 1 learning outcome as

measured by average course grade between a pair of two schedule schemes.

Hypothesis testing result: Reject no difference hypothesis at both 99% and 95% confidence levels for every pair of two sub courses as illustrated in Table 2.

Hypothesis testing finding: Schedule scheme 3 with offering one sub course every 5 weeks is the most suitable schedule for students to achieve the highest learning outcome. Schedule scheme 1 with offering three sub courses concurrently during the entire semester is significantly a better one than schedule scheme 2 with offering two sub courses every week within a semester for obtaining a higher learning outcome. There is a positive relationship trend between learning outcome, number of sub courses, and allowable study time in MIS 1.

Table 2: MIS 1 Z Value of Learning Outcome	
Schedule Scheme Pair	Z - Value
Schedule 1 vs 2	6.75
Schedule 1 vs 3	5.05
Schedule 2 vs 3	14.08

Learning behavior discussion: Table 3 presents the passing rate and failing rate of three schedule schemes. It is interesting to note that schedule scheme 2 with the lowest learning outcome produces the same passing rate (91%) and failing rate (9%) as schedule scheme 3 with the highest learning outcome. Schedule scheme 1 is not a suitable strategy for improving learning behavior since it generates the least passing rate (83%) and the highest failing rate (17%) as comparing with other two schedule schemes 2 and 3. Schedule scheme 3 is the most suitable scheduling strategy for MIS 1 students due to it can generate the highest learning effective impacts on learning performance as measured by learning outcome and learning behavior.

Table 3: MIS 1 Passing Rate and Failing Rate		
Schedule Scheme	Passing Rate	Failing Rate
Schedule 1	83%	17%
Schedule 2	91%	9%
Schedule 3	91%	9%

MIS 2 Spreadsheets: Tough Subject and Useful Skill for Professional Work Place

Tables 4 presents statistical summary of MIS 2 learning performance. Schedule scheme 3 has the highest average course grade with the least standard deviation. Students can achieve 8 or 7 points more of average course grade and 6 to 5 points less of standard deviation as comparing with schedule schemes 1 and 2. Students can do well if they only need to study one sub course material with a shortest 5 weeks learning period. There exists a 1 point difference in average course grade and its standard deviation between schedule schemes 1 and 2. Learning outcome does not change that much if students have to master a lesson each and submit an exam each of two or three sub courses in one or two weeks.

Table 4: MIS 2 Learning Performance

Schedule Scheme	Enrollment	No. of Credit	No. of No Credit	Average Point	Standard Deviation
Schedule 1	469	330	139	77.40	27.60
Schedule 2	483	342	141	78.27	26.49
Schedule 3	496	408	88	85.36	21.08

Hypothesis 2: There is no statistical significant difference of MIS 2 learning outcome as measured by average course grade between a pair of schedule schemes.

Hypothesis testing result: Accept no difference hypothesis at both 99% and 95% confidence levels for the pair of schedule schemes 1 and 2. Reject the no difference hypothesis at both 99% and 95% confidence levels for the pair of schedule schemes 1 and 3, and for the pair of schedule schemes 2 and 3 as illustrated in Table 5.

Table 5: MIS 2 Z Value of Learning Outcome

Schedule Scheme Pair	Z - Value
Schedule 1 vs 2	0.50
Schedule 1 vs 3	5.01
Schedule 2 vs 3	4.63

Hypothesis testing finding: Schedule scheme 3 with only offering MIS 2 from week 6 to week 10 is the most suitable schedule strategy to produce the highest learning outcome. Schedule schemes 1 and 2 with offering three or two sub courses concurrently during the semester result a no difference impact on learning outcome.

Learning behavior discussion: Schedule schemes 1 and 2 have achieved a very close two passing rates (70% or 71%) and two failing rates (30% or 29%) as illustrated in Table 6. These two schedule schemes have a same effect on MIS 2 learning behavior. On the other hand, the schedule scheme 3 yields 12% to 11% higher passing rate and the 11% to 12% lower failing rate as comparing with schedule schemes 1 and 2. Therefore, schedule scheme 3 not only achieves the highest learning outcome as measured by average course grade (85.36), but also generates a better learning behavior as measured by passing rate (82%) and failing rate (18%). Schedule scheme 3 is a suitable scheduling strategy to result the most positive learning effectiveness impacts on MIS 2.

Table 6: MIS 2 Passing Rate and Failing Rate		
Schedule Scheme	Passing Rate	Failing Rate
Schedule 1	70%	30%
Schedule 2	71%	29%
Schedule 3	82%	18%

MIS 3 Presentation Graphics: Easy Subject and Useful Skill for Other Courses

Tables 7 presents statistical summary of MIS 3 learning performance. Schedule scheme 3 has the highest average course grade with the least standard deviation. Students can achieve 1 to 3 points more of average course grade and 1 to 3 points less of standard deviation as comparing with schedule schemes 1 and 2. Students can put their entire studying effort in one sub course and perform well on exams with a short 5 weeks learning period. Schedule scheme 2 produces the lowest learning outcome with a 2 to 3 points less average course grade and standard deviation than other two schedule schemes 1 and 3. In general, students achieve very similar average course grade and standard deviation with every schedule scheme.

Table 7: MIS 3 Learning Performance					
Schedule Scheme	Enrollment	No. of Credit	No. of No Credit	Average Point	Standard Deviation
Schedule 1	456	363	93	87.58	19.51
Schedule 2	459	346	113	85.11	21.29
Schedule 3	498	406	92	88.55	18.73

Hypothesis 3: There is no statistical significant difference of MIS 3 learning outcome as measured by the average course grade between a pair of schedule schemes.

Table 8: MIS 3 Z Value of Learning Outcome	
Schedule Scheme Pair	Z - Value
Schedule 1 vs 2	1.81
Schedule 1 vs 3	0.78
Schedule 2 vs 3	2.64

Hypothesis testing result: Accept no difference hypothesis at both 99% and 95% confidence levels for the pair of schedule schemes 1 and 2, and for the pair of schedule schemes 1 and 3. Reject no difference hypothesis at both 99% and 95% confidence levels for the pair of schedule schemes 2 and 3 as illustrated in Table 8.

Hypothesis testing finding: Schedule scheme 3 is slightly a better strategy as comparing with scheme 2. There is no difference impact on learning outcome either between schedule schemes 1 and 3 or between schedule schemes 1 and 2. There is no superior scheduling strategy to have a significant impact on MIS 3 learning outcome.

Learning behavior discussion: Passing rate and failing rate of each schedule scheme are shown in Table 9. Schedule scheme 1 is a litter better strategy than schedule scheme 2 since it has achieved 5% higher passing rate and 5% less failing rate than schedule scheme 2. However, schedule scheme 3 achieves a 2% to 7 % higher passing rate and a 2% to 7% lower failing rate as comparing with schedule schemes 1 and 2. Therefore, schedule scheme 3 is an appropriate scheduling strategy to improve MIS 3 learning effectiveness since it has the highest average course grade (88.55), the highest passing rate (82%), and the lowest failing rate (18%).

Table 9: MIS 3 Passing Rate and Failing Rate

Schedule Scheme	Passing Rate	Failing Rate
Schedule 1	80%	20%
Schedule 2	75%	25%
Schedule 3	82%	18%

CONCLUSION

Student online learning performance as measured by learning outcome (average course grade) and learning behavior (passing rate and failing rate) have sufficient statistical evidences to support that the schedule scheme 3 is the most effective and promising scheduling strategy under the teacher-pushed online learning approach. It outperforms other two schedule schemes in every sub course regardless the course material difficulty, lengthy, and usefulness. Its superior result can be contributed to several factors. It only requires students to focus on one sub course material without joggling their studying time between two or three course materials within every 5 weeks. It allows students to complete each lesson of only one sub course by following a clearly defined and short learning period. Its exam submission deadline guides and regulates their studying time and effort (evenly pacing the studying within the five weeks). The reopening exam at the last three days gives students second chance that increases passing rate and reduces the failing rate.

The current teacher-guided online learning approach via guidance, encouragement, motivation and optimal schedule scheme does positively receive a better student online learning performance. However, this computer literacy course needs other educational motivation strategies to further improve passing and failing rates, especially for MIS 2 Spreadsheets (the hardest materials) and MIS 3 Graphic Presentation (limited studying time at the end of a semester). Based on this empirical study finding and instructional experience, two strategies might enhance learning effectiveness of all sub courses. These includes granting students more studying time for a subject and changing grade system. Restructure MIS 123 by splitting MIS 123 into two courses that can give students more studying time for a subject. Combine two easier subjects (MIS 1 Word Processing and MIS 3 Presentation Graphics) in one course where each subject offers longer studying time (7 and half weeks as comparing current 5 weeks for a subject). Create another independent course for MIS 2 Spreadsheets that allows students to have 15 weeks studying time to comprehend and master a difficult subject. Make the grade system of these two new courses to a regular letter grades (A, B, C, D and F) instead of credit/no Credit grade.

REFERENCES

Eom, S., Ashill, N., & Wen, J. 2006. “The determinants of students’ perceived learning outcomes and satisfaction in university online education: An empirical investigation,” *Decision Sciences Journal of Innovative Education* (4:2), pp. 102-115.

Guay, F., Morin, A. J., Litalien, D., Valois, P., & Vallerand, R.J. 2015. “Application of exploratory structural equation modeling to evaluate the academic motivation scale,” *Journal of Experimental Education* (83:1), pp. 51-82.

Klein, H. J., Noe, R. A., & Wang, C. 2006. “Motivation to learn and course outcomes: The impact of delivery mode, learning goal orientation, and perceived barriers and enablers,” *Personnel Psychology* (59:3), pp. 665-702.

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S. & Senecal, C. 2007. “Autonomous, controlled, and amotivated types of academic motivation: a person-oriented analysis,” *Journal of Educational Psychology* (99:4), pp. 734-746.

Rienties, B., Tempelaar, D., Van dan Bossche, P., Gijselaers, W., & Segers, M. 2009. “The role of academic motivation in computer-supported collaborative learning,” *Computers in Human Behavior* (25:6), pp. 1195-1206.

Simmering, M. J., Posey, C., & Piccoli, G. 2009. “Computer self-efficacy and motivation to learn in a self-directed online course,” *Decision Sciences Journal of Innovative Education* (7:1), pp. 99-121.

Smith, P. A. 2001. “Understanding self-regulated learning and its implications for accounting educators and researchers,” *Issues in Accounting Education* (*16:4),* pp. 663-700.

Wyss, J., Lee, S., Domina T., & MacGillivray, M. 2014. “Cotton Island: students’ learning motivation using a virtual world,” *Decision Sciences Journal of Innovative Education* (12:3), pp. 219-232.

DECISION SCIENCES INSTITUTE
Segmental Reporting and Earnings Management: Evidence from UK

Bilal ElSalem
College of Business of Economics
Qatar University
belsalem@qu.edu.qa

Ghassan H. Mardini
College of Business of Economics
Qatar University
ghassan.mardini@qu.edu.qa

ABSTRACT
The results of the current investigation lead to concludes that the IFRS-8's management approach is useful to internal and external decision makers especially that it explains its relationship with real earnings practices. For example, the management approach of IFRS-8 enhanced the perceptions of the managers about reliable segmental information.

KEYWORDS: Segmental Reporting, IFRS-8, Earnings Management, UK

1. INTRODUCTION
International Financial Reporting Standards (IFRS) have played an important role in the development of Earnings Management practices since their inception (Jeanjean and Stolowy, 2008). While Earnings Management (EM) is simply a practice whereby a company produces results in such a manner as to ensure that it looks as if it is performing better than it actually is (Jeanjean and Stolowy, 2008). In essence, Healy and Wahlen (1999) have argued that the company aims to portray a more positive outlook, thus securing the faith of its shareholders, future investment and cash flows. Hence, the IFRS was set up to meet the requirements of accounting standards, of which Earnings Management is one of the target areas. Accordingly, accounting standards have continuously been amended since the inception of IFRS in April, 2001, which has given rise to interpretations in different countries that depend on their own specific governance and other criteria. It is this interpretation issue that this study aims to address. The main objective of the current study is to assess this development from an interpretive angle, and with specific attention to the UK's Non-Financial FTSE-100 companies. Mostly, banks and financial institutions are excluded from earnings management research, as their characteristics are fundamentally different to other industrial companies (Peasnell et al., 2000).
The adoption of segmental reporting standards Internationally (IFRS 8) and in the USA (SFAS 131) offers new paradigms for a company's financial position and performance. This is called the management approach (Van Tendeloo, and Vanstraelen, 2005). A management approach is a key characteristic of IFRS 8 (Crawford et al., 2012). It underlines the strategy relating to the reporting of segmental information (André et al., 2016) and outlines the reporting choices that might be made by the Chief Operating Decision Maker (CODM). In this approach, the CODM is given a significant role in determining the segmental information that is to be disclosed for a particular financial year under IFRS 8 (Mardini et al., 2012).
A number of arguments have surfaced, and these have concentrated on justifying the effect of the segmental reporting requirements on earning management. The first argument that is of concern is clearing the air in regard to the relationship between the new segmental reporting standard and earnings (smoothing) management. Capkun et al., (2012) attributed the provision

of flexibility in the selection of accounting treatments to addressing earnings smoothing. The second arguments provide that the increase in disclosure in relation to the operating segments' items in the financial reports following the adoption of IFRS 8 in 2009, which may improve earning management for Jordanian listed companies (Mardini et al., 2012), as well as the UK's FTSE-100 companies (Mardini and Ammar, 2018). Moreover, According to Ettredge et al., (2005), the increase in available information relating to more reporting with regard to the segments' items (SFAS 131) strengthens the association between returns and future earnings. The adoption of the new IFRS standards as a mandatory requirement in the European zone is a matter of concern in relation to what the effects of better financial reporting turns are (Capkun et al., 2008). More specifically, the IFRS 8 standard is emphasised, since it encompasses the scope of reporting on operating segments that, in turn, influences the financial performance of the companies. The adoption of IFRS 8 articulates more specificity and exclusivity with regard to the earnings management of the reportable and operating segments (Jeanjean, and Stolowy, 2008; Zéghal et al., 2011). According to Mardini et al., (2012), the principle focus of IFRS is to enhance the understanding of a company's position and performance in relation to internal reports that are reviewed by the chief decision makers.

The arguments above have promoted a platform from which we can present a study based on the interpretation of accounting standards through paying attentions to IFRS 8 from the perspectives of UK listed companies that are subject to a thorough evaluation of Earnings Management. There are many advantages to the promotion and implementation of accounting standards worldwide. IFRS hopes to ensure all companies will use only one format, which increases transparency whilst promoting accountability. If economies aim to develop and grow, whilst being supported by good governance, one simple accounting standard should ensure that they do not become a group. This therefore overrates their company earnings and thus reduces the cash flow in the long run – a practice that has contributes to distrust amongst shareholders and also to the reduction of economic growth. IFRS promotes hopes of gaining the trust of investors, since they rely on companies reporting the correct information, particularly information relating to future earnings, and hence returns on investment. In terms of Earnings Management, over-valued financial reporting has a significant effect on subsequent operating performance. Some of the Earnings Management practices are associated with future operating performances. This reporting practice may have an effect on the subsequent years' return on assets, which are usually contrary to the stakeholders' expectations, and this then has an impact on future earnings and cash flows. The motivation of this study is thus threefold: (i) The study will evaluate the extent to which IFRS 8 has contributed to changes in Earnings Management practices in the UK non-financial FTSE-100 listed companies. All financial firms, including banks, insurance companies, and other financial institutions, have been excluded from the initial sample utilised in earnings management research, as the environments of their financial reporting are different from those of other industrial companies, and their accrual processes differ fundamentally where the expectation models of normal accrual activity are not able to capture (Peasnell et al., 2000). This can be evaluated in two time periods before the post-implementation review of IFRS 8 (2013), and after the review (2014-2016);(ii) the results will enable us to assess whether the practices have played a role in the country's economic performances or prosperity, particularly in light of the fact that poor practices are harmful to these causes; (iii) the importance is underpinned by inspecting and assessing whether performance differences played a part in Earnings Management, both before and after the post-implementation review of IFRS 8. Assessments can be made on comparisons drawn in order to provide results that will answer the question as to whether or not the differences that exist are in terms of the implementation of the amended standard, or if IFRS 8 is simply open to interpretation that depends on the governance of that country.

These motivations demonstrate a need for further research that explores earning management and segmental reporting under IFRS 8 in a different context. The implications of this have the potential to contribute to existing literature and also to assist policy makers and practitioners. The premise of this paper takes a more comprehensive approach that aims to address the aforementioned research gaps. What is missing in the literature, and the problem that is being addressed by this research, are the post-implementation review of IFRS 8 intercepts earnings management practices. Instead of replicating what is available in the existing literature, the current study may expand the present understandng of the IFRS 8 post-implementation review's impact on earning managements within an international context. This research also deals with some of the contradictory evidence by examining this aspect in relation to the FTSE-100. Importantly, this paper contributes to society through offering clarifications that advance the understandings of academics, policy makers, and practitioners about why different contexts that are implemented through IFRS 8 show different implications. In other words, prior studies on earnings management that have a segmental reporting perspective tend to focus on US companies. Hence, investigating this issue in relation to UK FTSE-100 firms may yield different results.

The structure of the current paper is as follows: Section 2 offers a literature review, theoretical framework and the hypothesis development, while Section 3 illustrate the methodology. Section 4 outlines the results and discusses the findings. Finally, Section 5 concludes the paper and identifies some research opportunities for future investigation.

2. LITERATURE REVIEW, THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

2.1 Literature Review

A massive amount of research on the benefit of IFRS is being conducted. According to studies carried out by Gassen and Sellhorn (2006), Barth et al., (2008), and Hung and Subramanyam, (2007), there was much improvement in the quality of earning after the voluntary adoption of the IFRS standards . However, there exists a concern over the predicted improvements that were reported on the quality of the financial reports after the adaptation of the IFRS may lack sustainability (Street et al., 2000, 1999).The IFRS standards have a provision for the utilisation of private information and the execution of judgment. The resultant effect is that managers are granted a substantial amount of discretion. Research shows that the discretion is a double-edged sword (Watts and Zimmerman, 1986). Discretion promotes an inexpensive method of reporting regulation, and it allows corporate insiders to assimilate the financial reports with the aim of reflecting on the economic reality as well as producing the information that lies within a given organisation. On the other hand, the discretion can be put to use optimistically. For instance, reporting discretion might be utilised by corporate managers to safeguard certain economic performances that are obfuscated, whether targets or earnings, and to avoid the violation of the set agreements (Leuz, 2010). It is evident that the insiders have the upper hand on the information over the outsiders, and this makes it difficult to curb such behaviour. The degree to which the discretion is applied pivots on the national legal institutions (Ball et al., 2003) and on specific organisational traits (incentives on reporting and operating characteristics) (Burgstahler et al., 2006; Christensen et al., 2015). This means that accounting standards are in constant interaction with a variety of contextual factors, such as the law, commercial codes, and the social norms that exist in various contexts and in different ways (Fearnley and Sunder, 2012).

The current work examining the issues relating to the voluntary adoption of the IFRS makes it difficult to carry out an extrapolation, especially in cases where the adoption of the IFRS has been mandatory. In cases where organisations are coerced to employ and adopt the IFRS, there exist different compliance incentives which make it easy for opportunists to manipulate financial reports due to the existence of massive flexibility (Soderstrom and Sun, 2007). The various ways through which the institutional factors influence the quality of financial reporting after forcing the adoption of IFRS remain an open and interesting question. Following a discussion on the merits of implementing IFRS from a global perspective, Ball (2006) argues that one fundamental advantage is information that is of high quality, which will be instrumental in promoting informed decision

making among investors. According to Barth et al. (2008), Cai et al. (2008), Gassen and Sellhorn (2006), and Hung and Subramanyam (2007), pieces of evidence have been discovered that offer support for this among the adopters of the IFRS globally. However, research conducted by Jeanjean and Stolowy (2008), Van Tendeloo and Vanstraelen (2005) and Goncharov (2005), demonstrates that the application of the IFRS has no particular impact on quality, and, most specifically, on the management of earnings. An explanation of the mixed findings can be partially explained via the country's influence on specific factors such as the adoption of one accounting standard that will be applicable around the world. Nevertheless, this may not lead to an improvement in homogeneity in the financial reporting for various companies and countries, due to different factors, such as the countries' legal systems, incentives for financial reporting, and political systems that may affect the quality of accounting (Soderstrom and Sun, 2007). It is evident that high quality accounting standards are necessary for the achievement of financial, economic systems that are cost effective (Ball, 2001). Country specific factors, such as the separation of financial reporting for the interests of the public, taxation, the existence of a legal system that is independent, the corporate governance and ownership structures, and a strong accounting profession, play a part in the creation of a public disclosure environment that is efficient (Ball, 2001). La Porta et al. (2000) put forward excellent work that leads to the popularisation of the conception of how country specific factors have made an impact on corporate governance. In particular, the strength of protecting investors which, in most cases, is usually based on the legal structure and the origin of a given country, and that is explained within the various practices of corporate governance. Financial reporting is a mechanism of corporate governance and any existing differences on the protection of investors may result in different accounting practices across different nations, even when they all apply the same accounting standards. Leuz et al. (2003), investigated the various ways through which the levels of earning management differed in the different countries. He discovered that countries that had strong protection for investors had lower levels of earnings management. Nobes (2011) suggests that countries that have congruent legal backgrounds, or those that have a shared history, can display similar accounting practices. After he classified into two groups the accounting practices of fourteen states in 1983, he discovered that these groupings remained in the post-IFRS era. These findings suggested that country specific factors tend to influence the various practices in accounting in ways that the harmonisation of international accounting may find difficult to overcome. In essence; the alignment of *de jure* accounting may not result to the harmonization of *de facto* accounting.
Generally, these studies have argued that useful segmental disclosure increases the liquidity of the market, reduces the information gap between management and investors, and lowers the capital costs per firm. Specifically, for the current study's objectives, a number of prior studies (e.g., Botosan and Stanford, 2005; Ettredge et al., 2005; Herrmann and Thomas, 2000; Johnston, 2001) have investigated the economic consequences of segmental reporting, and they have found that segmental disclosure enhances the market value of a firm and its returns. They also found that segmental reporting enhanced the ability of the firm to forecast and assess the likely persistence of future earnings. With data for 172 companies, Behn et al. (2002) examined whether the segmental information disclosed did help the users of financial statements to understand the risks and growth prospects of their investee companies. They found that in excess of 50% of their sample companies provided accurate segmental information; hence, they concluded that segmental reporting enhanced the ability of analysts to forecast a company's future earnings. Ettredge et al. (2005) documented that the segmental information disclosed increased the usefulness of financial statement disclosures for forecasting information, and it enhanced the predictive ability in relation to business segment earnings[1].

[1] They used the Future Earnings Response Coefficient (FERC) methodology to assess this argument; the FERC was obtained by regressing current year share returns, against the next year's corporate annual earnings plus the control variables (which include earnings growth, earnings persistence, and the information environment).

2.2. Theoretical Framework

The request for corporate information can be raised by information asymmetry between managers (agents) and investors (principals), which has serious importance for the well-organised allocation of existing resources, and thus for the running of a competent capital market. Firms deliver information in numerous forms, such as financial reports, and this is mandatory and must be audited. Firms also provide information through voluntary channels, such as press releases and the company's website. Furthermore, financial intermediaries, such as financial analysts and experts, also offer information about firms. Nevertheless, considerable managerial discretion is allowed in relation to how the companies, in generating financial reports, apply accounting standards. Firms with different characteristics take on different accounting policies. (Watts and Zimmerman, 1986; Jensen and Meckling, 1976) IFRS 8 puts in place disclosure requirements for the firms. The standard application and the disclosure of mandatory and voluntary segmental items depend on the firms' disclosure policy. The benefits and costs perceived by managers are connected with segment disclosure and these affect their segment disclosure interpretation of the standard. The current study adopts three theoretical perspectives, namely, agency, proprietary and regulation and enforcement theory. The current sub-section provides an overview of these theories, while the hypothesis development section demonstrate the linkages between the current study's objectives, the literature review, and the theoretical framework of the current study, in order to develop a hypothesis.

Agency theory

Agency theory, first detailed in Jensen and Meckling (1976) and explained in greater detail in Eisenhardt (1989), is the theoretical framework most often used to explain why managers might be interested in engaging in earnings management. In many modern corporations, the shareholders of the firm are the owners (the principals) and appoint managers as agents to act on their behalf. However, modern corporations are increasingly complex and, because shareholdings tend to be diffuse, it can be difficult for shareholders to carefully monitor managers. In particular, when managers are being rewarded for a firm's performance, they may have incentives for choosing the accounting measures or for choosing to record transactions in such a way that they maximise their own compensation (Healy, 1985). Apart from managing earnings for their own selfish purposes, managers in modern corporations could also be pressurised to earnings manage because of demanding shareholder expectations. Koller et al (2010, p. 43) describe an "expectations treadmill", in which shareholders demand that the managers of their companies constantly improve performance above expectations. When these demands are met, the expectations increase again, which puts additional pressure on managers. The continually ratcheting pressure on managers thus creates the potential for managers to consider earnings management to be an alternative way to meet shareholder expectations.

Proprietary Theory

Disclosure is put out to help to reduce information asymmetry. Common theories (e.g., agency, signalling, capital need and political cost) are used in this era, and they recommend that firms are motivated to disclose more financial information as requirements under accounting standards. However, financial information disclosure to outside parties may also include proprietary information that could be likewise taken and used by the company's competitors, by opposed shareholders and employees, or by tax authorities, with a negative effect on the disclosing firm. Accordingly, when the firm decides the volume of information to be disclosed, firms face a compromise between the benefits of disclosing information to capital markets and the drawbacks of revealing proprietary information. (Hayes and Lundholm, 1996; Leuz, 2004). The managers' motivation, to avoid possible negative effects and retain competitive advantage, touches the disclosure of segmental information provided by firms. Prior studies on segmental disclosure have revealed that the segment disclosure's quantity or quality is limited by the proprietary cost that is associated with the disclosure of more detailed information to the market (Leuz, 2004). The association between the competition level and the quantity or quality of the segment information disclosed by the firms in the market, seems to be positive. Consequently, the more competitive firms working in a market are, the more there is a propensity to disclose segmental information, as the higher competition level shrinks the competitive damage that is associated with segmental information. (Birt et al., 2006; Pisano and Landriana, 2012).

Regulation and Enforcement
Accounting is subjective, as it must fulfil the needs of several regulations, like the International Accounting Standards, the UK Companies Act of 2006, and those of the London Stock Exchange. The regulatory requirements will have no quality in relation to financial reporting if the firms do not fulfil the requirements of these regulations (Gaffikin, 2005). Prior studies' connotation is that the mechanism of enforcing and monitoring IFRS requirements, like an effective internal control system and the need for an independent auditor, are significant workings of accounting standards implementation. The other studies' results have also shown that compliance with IFRS demands the disclosure of more proprietary information.

2.3 Hypothesis Development

Prior studies propose that many factors could affect earnings management through financial reporting requirements. The results of these prior researches and the theory that is presented briefly in Section 2.2 are considered to develop the following hypotheses as an alternative means through which to answer the research questions raised by the association between earnings management (real earnings management) and the following variables.

Insert Table 1

Glaum and Street (2003)'s paper argues that a company with a higher growth rate has higher future funding requirements. In respect of agency theory, it has been found that there is a positive association between non-disclosures of geographical earnings and foreign sales growth (Hope and Thomas, 2008). Hope and Thomas (Ibid.) concluded their study by indicating that managers have a propensity to reduce the shareholder's abilities by reducing the disclosure quality, in order to observe their international operations' expansion. In addition, it can be argued that this occurs because for proprietary reasons too. The managers might attempt to shrink their competitors' ability, in different geographical locations, in order to monitor their business activities.

H_3: The growth rate is positively associated with real earnings management under the post-implementation review of IFRS 8.

Large firms are more likely to produce information with less cost, and to have the capability and resources to produce information adequately (Cooke, 1989; Ali et al., 2004). Nevertheless, agency, proprietorship and competition-like entry barriers similarly can encourage companies to disclose more (Jensen and Meckling, 1976; Watts and Zimmerman, 1978 and 1990; Dye, 1986; Hayes and Lundholm, 1996; Nagarajan and Sridhar, 1996) when conforming to accounting standards and regulations. In alignment with regulation and enforcement theory, it has been argued that the financial reporting regulators might be blamed for any possible future crisis. Further, proprietary theory has argued that the detailed disclosure in relation to segmental information might harm smaller firms more than the bigger ones (Katselas et al., 2011; Bens et al., 2009). Prior studies have used total assets as a proxy for size (Prather-Kinsey and Meek, 2004; Al-Ulis, 2006; Al-Jabri, 2008; Bova and Pereira, 2012; Fekete et al., 2008; Al- Shammari et al., 2008; De Vicente Lama et al., 2011; Pisano and Landriana, 2012), other studies have used total sales (Hodgdon et al., 2009; Fekete et al., 2008).Market capitalisation has also been used in previous studies (Al-Jabri, 2008). Furthermore, Crawford et al's (2012) study found differences between the segmental information disclosure of the FTSE 100 and of the FTSE 250 firms.

H_4: The size is negatively associated with real earnings management under the post-implementation review of IFRS 8.

As a proxy for the financial risk of a company, gearing can be used. With a higher gearing, the more unpleasant the debt covenant restraint is for the firm, therefore the higher the likelihood is that the firm will abuse the debt covenant. Where companies are experiencing difficult financial times, administrators are more motivated to violate Earnings Management. Callao and Jarne (2010), Vander et al. (2003) and Jelinek (2007) have proved that higher gearing gives rise to more prominent degrees of Earnings Management.

H_5: The gearing is negatively associated with real earnings management under the post-implementation review of IFRS 8.

Watts and Zimmermann (1986)'s paper has indicated that companies with higher profitability have a greater propensity to disclose financial information in detail in their annual reports in order to act as a good agent (Agency theory). Profitability is a common factor in compliance studies. Prior research has discovered negative associations between Earnings Management and ROA (Young, 1999; Dechow et al, 1995; Duke and Hunt, 1990).

H_6: The profitability is negatively associated with real earnings management under the post-implementation review of IFRS 8.

Liquidity is the ability of a company to meet its short-term financial liabilities and it is one of the vital factors when financial statements' users assess a company. The argument here is that the better liquidity a company has, the greater the incentive to engage in earnings management practices. Firms with a higher liquidity ratio are thus more likely to present more disclosure and to have a greater propensity to manage earnings.

H_7: The liquidity is positively associated with real earnings management under the post-implementation review of IFRS 8.

Regarding agency and regulation theories, companies that are internationally visible have varied stakeholder groups (working under different tax authorities and stock exchange markets' regulators, more overseas suppliers and customers, financiers, personnel, etc.). One can argue that a company with more international visibility has a greater disclosure quality so as to meet the information requirements of its stakeholders, and thus it has the propensity to engage in earnings management practices. Previous studies have suggested that there is a significant association between listing status and International Accounting Standards requirements. They have found that companies have more compliance when they are listed on more stock exchange markets (El-Gazzar et al., 1999; and Prather-Kinsey and Meek; 2004). Here, the argument may be that the international visibility of a company affects the quality of its disclosure.

H_8: The international visibility is positively associated with real earnings management under the post-implementation review of IFRS 8.

Previous studies (e.g., Hayes and Lundholm, 1996; Harris, 1998; Botosan and Stanford, 2005; Birt et al., 2006; Pisano and Landriana, 2012) have indicated that the competitive harm may decrease under a higher level of competition that is associated with segmental information disclosure. Firms therefore have more incentive to disclose segmental information when they operate in industries that are more competitive. This is aligned with Proprietary theory "In fact, the release of additional information could benefit the firm by reducing information asymmetries between management and shareholders." (Birt, et al., 2006, p236). The proxy for industry competition is the Herfindahl-Hirschamn Index (HHI), because it takes into consideration the company's relative size and the companies' distribution in a market. HHI indicates that a higher index means a higher concentration in a market and lower competition in the market. While Capital intensity proxies the barriers to entry in a market and is a negative proxy for market competition (Givoly et al., 1999; and Wang et al., 2011). This paper used each industry sector along with the HHI and the capital intensity of the companies to measure the market competition.

H_9: The capital intensity is negatively associated with real earnings management under the post-implementation review of IFRS 8.

H_{10}: The market concentration is positively associated with real earnings management under the post-implementation review of IFRS 8.

3. METHODOLOGY

3.1 Sample

To address the questions being tackled in the current investigation, the initial sample included all of the FTSE-100 firms. However, the current research includes some sampling criteria for the purposes of its analysis. First, the firm must provide segmental information in its annual report. Specifically, some FTSE-100 companies may only sell/produce one product, or may provide only one service and operate locally. As a result, they may not disclose any segmental information in their annual reports. Second, all financial firms, including banks, insurance companies, and other financial institutions, have been excluded from the initial sample in earnings management research, because the environments of their financial reporting are different from those of other industrial companies, and their accrual processes differ fundamentally which means that the expectation models of normal accrual activity are not able to capture them (Peasnell et al., 2000). Table 2 indicates that this has led to the exclusion of 15 companies which did not have any segmental reporting. It has thus resulted in 85 companies, which were FTSE 100 companies with segmental reports. Additionally, the sample taken for this particular study also concentrated on eliminating banks along with financial companies, which did not have any available data (29 companies). The final sample taken in order to conduct this study was comprised of 56 companies. These companies were selected on the basis of the firm's years from 2012 to 2016. Hence, the companies included in this study were included in the FTSE-100 for 2012-2016, based on which it can be stated that some of the factors affected the earnings management with respect to financial reporting.

<u>**Insert Table 2**</u>

3.2 Model

The independent variables are the adoption of IFRS 8[2], the size of the firm (Market Capitalism). For the control variables, we chose gearing, growth, return on assets and loss, and market

[2] For the purposes of the current study, the checklist collects data on the number of segments published and of segmental items for the operating segments that were reported, based on the IFRS 8 requirements. It also captures any voluntary disclosures that were provided over and above the IFRS 8's requirements. The checklist includes 18 mandatory segmental items and 5 voluntary items, if viewed by the CODM as needing to be reported. Companies were not penalised for non-disclosure of items, which did not apply to their circumstances; the relevance of each

competition. In this model, gearing is used as an indicator to replace 'debt covenant restraints tightness'. The control variables are revenue growth and the Return on Assets (ROA), and these variables are included to account for differences in performance. Skinner and Sloan (2002) stipulate that the growing companies are more attracted to the activities of Earnings Management because part of their incentive is not to disappoint their shareholders. Their companies are expected to show signs of growth year after year, and therefore their company directors would like to make market commentators aware of their performances and thus the earnings report may be over inflated.

Past studies discover negative relationships between Earnings Management and ROA (Young, 1999; Dechow et al, 1995; Duke and Hunt, 1990). With a higher gearing, the more unpleasant the debt covenant restraint is for the firm, the higher the likelihood there is for the firm to abuse the debt covenant. Where companies are experiencing difficult financial times, administrators are more motivated to violate Earnings Management. Callao and Jarne (2010), Vander et al. (2003) and Jelinek (2007) have proved that higher gearing gives rise to more prominent degrees of Earnings Management. Loss is incorporated into the control variables and Francis and Yu (2009) find that organisations reporting a loss the previous year are more averse to Earnings Management than those firms that report good profits. Finally, it is clarified that organisations with good profits are more motivated to take part in Earnings Management in order to keep their shareholders content. In this model, we will add dummies, which are 'industry' and 'year', to control for the effects of industry and the year. Whilst this methodology, being a popular adoption in the literature, will allow us to assess the companies cross-sectionally and to assess the differences in both time periods and in looking for evidence of their interpretation of IFRS 8 rules, it does not necessarily assist in measuring the interpretation of the standard. Compared to accruals management, real earnings management has not been extensively examined in the literature. The models described in this section are those of Roychowdhury (2006), and Cohen and Zarowin (2010). According to Cohen and Zarowin (2010), we start by estimating "normal" levels of CFO as a linear function of revenues and changes in revenues. Using the following panel regression for each industry and year, we have the following model:

$$\frac{CFO_{it}}{Assets_{i,t-1}} = \frac{k_{1t}}{Assets_{i,t-1}} + \frac{k_2 Sales_{it}}{Asset_{i,t-1}} + \frac{k_3 \Delta Sales_{it}}{Assets_{i,t-1}} + \varepsilon_{it}$$

The figures provided for this model are the "normal" CFO levels, and the abnormal CFOs are thus simply the actual CFO – normal CFO. Cohen and Zarowin (2010) also explained that production costs and discretionary expenses can be manipulated to manage earnings. The following models relate to production costs and discretionary expenses to sales respectively:

$$\frac{Prod_{it}}{Assets_{i,t-1}} = \frac{k_{1t}}{Assets_{i,t-1}} + \frac{k_2 Sales_{it}}{Assets_{i,t-1}} + \frac{k_3 \Delta Sales_{it}}{Assets_{i,t-1}} + \frac{k_4 \Delta Sales_{it-1}}{Assets_{i,t-1}} + \varepsilon_{it}$$

$$\frac{DiscExp_{it}}{Assets_{i,t-1}} = \frac{k_{1t}}{Assets_{i,t-1}} + \frac{k_2 Sales_{i,t-1}}{Assets_{i,t-1}} + \varepsilon_{it}$$

absent item was investigated, and the item was then classified as a "non-disclosure", or as being "non-applicable". In other words, the IFRS 8 mandatory items may not be applicable to all companies; the applicability of the items included in the checklist was based on the company and its operations. For instance, not all companies have joint ventures or associate companies, so the segmental item for 'profit from associates and joint ventures', which is required by IFRS 8 (if reviewed by the CODM), may also not be relevant. (See Appendix 1 for details).

Finally, the various real earnings management variables are combined into a single comprehensive measure that is used to summarise the extent to which real earnings management has been undertaken. The procedure for doing this is described with the following equations:

$$R_{CFO} = \frac{CFO_{it}}{Assets_{i,t-1}} * (-1)$$
$$R_{DISC} = \frac{DiscExp_{it}}{Assets_{i,t-1}} * (-1)$$
$$R_{PROD} = \frac{Prod_{it}}{Assets_{i,t-1}}$$
$$R_{PROXY} = R_{CFO} + R_{DISC} + R_{PROD}$$

Based on the hypothesis given in Section 2.3 and the description of the methodology, above, the following regression model has been built:

$$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \varepsilon$$

Where

$Y_j=$	Each sample company's earnings management proxies
$X_1=$	Number of reported segments
$X_2=$	Percentage of disclosed items mandatory, voluntary and total (IFRS8)
$X_3=$	Growth
$X_4=$	Size
$X_5=$	Gearing
$X_6=$	Profitability
$X_7=$	Liquidity
$X_8=$	International visibility
$X_9=$	Capital intensity
$X_{10}=$	Market concentration
$\varepsilon_j=$	error

Insert Table 3

Based on Table 3, above, it can be stated that the independent variables included a number of reported segments along with the percentage of disclosed items, growth, gearing, size, profitability, liquidity, capital intensity, market concentration, and international visibility. However, based on the study findings of Street et al. (2000), a significant concern related to predicted improvements can be observed. These were reported based on the overall quality of the financial reports, following the implementation of IFRS, and this was due to the lack of sustainability.

3.3 Checking outliers, multicollinearity

Insert Table 4

4. RESULTS AND DISCUSSION

This section presents the results of the analyses of all the data collected during the course of this study. The empirical results are then presented, beginning with the analysis of the relationship between earnings management, as estimated by the real earnings model. The results of the analysis showed that the post-implementation of IFRS 8 did lead to significant changes in earnings management behaviour, as measured on the earnings management model. The data for this study were collected manually from financial statements and from the OSIRIS database for the period (2012-2016). Then, the obtained data was transformed into Excel sheet form along with company characteristics (e.g., growth, market cap, ROA, etc.) and the disclosure of IFRS 8.

Insert Table 5

Table 5, above, focuses on highlighting the correlation among the number of items, % mandatory items, % voluntary items, % total items, growth, marketing capitalism, gearing, ROA, liquidity, international visibility, capital intensity, and the market concentration. In this context, it can be stated that the relationships between the variables have been observed to be both positive and negative. For instance, it has been observed that the relationship between the number of segments and the R-proxy is calculated to be -0.1781, which is negative. This indicates that if the value of one variable increases, the value for the other variable decreases. On the other hand, a positive relationship was noted between % voluntary items and R-proxy. This indicates the fact that the hike in the value of one variable leads to a simultaneous increase in the value of another variable. In this context, Goncharov (2005) has stated that the implementation of the IFRS practice has no effects on the management of the earnings of a company. In addition, it will degrade the quality of financial reporting.

Insert Table 6

Table 6, above, focuses on highlighting the regression of real earnings management against the three segments' models. In other words, it illustrates the results so as to answer the main research question of the paper, which is to investigate an association between earnings management and segmental reporting. The adjusted-R^2 of the mandatory segment has been calculated as 56.80%. The same is found for the voluntary segment model, and the results of the total segment model are recorded to be 54.65% and 56.53%, respectively[3]. Hence, this has helped in comparing the overall explanatory power that is related to the three models that comprise the various predictors. In this case, the predictors have included the number of reported segments, the percentage of disclosed segmental items, growth, market capitalism, gearing, liquidity, ROA, market intensity, international visibility, and market concentration. In this particular context, as presented in Table 6, most of the null hypotheses are being rejected for those in which an alternative hypothesis can be used. For example, with respect to the mandatory segments model hypothesis, which are being rejected at 0.01 significance level, this includes ROA (-0.03) and market intensity (-0.07). On the other hand, the null hypothesis, which is rejected at 0.05 significance level includes the market relating to the voluntary segment model, which is market concentration (0.01). The reason for accepting a null hypothesis is that it can help in claiming a level of significance of $P<0.05$ (Filho et al., 2013). Hence, in this context, it has been evident that standards relating to the accounting needs for high quality are used in order to maintain an accurate level of significance. This is because it will help to achieve and maintain economic and financial systems (Ball, 2001).

[3] Moreover, the value helps in predicting the adjusted-$R^{2,}$ which is often perceived to be an upgraded version of $R^{2.}$ In this context, the adjusted-R^2 is observed to increase if it helps in improving all the predictors that are related to the respective models mentioned above. On the other hand, the R^2 value decreases when the predictor is observed to have a negative impact on the three models. This can thus help in assessing the development of a firm from the various interpretative angles.

The results of regression analysis with respect to the three models of the segmental reporting have represented slight variations in the obtained F-values. Contextually, the F-value of the mandatory segment model was calculated to be 13.75%, which was relatively more in comparison to the other two models. The F-value of the voluntary was again observed to be 12.69%, while the data recorded for the total segments model was 13.61%. The F value in the regression refers to the test outcomes, where a null hypothesis implies that the regression coefficients are zero. Since the F-values of all of the models are high, it does not support the null hypothesis. An alternative hypothesis is more compatible in this context. This is because the null hypothesis refers to the variances in the errors, which are often found to be equivalent to one another (Breusch & Pagan, 1979). Hence, in this regard, the disclosure of this information helps to decrease information asymmetry. However, it is essential to disclose the company's segmental information, since it has become mandatory under IFRS 8. However, the quantity of the information being disclosed in the capital markets depends on the company, which has its own advantages, as well as its disadvantages (Hayes and Lundholm, 1996; Leuz, 2004). Following proprietary theory, the disclosure of the information can help in significantly reducing its asymmetry. Hence, it has been evident that agency, along with capital need, signalling, and the political cost theories, assist in recommending that the companies reveal information with respect to their financial conduct in a stated period. This also includes complying with all the mandatory segmental disclosures needed for the information on the basis of the IFRS 8 management approach. Additionally, it is also essential to maintain the amount of information that will be shared in the markets (Hayes and Lundholm, 1996; Leuz, 2004). From an overall perspective it can thus be stated that this will significantly help in understanding and interpreting the accounting standards that the companies need to follow in order to assess their development from the interpretive angles.

The increasing and positive value of the adjusted-R^2 is good for the performance of the company. In order to maintain the adjusted R^2 value, the managers thus need to become involved in enhancing the overall earnings management. According to Jensen and Meckling (1976), this can also help to manage earnings better. It was further evident that the managers are pressurised to manage earnings due to the shareholders' demands. The adjusted-R^2 can help the managers of the company to assess the predictors relating to the model so as to enhance organisational performance. In this particular context, it has thus been observed that using the adjusted-R^2 has helped in comparing the overall suitability of fit relating to the regression model. This has been compared by taking into consideration the different types of independent variables, as in Table 6, above. Contextually, the agency theory offered by Jensen and Meckling (1976) was significantly detailed in the theoretical framework presented by the Eisenhardt (1989). This particular framework has significantly focused on the reasons for which the managers may be interested in becoming involved in the management of the company's earnings. In respect of the modern corporations, it has therefore been observed that the shareholders of the company are the owners, and it is they who appoint managers to act on their behalf. Additionally, the managers are rewarded if they perform well for the company. It will thus be essential for the companies, as well as for the managers, to act according to the accounting standards or measures in order to select the recorded transactions, and this can assist in maximising the overall compensation (Healey, 1985).

The results of regression analysis with respect to the three models of the segmental reporting show that the number of reported segments is associated negatively with real earnings management. This indicates that managers engage in fewer real earnings management practices when the reported number of segments is increasing. This result is aligned with regulation theory, since the greater the segmental number that is disclosed the more this makes the stakeholders more likely to be aware of setting internal control procedures and policies in order to stop managers from being engaged in real earnings management practices. Further, the results show a positive association between real earnings management and the parentage of the reported segment items under the mandatory and total disclosure models. However, the empirical results show that the voluntary disclosure of segmental items has insignificant effects on real earnings activities. For instance, Emmanuel (2012) found that UK firms that voluntarily adopted IFRS tended to have a more significant positive change in equity and earnings than firms that adopted IFRS on a non-voluntary basis. The firms in this study's samples that adopted IFRS were thus likely to have been a self-selected sample that may have adopted it because of a higher level of financial sophistication. This, in turn, could mean that the firms in the sample that adopted IFRS had more knowledge about the ways in which earnings management could be performed. This effect may have led to the positive relationship between IFRS adoption and earnings management measured on their real earnings model.

Under ruining the regression of the three above models, growth and market concentration show a significant positive association with real earnings management. While a negative association has resulted in market capitalism, ROA and capital intensity. Additionally, this has further assisted in comparing the different variables of the mandatory segments model, the voluntary segments model, and the total segments model. This includes growth, market capitalism, ROA, gearing, liquidity, market intensity, and market concentration. However, there are several factors which affect the overall requirements of the earnings management and financial reporting of the company. In summary, these results indicate that segment-related disclosures are of interest, and real earnings managers are using them when they are making investment decisions. Accordingly, it is relevant for the segmental information disclosed by UK FTSE-100 firms to rely on earnings management and this may explain the variations in earning management practices.

5. CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

The results of the current investigation lead to concludes that the IFRS 8's management approach is useful to internal and external decision makers especially that it explains its relation with real earnings practices. For example, the current paper results suggest that the management approach of IFRS 8 enhanced the perceptions of the managers about the reliable segmental information with further insights to publish such information to the shareholders. The current paper results also suggest that the segmental earnings effect on the extent of segment reports, rather than by equity of the firm. The implication of these findings for regulators is thus that allowing firms to disclose voluntary segmental information in accounting standards will not necessarily lead to improved earnings quality on their own: it is likely that companies will opt to report using which ever accounting standard allows them to report the "best" earnings. DeAngelo et al (1996) have suggested that the managers of public firms have strong incentives to avoid losses in reporting earnings; it is therefore likely that the incentive to avoid losses in this case would trump any desire to use a more stringent financial reporting standard.

The current study has some limitations. Our sample size was reduced, due to a lack of segmental information for several of the UK FTSE-100 firms, and thus many firms had to be excluded for data availability reasons. The effect of industry as a control variable was not included when the models were run, yet the industry effect has been run in measuring real earnings management.

REFERENCES

American Instituite of Certifieid Public Accountants (2017) IFRS Resources [online] http://www.ifrs.com/ifrs_faqs.html (accessed 28 March 2017).

Al-Jabri, H., 2008. *Financial reporting practices in Oman and compliance with disclosure requirements of international reporting standards* (Doctoral dissertation, Cardiff University).

Ali, M.J., Ahmed, K. and Henry, D., 2004. Disclosure compliance with national accounting standards by listed companies in South Asia. *Accounting and Business Research*, *34*(3), pp.183-199.

Al-Shammari, B., Brown, P. and Tarca, A., 2008. An investigation of compliance with international accounting standards by listed companies in the Gulf Co-Operation Council member states. *The International Journal of Accounting*, *43*(4), pp.425-447.
Al-Ulis, I., 2006. *Corporate financial reporting reform in economics in transition: the extent of compliance with international financial reporting standards disclosure requirements in annual reports* (Doctoral dissertation, University of Surrey).
André, P., Filip, A. and Moldovan, R. (2016), "Segment Disclosure Quantity and Quality under IFRS 8: Determinants and the Effect on Financial Analysts' Earnings Forecast Errors", The International Journal of Accounting, Vol. 51 No. 4, pp. 443–461.
Balakrishnan, R., Harris, T.S. and Sen, P.K., 1990. The predictive ability of geographic segment disclosures. *Journal of Accounting research*, pp.305-325.
Ball, R., 2001. Infrastructure requirements for an economically efficient system of public financial reporting and disclosure. Brookings-Wharton papers on financial services, 2001(1), pp.127-169.
Ball, R., 2006. International Financial Reporting Standards (IFRS): pros and cons for investors. Accounting and business research, 36(sup1), pp.5-27.
Ball, R., Robin, A. and Wu, J.S., 2003. Incentives versus standards: properties of accounting income in four East Asian countries. Journal of accounting and economics, 36(1), pp.235-270.
Barth, M.E., Landsman, W.R. and Lang, M.H., 2008. International accounting standards and accounting quality. Journal of accounting research, 46(3), pp.467-498.
Behn, B.K., Nichols, N.B. and Street, D.L., 2002. The predictive ability of geographic segment disclosures by US companies: SFAS No. 131 vs. SFAS No. 14. *Journal of International Accounting Research*, *1*(1), pp.31-44.
Bens, D.A. and Monahan, S.J., 2004. Disclosure quality and the excess value of diversification. *Journal of Accounting Research*, *42*(4), pp.691-730.
Birt, J.L., Bilson, C.M., Smith, T. and Whaley, R.E., 2006. Ownership, competition, and financial disclosure. *Australian Journal of Management*, *31*(2), pp.235-263.
Birt, J., Joshi, M. and Kend, M. (2017), "Segment reporting in a developing economy: the Indian banking sector", Asian Review of Accounting, Vol. 25 No. 1, pp. 127–147.
Botosan, C.A. and Stanford, M., 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. *The Accounting Review*, *80*(3), pp.751-772.
Bova, F. and Pereira, R., 2012. The determinants and consequences of heterogeneous IFRS compliance levels following mandatory IFRS adoption: Evidence from a developing country. *Journal of International Accounting Research*, *11*(1), pp.83-111.
Breusch, T.S. and Pagan, A.R., 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pp.1287-1294.
Bugeja, M., Czernkowski, R. and Moran, D. (2015), "The impact of the management approach on segment reporting", Journal of Business Finance and Accounting, Vol. 42 No. 3–4, pp. 310–366.
Burgstahler, D.C., Hail, L. and Leuz, C., 2006. The importance of reporting incentives: Earnings management in European private and public firms. The accounting review, 81(5), pp.983-1016.
Cai, L., Rahman, A.R. and Courtenay, S.M., 2008. The effect of IFRS and its enforcement on earnings management: An international comparison.
Callao, S. and Jarne, J.I., 2010. Have IFRS affected earnings management in the European Union?. Accounting in Europe, 7(2), pp.159-189.
Capkun, V., Collins, D.W. and Jeanjean, T., 2012. Does adoption of IAS/IFRS deter earnings management. Google Scholar, SSRN, uiova. edu, accessed, 22(8), p.2014.
Capkun, V., Jeny-Cazavan, A., Jeanjean, T. and Weiss, L.A., 2008. Earnings management and value relevance during the mandatory transition from local GAAPs to IFRS in Europe.
Christensen, H.B., Lee, E., Walker, M. and Zeng, C., 2015. Incentives or standards: What determines accounting quality changes around IFRS adoption?. European Accounting Review, 24(1), pp.31-61.

Cohen, D.A. and Zarowin, P., 2010. Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of accounting and Economics*, *50*(1), pp.2-19.

Cooke, T.E., 1989. Disclosure in the corporate annual reports of Swedish companies. *Accounting and business research*, *19*(74), pp.113-124.

Cooke, T. E., & Wallace, R. S. O. (1989), "Global surveys of corporate disclosure practices and audit firms: a review essay", Accounting and Business Research, Vol. 20 No. 77, pp. 47-57.

Crawford, L, Extance, H., Helliar, C.V. and Power, D.M., 2012. Operating segments: The usefulness of IFRS 8. [online] An ICAS Insight Publication, available: http://search.ror.unisa.edu.au/record/UNISA_ALMA11143738450001831 [accessed 17 June 2018].

DeAngelo, H., DeAngelo, L. and Skinner, D.J., 1996. Reversal of fortune dividend signaling and the disappearance of sustained earnings growth. *Journal of financial Economics*, *40*(3), pp.341-371.

Dechow, P.M., Sloan, R.G. and Sweeney, A.P., 1995. Detecting earnings management. Accounting review, pp.193-225.

De Vicente Lama, M., Sánchez, H.M. and Sobrino, J.N.R., 2011. Compliance with disclosure requirements for investment properties: A comparative study between Spain and UK. [online], available: http://www.aeca1.org/pub/on_line/comunicaciones_xivencuentroaeca/cd/126a.pdf [accessed 17 June 2018].

Di Carlo, F. and Lucchese, M. (2016), "The Impact of IFRS 8 on Segment Disclosure Practice: Panel Evidence from Italy", International Journal of Accounting and Financial Reporting, Vol. 6 No. 1, pp. 96–126.

Duke, J.C. and Hunt, H.G., 1990. An empirical examination of debt covenant restrictions and accounting-related debt proxies. Journal of accounting and Economics, 12(1-3), pp.45-63.

Dulitz, L., 2009. IFRS: A Preparer's Point of View. Journal of accountancy, 207(4), p.46.

Dye, R.A., 1986. Proprietary and nonproprietary disclosures. *Journal of business*, pp.331-366.

Eisenhardt, K.M., 1989. Agency theory: An assessment and review. *Academy of management review*, *14*(1), pp.57-74.

El-Gazzar, S.M., Finn, P.M. and Jacob, R., 1999. An empirical investigation of multinational firms' compliance with international accounting standards. *The International Journal of Accounting*, *34*(2), pp.239-248.

Emmanuel Iatridis, G., 2012. Voluntary IFRS disclosures: evidence from the transition from UK GAAP to IFRSs. *Managerial Auditing Journal*, *27*(6), pp.573-597.

Ettredge, M.L., Kwon, S.Y., Smith, D.B. and Zarowin, P.A., (2005). The impact of SFAS No. 131 business segment data on the market's ability to anticipate future earnings. The Accounting Review, 80(3), pp.773-804.

Iwata, E., 2010. Will going global extend to accounting? Switching to international rules has its pros and cons. USA TODAY B, 1, p.2010.

Fearnley, S. and Sunder, S., 2012. Global Accounting Rules–an unfeasible aim. Financial Times: June, 3.

Fekete, S., Matis, D. and Lukács, J., 2008. Factors influencing the extent of corporate compliance with IFRS-The Case of Hungarian listed companies. [online], available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295722 [accessed 17 June 2018].

Filho, D. B. F., Paranhos, R., da Rocha, E. C., Batista, M., da Silva, J. A., Santos, M. L. W. D. and Marino, J. C., 2013. When Is Statistical Significance Not Significant. Brazilian Political Science Review, 7 (1), 31-55.

Francis, J.R. and Yu, M.D., 2009. Big 4 office size and audit quality. The Accounting Review, 84(5), pp.1521-1552.

Gaffikin, M., 2005. Regulation as accounting theory. [online], University of Wollongong School of Accounting and Finance Working Paper No. 9, available: http://ro.uow.edu.au/accfinwp/50/ [accessed 17 June 2018]

Gassen, J., Uwe Fülbier, R. and Sellhorn, T., 2006. International differences in conditional conservatism–the role of unconditional conservatism and income smoothing. European Accounting Review, 15(4), pp.527-564.
Givoly, D., Hayn, C. and D'souza, J., 2000. Measurement errors and information content of segment reporting. *Review of Accounting Studies*, *4*(1), pp.15-43.
Glaum, M. and Street, D.L., 2003. Compliance with the disclosure requirements of Germany's new market: IAS versus US GAAP. *Journal of International Financial Management & Accounting*, *14*(1), pp.64-100.
Goncharov, I., 2005. Earnings management and its determinants: Closing gaps in empirical accounting research (Vol. 5). Peter Lang Pub Inc.
Hayes, R.M. and Lundholm, R., 1996. Segment reporting to the capital market in the presence of a competitor. *Journal of Accounting Research*, pp.261-279.
Healy, P.M. and Wahlen, J.M., 1999. A review of the earnings management literature and its implications for standard setting. Accounting horizons, 13(4), pp.365-383.
Healy, P.M., 1985. The effect of bonus schemes on accounting decisions. *Journal of accounting and economics*, *7*(1-3), pp.85-107.
Herrmann, D. & Thomas, W. (2000) A Model of Forecast Precision Using Segment Disclosures: Implications for SFAS No.131. Journal of International Accounting, Auditing and Taxation, 9, 1, 1-18.
Hodgdon, C., Tondkar, R.H., Adhikari, A. and Harless, D.W., 2009. Compliance with International Financial Reporting Standards and auditor choice: New evidence on the importance of the statutory audit. *The International Journal of Accounting*, *44*(1), pp.33-55.
Hope, O.K. and Thomas, W.B., 2008. Managerial empire building and firm disclosure. *Journal of Accounting Research*, *46*(3), pp.591-626.
Hope, O., Kang, T., Thomas, W.B. and Vasvary, F. (2008), “Pricing and mispricing effects of SFAS 131”, Journal of Business Finance & Accounting, Vol. 35 Nos 3/4, pp. 281-306.
Hung, M. and Subramanyam, K.R., 2007. Financial statement effects of adopting international accounting standards: the case of Germany. Review of accounting studies, 12(4), pp.623-657.
International Accounting Standards Board (2013), “IASB completes post-implementation review of IFRS 8 operating segments”, 18 July, available at: www.ifrs.org/Alerts/ProjectUpdate/Pages/IASB-Completes-Post-Implementation-Review- of-IFRS-8-Operating-Segments-July-2013.aspx (accessed 7 October 2016).

International Accounting Standards Board (IASB) (2006), International Financial Reporting Standard No. 8 Operating Segments, International Accounting Standards Board, London.
Jeanjean, T. and Stolowy, H., 2008. Do accounting standards matter? An exploratory analysis of earnings management before and after IFRS adoption. Journal of accounting and public policy, 27(6), pp.480-494.
Jelinek, K., 2007. The effect of leverage increases on earnings management. The Journal of Business and Economic Studies, 13(2), p.24.
Jensen, M.C. and Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, *3*(4), pp.305-360.
Kajüter, P. and Nienhaus, M. (2017), “The Impact of IFRS 8 Adoption on the Value Relevance of Segment Reports”, A Journal of Accounting, Finance and Business Studies, Vol. 53 No. 1, pp. 28–58.
Katselas, D., Birt, J. and Kang, X.H., 2011. International firm lobbying and ED 8 operating segments. *Australian Accounting Review*, *21*(2), pp.154-166.
Koller, T., Goedhart, M. and Wessels, D., 2010. *Valuation: measuring and managing the value of companies* (Vol. 499). john Wiley and sons.
Leung, E. and Verriest, A. (2015), “The Impact of IFRS 8 on Geographical Segment Information”, Journal of Business Finance and Accounting, Vol. 42 No. 3–4, pp. 273–309.

Leuz, C., 2010. Different approaches to corporate reporting regulation: How jurisdictions differ and why. Accounting and business research, 40(3), pp.229-256.
Leuz, C., 2004. Proprietary versus non-proprietary disclosures: Evidence from Germany. *The economics and politics of accounting*, *13*(4), pp.164-199.
Leuz, C., Nanda, D. and Wysocki, P.D., 2003. Earnings management and investor protection: an international comparison. Journal of financial economics, 69(3), pp.505-527.
Lim, S.J., White, G., Lee, A. and Yuni, Y. (2017), "A longitudinal study of voluntary disclosure quality in the annual reports of innovative firms", Accounting Research Journal, Vol. 30 No. 1, pp. 1–27.
Mardini G, H. and Ammar, S (Forthcoming) "Quality and Quantity of FTSE-100 Segmental Information Reporting", Accounting Research Journal, Vol. X, No. X, pp. XXX-XXX.
Mardini, G, H., Tahat Y, A. and Power D, M. (2018) "The Extent of Segmental Reporting and its Value Relevance: Cross-Country Evidence", Journal of Accounting in Emerging Economies, Vol. 8, No. 2, pp. 279-299.
Mardini, G.H., Crawford, L. and Power, D.M., 2012. The impact of IFRS 8 on disclosure practices of Jordanian listed companies. Journal of Accounting in Emerging Economies, 2(1), pp.67-90.
Nagarajan, N.J. and Sridhar, S.S., 1996. Corporate responses to segment disclosure requirements. *Journal of Accounting and Economics*, *21*(2), pp.253-275.
Niemeier, C.D., 2008. Keynote address on recent international initiatives. 2008 Sarbanes-Oxley, SEC and PCAOB Conference, New York State Society of CPAs, New York (September 10).
Nobes, C., 2011. IFRS practices and the persistence of accounting system classification. Abacus, 47(3), pp.267-283.
Peasnell, K.V., Pope, P.F. and Young, S., 2000. Detecting earnings management using cross-sectional abnormal accruals models. Accounting and Business research, 30(4), pp.313-326.
Pisano, S. and Landriani, L., 2012. The determinants of segment disclosure: an empirical analysis on Italian listed companies. *Financial reporting*. (1), pp113-132.
Prather-Kinsey, J. and Meek, G.K., 2004. The effect of revised IAS 14 on segment reporting by IAS companies. *European Accounting Review*, *13*(2), pp.213-234.
Ramsey, J.B., 1969. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.350-371.
Roychowdhury, S., 2006. Earnings management through real activities manipulation. *Journal of accounting and economics*, *42*(3), pp.335-370.
Soderstrom, N.S. and Sun, K.J., 2007. IFRS adoption and accounting quality: a review. European Accounting Review, 16(4), pp.675-702.
Skinner, D.J. and Sloan, R.G., 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. Review of accounting studies, 7(2), pp.289-312.
Street, D.L., Gray, S.J. and Bryant, S.M., 1999. Acceptance and observance of international accounting standards: An empirical study of companies claiming to comply with IASs. The International Journal of Accounting, 34(1), pp.11-48.
Street, D.L., Nichols, N.B. and Gray, S.J., 2000. Assessing the acceptability of international accounting standards in the US: An empirical study of the materiality of US GAAP reconciliations by non-US companies complying with IASC standards. The International Journal of Accounting, 35(1), pp.27-63.
Tsakumis, G.T., Doupnik, T.S. and Seese, L.P., 2006. Competitive harm and geographic area disclosure under SFAS 131. *Journal of international accounting, auditing and taxation*, *15*(1), pp.32-47.
Vander Bauwhede, H., Willekens, M. and Gaeremynck, A., 2003. Audit firm size, public ownership, and firms' discretionary accruals management. The International Journal of Accounting, 38(1), pp.1-22.

Van Tendeloo, B. and Vanstraelen, A., 2005. Earnings management under German GAAP versus IFRS. European Accounting Review, 14(1), pp.155-180.
Wang, Q., Ettredge, M., Huang, Y. and Sun, L., 2011. Strategic revelation of differences in segment earnings growth. *Journal of Accounting and Public Policy*, *30*(4), pp.383-392.
Watts, R.L. and Zimmerman, J.L., 1986. Positive accounting theory. Prentice – Hall Inc.
Watts, R.L. and Zimmerman, J.L., 1978. Towards a positive theory of the determination of accounting standards. *Accounting review*, pp.112-134.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, pp.817-838.
Young, S., 1999. Systematic measurement error in the estimation of discretionary accruals: An evaluation of alternative modelling procedures. Journal of Business Finance & Accounting, 26(7-8), pp.833-862.
Zéghal D, Chtourou S, Sellami YM. An analysis of the effect of mandatory adoption of IAS/IFRS on earnings management. Journal of international accounting, auditing and taxation. 2011 Jan 1;20(2):61-72.

Table 1: Summary of the possible links between the different company characteristics and the theoretical framework of corporate earnings management behaviour

Hypothesis	Factor	Association	Theory		
			Agency	**Proprietary**	**Regulation**
H_{a1}	Number of Segments	-	X		X
H_{a2}	Adoption of Post IFRS8	+	X		X
H_{a3}	Growth rate	+	X	X	
H_{a4}	Size	-	X	X	X
H_{a5}	Gearing	-	X		
H_{a6}	Profitability	-	X		
H_{a7}	Liquidity	+	X		
H_{a8}	International visibility	+	X	X	X
H_{a9}	Capital intensity	-		X	
H_{a10}	Market concentration	+		X	

Table 2: Sampling

Panel A: Sampling Process Employed	
FTSE -100	100 companies
Less: Companies do not have Segmental Report	15
FTSE-100 with Segmental Report	85
Less: Banks and Financial Companies, and not available data	29
Total	56
Firm-Year (2012-2016)	**280 observations**
Panel B: Final Sample per Sector	
Manufacturing	19
Services	32
Retail	5
Total	56

Salem and Mardini, Post-Implementation Review of IFRS 8 Interpret Earnings Management

Firm-Year (2012-2016)	**280 observations**

Table 3: Description of the independent variables, their measurement and the source of the information

Independent variable		**Measurement**	**Source of information**
Number of reported segments	X_1	Number of business segments	Annual Report + EIKON
Percentage of disclosed items	X_2	The model 3 times for mandatory, voluntary, and total items disclosed. 1 if item disclosed; 0 otherwise	Annual Report + EIKON
Growth	X_3	Growth rate of sales revenue	Annual Report
Size	X_4	Market Capitalism	OSIRIS
Gearing	X_5	Shareholder's funds to total liabilities	OSIRIS
Profitability	X_6	ROA (Net income to Total Assets)	OSIRIS
Liquidity	X_7	Current ratio (Current Assets to Current Liabilities)	OSIRIS
International visibility	X_8	Number of subsidiaries in foreign countries	OSIRIS
Capital intensity	X_9	Total PPE to total Assets	Annual Report
Market concentration	X_{10}	HHI (based on sales)	Annual Report

Table 4: Validity test of real earnings management against the three models of segmental reporting[4]:			
Dependent variable Real Earnings Management (R_{proxy})	**Mandatory Segments Model**	**Voluntary Segments Model**	**Total Segments Model**
Mean VIF	1.60	1.64	1.62
Breusch-Pagan Test	0.63	0.18	0.25
Probability	0.43	0.67	0.62

Table 5: Correlation

	R_{proxy}	No. segments	%Mandatory items	%Voluntary items	%Total items	Growth	Market Capitalism	Gearing	ROA	Liquidity	International Visibility
R_{proxy}	**1.0000**										
No. segments	**-0.1781**	**1.0000**									
%Mandatory items	**0.1465**	**0.1789***	**1.0000**								
%Voluntary items	**0.0328**	**-0.0429**	**-0.0820**	**1.0000**							
%Total items	**0.1465**	**0.1521***	**0.8748***	**0.4086***	**1.0000**						
Growth	**0.3202***	**-0.1102**	**-0.0727**	**-0.0708**	**-0.1068**	**1.0000**					
Market Capitalism	**-0.3225***	**0.1615***	**-0.0743**	**-0.1044**	**-0.1070**	**-0.1276**	**1.0000**				
Gearing	**-0.0439**	**0.1624***	**0.0919**	**0.2164***	**0.1918***	**-0.0131**	**0.1159**	**1.0000**			
ROA	**-0.4557***	**-0.0431**	**-0.2237***	**-0.0892**	**-0.2603***	**0.1792***	**0.1025**	**-0.0740**	**1.0000**		
Liquidity	**-0.0056**	**0.1955***	**-0.0529**	**0.0103**	**-0.0440**	**0.0670**	**-0.2157***	**-0.3463***	**0.1622***	**1.0000**	
International Visibility	**0.2130***	**0.0153**	**0.0878**	**-0.1203**	**0.0339**	**-0.0182**	**0.2549***	**0.0933**	**-0.2103***	**-0.1678***	**1.000**
Capital Intensity	**-0.3291***	**0.1589***	**0.0967**	**0.0480**	**0.1206**	**-0.0701**	0.0980	**0.0610**	**-0.1996***	**-0.0630**	**-0.364**
Market Concentration	**0.0344**	**0.2104***	**0.1224**	**0.0472**	**0.1448***	**-0.1858***	**0.6453***	**0.1255**	**-0.3464***	**-0.3350***	**0.217**

[4] On the basis of the Table 4 presented above, it focuses on depicting the overall validity test relating to the real earnings management. The mean VIF relating to mandatory segments model has been calculated to be the1.60, which indicates for voluntary segment model, it was 1.64, and the total segments model has been calculated to be 1.62. Since, the VIFS for all the models are exceeding the value 0.1, it can be stated the collinearity can cause problems (Ramsey, 1969). On the contrary, Breusch-Pagan test three models were recorded to be 0.63, 0.18, and 0.25 respectively. It can thus be depicted that the values of Breusch-Pagan test related are not equal to one another. However, the Breusch-Pagan tests indicate that the null hypothesis associated with different variances of the errors is observed to be equivalent to one another. Furthermore, probability was obtained for the mandatory segments model was calculated to be 0.43. The same for voluntary segments model was 0.67 and for total segments model was recorded to be 0.62 correspondingly. This indicates that the White test has an advantage in reducing biases using the Breusch-Pagan test, which occur when errors are abnormally distributed (White, 1980).

Salem and Mardini, Post-Implementation Review of IFRS 8 Interpret Earnings Management

Table 6: Regression of real earnings management against the three models of segmental reporting:			
Dependent variable Real Earnings Management (R_{proxy})	**Mandatory Segments Model**	**Voluntary Segments Model**	**Total Segments Model**
F	13.75 ***	12.69 ***	13.61 ***
Adjusted-R^2	56.80%	54.65%	56.53%
Number of reported segments	-0.01 (0.00)***	-0.01 (0.00)***	-0.01 (0.00)***
Segmental items	0.07 (0.03)**	0.00 (0.01)	0.08 (0.04)**
Growth	0.12 (0.04)***	0.13 (0.04)***	0.13 (0.04)***
Market Capitalism	-0.02 (0.01)***	-0.02 (0.01)***	-0.02 (0.01)***
Gearing	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
ROA	-0.03 (0.01)***	-0.03 (0.01)***	-0.03 (0.01)***
Liquidity	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
International Visibility	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
Capital Intensity	-0.07 (0.00)***	-0.07 (0.02)***	-0.08 (0.02)***
Market Concentration	0.01 (0.00)**	0.01 (0.00)**	0.01 (0.00)**

*** Denotes rejection of the null hypothesis at the 0.01 significance level.
** Denotes rejection of the null hypothesis at the 0.05 significance level.
* Denotes rejection of the null hypothesis at the 0.1 significance level.

Salem and Mardini, Post-Implementation Review of IFRS 8 Interpret Earnings Management

Appendix 1 Disclosure Index Checklist (Variable: Segmental Items)
IFRS 8 Mandatory Disclosures for Operating Segments (if reviewed by the CODM)
Profit
Assets
Liabilities
Revenue (external)
Depreciation & amortization
Other non-cash expenses
Reconciliation to consolidated accounts
Revenue (internal)
Basis of inter-segment pricing
Profit from associates and joint ventures
Basis of measurement
Interest revenue
Interest expense
Income tax expense
Factors used to identify the entity's segments
Entity-Wide (major customers)
Entity-Wide (products and services)
Entity-Wide (Geographic Information)
Voluntary Disclosures
Number of Employees by Segment
Capital Expenditure on plant and equipment
Intangible Assets by Segment
Non-Current Assets
Investment Activities

DECISION SCIENCES INSTITUTE

A study of earnings quality and earnings management of companies accused by the Securities and Exchange Commission: 2007-2016

Manuel A. Sánchez
Inter American University of Puerto Rico
Email: msanchez@arecibo.inter.edu

Ahmad H. Juma'h
University of Illinois Springfield
Email: jumah@uis.edu

ABSTRACT

This article relates earnings quality with earnings management to determine if good or poor earnings quality is an indicator for financial or accounting irregularities by considering the persistence of earnings, cash flows and accruals between companies with high or low earnings quality. A sample of 61 companies accused by the Securities and Exchange Commission of accounting irregularities and 61 control companies were used. The data analysis showed that accused companies have poorer quality than control ones; similarly, control companies presented less discretionary accruals and better return on assets than accused ones.

KEYWORDS: Earnings Management, Quality, Persistence, Fraud, Manipulation.

INTRODUCTION

Financial reporting shenanigans tend to be mostly associated with accrual-based earnings. Further, an important red flag is when cash flow does not move in the same direction as revenue, accounts receivable, and net income (Deo & Cathy, 2016), affecting the quality of earnings and signaling possible earnings management. Earnings management is an umbrella for managers' discretion that affects the reported accounting earnings or their interpretation. It starts with production and investments decisions, the choice of accounting treatment, the size of accruals, and decisions that affect the interpretation of reported earnings. Healy & Wahlen (1999) stated that "e*arnings management occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers"* (p. 368).

There are several other earnings management definitions, but most of them are difficult to operationalize directly using attributes of reported accounting numbers since they center on management intention which is unobservable (Dechow & Skinner, 2000). Although not to be confused with illegal activities to manipulate financial reports that do not reflect or show the economic reality of the company, this practice has been used for many years. Financial reporting requires managers to estimate future benefits and obligations such as allowance for doubtful accounts, and deferred revenues. These estimates allow managers to choose from different accounting procedures giving them the option to influence financial results. The problem arises when managers ignore those rules to show a more successful organization to influence shareholders' decisions. Researchers had been trying to predict earnings management for a long time. Since it is not possible to observe directly, researchers have been using several models to predict whether companies are involved in this practice.

LITERATURE REVIEW

The external pressure from shareholders, financial analysts and credit rating agencies take its toll on managers who struggle to make the financial results forecasted and try to achieve their goals through aggressive accounting procedures to give the illusion of financial strength, for example, the Enron case where executives engaged in deceitful practices to inflate earnings and increase stocks prices to profit themselves, this is consistent with the 2007 subprime crisis and the 2008 collapses of Bear Sterns and Lehman Brothers. During the 1960s and throughout the 1970s, the earnings management literature was focused mainly on capital markets, starting with the Mechanical Hypothesis that postulates a mechanical relationship between stock prices and accounting earnings. This hypothesis predicts that stock price is associated with accounting information changes and concluded that the market is systematically misled by accounting procedures. Kaplan and Roll (1972) stated that when the mechanical hypothesis is prevalent, investors and practitioners believe that their stock prices are inflated as reported earnings are inflated by corporate managers.

Healy (1985) tested for earnings management by comparing the mean of total accruals, scaled by lagged total assets, across the earnings management partitioning variable. Healy's study differs from most other earnings management studies predicting that systematic earnings management occurs in every period. DeAngelo (1986) used the last period's total accruals as the measure of non-discretionary accruals to investigate the accounting decisions made by managers who proposed a management buyout. Jones (1991) proposed a model that relaxes the assumption that nondiscretionary accruals are constant. Her model tried to control for changes in a firm's economic situations on nondiscretionary accruals. Dechow, Sloan and Sweeney (1995) developed the Modified Jones Model by adjusting for the changes in receivables in the event period.

Since the 1990s, earnings management have shifted back to capital markets incentives. Dechow & Skinner (2000) pointed out three reasons for this shift. One of the reasons is the market efficiency concept prevailing among early accounting researchers. Also, Capital Markets are usually available to a much wider spectrum of users of financial statements, making earnings management more likely to be effective as the cumulative effects of information asymmetry between investors and managers are significant. In addition, since the 1990's the stock market valuation based on accounting benchmarks substantially increased, making stock prices highly sensitive to accounting measures such as earnings. Earnings are the net benefits of a firm's operations according to the SFAC No. 2, the primary determinants of earnings quality are relevance and reliability, and these two dimensions make accounting information useful for decision making (FASB 1980).

Earnings Quality

Earnings quality is one of the most popular topics in the accounting literature. The subject is also of great interest to practitioners in the investing community such as analysts (both buy and sell side), hedge funds, and institutional investors (Khan, Peddireddy, & Rajgopal, 2019). Recent research has changed the focus to earnings quality and quality of accruals and the relation between earnings management and earnings quality. Earnings quality focuses on the extent to which the reported net income deviates from the actual, true earnings of the company

(Warshavsky, 2012). Most researchers agree that highly managed earnings have low quality. Dechow, Ge and Schrand (2010) provided a broad and very detailed review of earnings quality research by reviewing over 300 studies trying to define earnings quality broadly to help with decision making. Schipper and Vincent (2003) defined earnings quality as the extent which reported earnings faithfully represent Hicksian income, where representational faithfulness means "correspondence or agreement between a measure or description and the phenomenon that it purports to represent" (FASB Concepts Statement No. 2 para. 63). Dechow and Schrand (2004) argued that a high-quality earnings number is one that accurately reflects the company's current operating performance, is a good indicator of future operating performance, and is a useful summary measure for assessing firm value. Nakashima & Ziebart (2016) used accrual quality, discretionary accruals, and the accuracy of cash flows predictions as surrogates for earnings quality and found mixed results between tone at the top and those variables. They defined tone at the top as the ethical atmosphere created in the workplace by the organization's leadership and stated that the CEO sets the tone at the top, and it affects integrity, ethics, and other factors underlying a positive control environment.

The quality of earnings is of interest to those who use financial reports for contracting purposes and investments decisions (Schipper & Vincent, 2003). Decisions based on low-quality earnings will induce unintended wealth transfers, and from the investment side, low-quality earnings might provide a defective resource allocation signal (Schipper & Vincent, 2003). Earnings quality can vary among companies as a function of accruals, even in the absence of intentional earnings manipulation (Dechow & Schrand, 2004). One role of accruals is to shift or adjust the recognition of cash flows over time so that the adjusted numbers (earnings) better measure firm performance. Since accruals require assumptions and estimates of future cash flows, Dechow and Dichev (2002) argue that the quality of accruals and earnings is decreasing in the magnitude of estimation error in accruals. If net income is high, but cash flow from operations is low, then there is a need to scrutinize the quality of the net income with particular attention to the type of industry and the life cycle of the firm (Deo & Cathy, 2016). Nakashima & Ziebart (2016) argued that if the internal control system is good, segregation of duties will be the norm, and earnings management would be controlled, leading to an improvement in earnings quality. Accrual analysis is comprehensive- measuring both the effect of accounting method choices and operating, financing, and investments decisions insofar as they affect accruals, but there is a drawback; accruals models generate low-power tests of earnings management even for fairly high levels of manipulation as high as 5% of total assets (Jiambalvo, 1996).

For the estimation of unexpected accruals, some studies start with total accruals, measured as the difference between reported net income and cash from operations. Total accruals are regressed on variables that are proxies for normal accruals to allow for typical working capital needs, and gross fixed assets to allow for normal depreciation; unexpected accruals are thus the unexplained or residual components of total accruals (Healy & Wahlen, 1999). These unexpected or discretionary accruals are used by managers to manage earnings. A simple approach to measuring earnings quality as the inverse of estimates and judgments embedded in accruals is based on changes in total accruals. As long as some portion of accruals is both non-manipulated and approximately constant over time, changes in total accruals measure managerial manipulations and provide an inverse measure of earnings quality (Schipper & Vincent, 2003). Earnings quality can be improved when accruals smooth out value-irrelevant changes in cash flows, but earnings quality is reduced when accruals are used to hide value-relevant changes in cash flows (Dechow & Schrand, 2004). Less reliable accruals or earnings quality lead to lower earnings persistence.

Earnings Persistence

Earnings volatility or earnings persistence is the continuity and durability of the current earnings and denotes how constant or unstable earnings are. A company with fluctuating earnings may be seen as a risky investment, especially when cash is needed for capital investments because future cash flows can be uncertain. Earnings are, on average, more persistence that cash flows, but the cash flow component of earnings is more persistence than accruals (Dechow & Schrand, 2004). Sloan (1996) evaluated whether cash flow from operations and accruals had different implications for the persistence of future earnings and showed that the extent to which current earnings performance persists into the future is shown to depend on the relative magnitudes of the cash and accrual components of current earnings.

Dichev & Tang (2008) investigated the link between earnings volatility and earnings predictability and found that the consideration of earnings volatility brings substantial improvement in the prediction of both short-term and long-term earnings. Artikis & Papanastasopoulos (2016) examined the persistence, pricing, and economic significance of the cash component of earnings in the United Kingdom listed firms from 1981 to 2013. The results from the persistence tests indicated that there were systematic differences in the persistence among the cash subcomponents of earnings and that the cash component of earnings cannot be treated as a homogeneous unit. Richardson, Sloan, Soliman & Tuna (2005), linked accrual reliability to earnings persistence by constructing a model showing that less reliable accruals led to lower earnings persistence, also confirmed that less reliable accruals lead to lower earnings persistence and that investors do not fully anticipate the lower earnings persistence, leading to significant security mispricing. Dechow, Larson & Resutek (2019) stated that empirical regularity that accruals persist into future earnings at a lower rate than cash flows is frequently interpreted as evidence of low accrual quality. They predicted that accrual heterogeneity significantly impacts this empirical regularity and showed that their consistency measure captures unique aspects of reporting quality and that reporting consistency can induce both higher or lower earnings persistence depending on the measurement rules attributable to the underlying accrual.

RESEARCH STATEMENTS DEVELOPMENT

To detect earnings management or financial manipulation, we used a modified version of the Dechow and Dichev (2002) accrual quality model that classifies the quality of accruals and earnings in low and high quality, utilizing changes in working capital as a measure of accruals and cash flow from operations as a proxy for cash flow related to accruals. The residuals from the regression reflect the accruals that are not related to cash flow realizations, and their standard deviation is a firm-level measure of accrual and earnings quality. Our aim is to verify the following statement:

$S_{1:}$ firms with low-quality (high-quality) accruals are more likely to engage (not engage) in earnings management.

Dechow & Dichev (2002) used the standard deviation of the residuals from their regression as a firm-specific measure of accrual quality, where a higher standard deviation signifies lower quality. Therefore, we expect net income and CFO measures to be much closer together for firms with good earnings quality. CFOs feel that most earnings misrepresentation occurs in an attempt to influence stock price and to avoid adverse compensation and career consequences for senior executives (Dichev, Graham, Harvey, & Rajgopal, 2013).They also pointed out that the two most common red flags to potential misrepresentation are: persistent deviations between earnings and the underlying cash flows and deviations from industry and other peer experience. In this research, the first statement attempts to verify that firms with these characteristics have a higher probability of earnings management.

Most practitioners and academics agree that cash flows are more reliable than earnings, and unlike earnings that contain accruals, cash flows are not estimated, and for a company to manipulate it, it must incur real costs. Earnings are, on average, more persistence than cash flows, but the cash flow component of earnings is more persistence than accruals (Dechow & Schrand, 2004). Sloan (1996) stated that the extent to which current earnings performance persists into the future is shown to depend on the relative magnitudes of the cash and accrual components of current earnings. Kubota & Takehara (2019) investigated the time-series properties of accounting earnings and their components and proposed a new measure of earnings persistence in accordance with the vector autoregressive (VAR) model- linking earnings and stock returns. As a preliminary analysis, they estimated the first-order autocorrelations and test the stationary of five variables: earnings, cash flow from operations, total accruals, current accruals, and non-current accruals. They argued that earnings and noncurrent accruals have a more persistent time-series than cash flows and current accruals. Following the literature on earnings persistence, this research also verifies the following statements:

S_2: There is a difference (no difference) in the persistence of cash flow (earnings) between companies with high or low earnings quality.

RESEARCH METHODOLOGY

Research methodology utilized in earnings management and fraud detection varies depending on the question been asked or the hypothesis the researcher is trying to develop. For the manipulation of earnings, managers have a variety of choices to increase or decrease earnings. In the 1970s and early 1980s, many studies found that managers can exercise discretion through the choice of accounting methods or policies (Sun & Rath, 2010). Executives or managers can use specific accounting policies to manipulate earnings (Juma'h, 2014, 2019). Researchers have tried to detect earnings management by dividing accruals into two components: discretionary and non-discretionary. Discretionary accruals are the ones that managers can choose within the flexibility of accounting or from Generally Accepted Accounting Principles (GAAP) to adjust the firm's cash flows. Non-discretionary accruals result from normal business operations. Since discretionary accruals are the management choices to make accounting adjustments, they are of our interest because of the opportunity they represent to manipulate earnings. Also, those discretionary accruals can determine the quality of earnings of the company. Earnings management literature agreed that a higher number of accruals denotes the poor quality of earnings and the financial press suggest that a company has an earnings-quality problem if earnings contain unusual items or lack transparency, even if reported earnings and the related disclosures are in accordance with GAAP (Dechow & Schrand, 2004).

The main research question is how the quality of earnings reported by a firm can point out or signal earnings management or financial statements fraud. Good earnings quality will give the investor a sure ground to analyze companies and make an investment in a well-informed environment. The key objective of this research is to examine the relation between earnings quality and earnings management and how these variables, among others, affect the financial reporting of firms; or similarly if firms with low-quality accruals are more prone to engage in earnings management or financial reporting irregularities. To obtain an earnings quality measure, we used a regression method based on past, present, and future cash flows from operations.

Past researches (Sloan, 1996, Dechow & Schrand, 2004, Richardson *et al.* 2005) showed that the accrual component of earnings is less persistent than the cash flow component of earnings and attributes this difference to the greater subjectivity of accruals. Earnings volatility or earnings persistence is the continuity and durability of the current earnings and denotes how constant or unstable earnings are. A company with fluctuating earnings may be seen as a risky investment, especially when cash is needed for capital investments because future cash flows can be uncertain. Richardson *et al.* (2005) suggested that the recognition of less reliable accrual estimates introduces measurement error that reduces earnings persistence and leads to significant security mispricing, allowing for less verifiable and less reliable estimates into accounting numbers that can compromise their usefulness.

Research design

We use multiple linear regression with dummy variables to estimate the nondiscretionary component of accruals. The nondiscretionary component reflects the business condition that naturally creates and destroy accruals, while discretionary accruals identify management choices. The purpose of separate discretionary accruals from total accruals is to determine the number of accruals that are due to management's choices; this is a better proxy for earnings quality.

The equation to obtain total accruals is derived by using:

$$TA_{i,t} = \Delta CA_{i,t} - \Delta CL_{i,t} - \Delta CASH_{i,t} + \Delta STDebt_{i,t} - DEPN_{i,t} / (A_{t-1})$$

Where,

$TA_{i,t}$ = Total accruals

$\Delta CA_{i,t}$ = change in current assets in year t

$\Delta CL_{i,t}$ =change in current liabilities in year t.

$\Delta CASH_{i,t}$ = change in cash and cash equivalent in year t.

$\Delta STDebt_{i,t}$ = change in short-term debt in year t.

$\Delta DEPN_{i,t}$ = depreciation and amortization expense in year t.

A = Total assets.

This equation reflects a balance sheet approach used in prior researches by Healy (1985), Jones (1991) and Dechow *et al*. (1995). After estimating total accruals, the Jones Model is used to estimate non-discretionary accruals. Non-discretionary accruals are the residuals of the following regression:

$$TA_{i\,t} / A_{i\,t-1} = \alpha_{I\,t} (1/A_{i\,t}) + \beta_1 (\Delta REV_{i\,t} /1/A_{i\,t}) + \beta_2 (PPE_{i\,t} /1/A_{i\,t}) + e_t$$

Where,

$TA_{i\,t}$ = total accruals for a company i in year t

ΔREV = revenues in year t minus revenues in year t-1

$PPE_{i\,t}$ = gross property plant and equipment in year t

$A_{i\,t-1}$ = total assets in year t-1

e_t = error term in year t.

Accruals require estimation from the management; these estimates allow management the ability to choose different accounting procedures giving them the option to influence financial results presenting an economic picture different from the real performance. Also, these estimations, either intentional or unintentional, reduces the ability of earnings to reflect future cash flows (Dechow & Schrand, 2004). Earnings or cash flows are judged to be of higher quality when they are: more persistent and less volatile, more strongly associated with future cash flow realizations and more strongly associated with contemporaneous stock performance or market value (Dechow & Schrand, 2004, Dechow, Larson & Resutek, 2019).

Similar to Dechow & Dichev (2002), to obtain a practical measure of working capital accruals quality, a firm-level time-series regression is used: $\Delta WC = \beta_0 + \beta_1 (CFO_{t-1}) + \beta_2 (CFO_t) + \beta_3 (CFO_{t+1}) + e_t$; where ΔWC is the change in working capital and CFO is cash from operations. The model was modified to include seven years, five years before the fraud was discovered (CFO_{t-5}, CFO_{t-4}, CFO_{t-3}, CFO_{t-2}, CFO_{t-1}) and one year after (CFO_{t+1}) to capture long-term accruals. The residuals from this regression reflect the accruals that are unrelated to cash flow realizations (Dechow & Dichev, 2002). The standard deviation of these residuals is a firm-level measure of accrual quality, where higher standard deviation denotes lower quality.

To measure the persistence of earnings and cash flow from operations, this research follows Dechow & Schrand's (2004) regression of the future value of the variable on its current value: $X_{t+1} = \alpha_t + \beta(X_t) + e_t$. The closer β is to 1, the more persistent the variable X_t is. The following regressions models estimate the persistence of earnings and cash flows respectively: $Earnings_{t+1} = \alpha + \beta(Earnings_t) + e_t$; $CFO_{t+1} = \alpha + \beta(CFO_t) + e_t$; where β is the persistence parameter.

Sample

The sample's financial data was obtained from Compustat and consist of 61 firms accused by the Securities and Exchange Commission (SEC) of fraudulent financial reporting in Accounting and Auditing Enforcement Releases (AAER) issued during a ten-year period between January 2007 and December 2016. The sample is limited to firms with complete data for assets, earnings, cash flow from operations, changes in account receivable, and change in inventory. Excluded from the sample are Financial firms (SIC codes 6000-6999) and firms without complete data. Of the total sample of 303 firms, 51 financial firms were excluded because of the regulation of such firms along with 104 CPA's firms and individuals accused in the sample period, 5 releases involving the same firm and 82 firms with no or incomplete data. A group of non-accused firms (control firms) was used to determine differences in accruals and earnings quality. The control firms are of the same size and industries as the accused firms. That results in a final sample of 122 firms. See Table 1.

Table 1
Companies accused by the Securities and Exchange Commission from 2007-2016, Excluded Firms and Final Sample for Research

Accounting and Auditing Enforcement Releases between 2007 and 2016	303
Financial Firms	(51)
Auditing and other accusations not related to research	(104)
Releases involving the same firm	(5)
Companies with no or incomplete data	(82)
Total accused companies by SEC included in the Sample	61
Selected companies (competitors) for comparison	61
Final sample	122

The Appendix shows descriptive statistics and correlation for accused firms from 2007 to 2016.

RESULTS

According to our sample, the industry with the highest total of accusations is Business Services (SIC code 73) with 10 accusations; Electronic equipment except computers (SIC code 36) with 6 accusations follow by Food and Kindred Products (SIC code 20) with 5. Based upon information included in the AAERs we classify the methods used to deceptively reporting the financial statement information and the SEC definition of the alleged violation; the two most common techniques used to deceive investors were overstated earnings (64%) and the presentation of False Financial Statements (18%).

The most common techniques used to deceive investors (overstated earnings and false financial statements) take over six years to be detected. However, for 7 firms in the sample, the time between fraud year and accusation year was more than 10 years; for 17 firms, the time was between 7 and 9 years and for 37 firms, the time was 6 years or less.

Considering the Auditor's Opinion on the Financial Statements, all accused companies received an Unqualified Opinion, and 18 firms received an Adverse Opinion on internal control due to a material weakness. Two firms showed doubts about their ability to continue as a going concern, and two did not receive an opinion on internal control. This is consistent with the argument that the main objective of an audit is to issue a statement or an opinion as regards the quality or integrity of the company's reported financial information and not to detect fraud.

As shown in Table 2, the mean for total accruals of accused firms is -0.07 compared with a -0.066 for control firms. That result could be expected since a high percentage of these companies were accused of overstated earnings.

Table 2: *Descriptive Statistics Total Accruals Modified Jones Model (Accused companies)*

Mean	-0.070
Std Error	0.021
Median	-0.051
Std Deviation	0.162
Control companies	
Mean	-0.066
Std Error	0.018
Median	-0.044
Std Deviation	0.145

Descriptive statistics for the variables used in the Jones Model to determine non-discretionary accruals are shown in Table 3. Total accruals are lower for control firms (-0.088) compared with the accused firm's (0.001). The results were expected since a high percentage of the accused firms were accused of overstating revenues. Also, the data showed that changes in revenues for the accused firms are greater than the control sample. The discretionary accruals are higher for accused firms compare to the control sample. This result can be expected since discretionary accruals often provide managers the opportunities to manipulate earnings. Also, can be interpreted as the control companies having better earnings quality than the accused companies, since more accruals indicate more estimation and errors of estimation, and therefore lower quality of accruals (Dechow & Dichev, 2002). Correlations for both sub-samples are shown in Table 4.

Table 3: *Descriptive statistics of Jones Model Variables for Accused and Control Firms*

Accused firms	TA/Ai t-1	1/Ai t-1	∆REV/Ai t-1	PPE/Ai t-1
Mean	0.001	0.013	0.784	0.664
Std Error	0.002	0.006	0.481	0.287
Median	0.000	0.001	0.078	0.356
Std Deviation	0.019	0.047	3.757	2.241
Minimum	-0.051	0.000	-0.424	0.007
Maximum	0.140	0.286	29.280	17.783
Control Firms				
Mean	-0.088	0.010	0.016	0.232
Std Error	0.087	0.094	0.053	0.026
Median	0.000	0.002	0.092	0.153
Std Deviation	0.677	0.735	0.412	0.200

Table 4: *Pearson Correlation for Variables of Accused and Control Firms for Jones Model*

Control Firms

Variables	TA/Ai t-1	1/Ai t-1	∆REV/Ai t-1	PPE t/Ai t-1
TA/Ai t-1	1			
1/Ai t-1	-0.986	1		
∆REV/Ai t-1	0.346	-0.344	1	
PPE t/Ai t-1	0.093	-0.038	-0.071	1

Correlations (Accused firms)

Variables	TA/Ai t-1	1/Ai t-1	∆REV/Ai t-1	PPE/Ai t-1
TA/Ai t-1	1			
1/Ai t-1	0.506	1		
∆REV/Ai t-1	0.001	0.028	1	
PPE/Ai t-1	-0.065	-0.027	0.977	1

∆REV= changes in revenues, PPE = property, plant and equipment, Ai = total assets.

Notice the negative relation between total accruals and Gross PPE for the accused sample. A negative relation means that as one variable increases, the other decreases. In this case as total accruals increases, gross property, plant and equipment decreases. The decreasing account is most likely be the account of net property, plant and equipment since the adjusting entry increases an expense account (depreciation expense) and a contra-asset account (accumulated depreciation) decreasing the book value or the net property, plant and equipment. The positive correlation between total accruals and gross PPE for the control firms resembles the normal operations of a business. Another atypical correlation arises between ∆REV and total accruals for the accused firms. The 0.001 coefficient shows almost no relation between variables. But keep in mind that among the accused sample, there is a high percentage of overstated-earnings accusations, so this coefficient at least should show a medium to high correlation. Also, there is a high correlation between PPE and ∆REV for the accused sample giving the impression of a high fixed-assets turnover ratio. This ratio is used by analysts to measure operating performance, specifically measures a company's ability to generate net sales from fixed-asset investments.

For model developing, residuals from a firm-level time-series regression are obtained to determine the quality of the earnings, where the standard deviation of these residuals is a measure of accrual quality where higher standard deviation denotes the lower or poor quality of earnings. Descriptive statistics and correlations for our sample of accused firms are provided in Tables 5 and 6.

Table 5: *Descriptive Statistics Accused Companies (Pool)*

	∆WC	Earn	Prof	Accruals	CFO t_{-1}	CFO t_0	CFO $_{t+1}$	TA $_t$
Mean	0.021	0.047	-0.043	-0.068	0.060	0.056	0.018	6,136
Median	0.027	0.101	0.031	-0.034	0.070	0.066	0.060	1,101
Std Deviation	0.115	0.411	0.567	0.259	0.249	0.294	0.393	13,675

All variables scaled by average total assets

An examination of the descriptive statistics shows that the results are in line with previous research with comparable variables (e.g., Barth et al. 2001, Dechow & Dichev, 2002). Earnings are less than $CFO_{t,}$ suggesting that accruals are mostly negatives. Accruals are negative mainly because of depreciation.

Table 6: *Correlations Accused Companies (Pool)*

	∆WC	Earn	Prof	Accruals	TA $_t$	CFO $_{t-1}$	CFO $_t$	CFO $_{t+1}$
∆WC	1							
Earn	0.705	1						
Prof	0.567	0.952	1					
Accruals	0.566	0.811	0.93	1				
TAt	-0.043	0.055	0.076	0.06	1			
CFO $_{t-1}$	0.439	0.884	0.856	0.665	0.003	1		
CFO $_t$	0.452	0.922	0.877	0.658	0.055	0.956	1	
CFO $_{t+1}$	0.449	0.926	0.916	0.734	0.067	0.912	0.957	1

All variables scaled by average total assets

There is a positive relation between Earnings (Earn) and Cash flow from Operations ($CFO_{t\,0}$) (0.92), between Earnings and ∆WC (0.71) and between Accruals and ∆WC (0.57). Tables 7 and 8 show descriptive statistics and correlations for a sample of control companies. As with the accused companies' sample, earnings are lower than CFO and accruals are negatives. ∆WC is negatively correlated with CFO_{t+1} and $CFO_{t,}$ and CFO_t is positively correlated with CFO_{t+1}. These results agree with previous researches.

Table 7: *Descriptive statistics (Control Firms)*

	CFO_t	$∆WC_t$	$Earn_t$	$Prof_t$	$Accruals_t$	$TA_{t\ (in\ million)}$
Mean	-0.061	-0.199	-0.162	-0.018	-0.031	3,995
Median	0.099	0.014	0.098	0.044	-0.054	587
Std Deviation	0.842	1.605	1.624	0.223	0.364	9,391

All variables scaled by average total assets

Correlations are in line with Dechow & Dichev (2002), CFO is negatively correlated to ∆WC and accruals and positively correlated to CFO_{t+1}, CFO_{t-1} and $Earn_{t+1.}$ Accruals are positively correlated with Profits and ∆WC as expected and negatively correlated with CFO_t, CFO_{t-1}, and CFO_{t+1}. Also, accruals are negatively correlated with earnings $_t$ and earnings $_{t+1}$ suggesting income-decreasing accruals such as depreciation.

Table 8: *Pearson correlations of control companies*

	∆WC	Accruals	CFO	CFO $_{t-1}$	CFO $_{t+1}$	Earn	Earn $_{t+1}$	Prof
∆WC	1.000							
Accruals	0.051	1.000						
CFO	-0.032	-0.001	1.000					
CFO $_{t-1}$	-0.018	-0.029	0.002	1.000				
CFO $_{t-1}$	-0.002	-0.703	0.034	0.070	1.000			
Earn	0.971	-0.154	-0.026	-0.001	0.224	1.000		
Earn $_{t+1}$	-0.083	-0.694	0.048	0.074	0.984	0.141	1.000	
Prof	-0.053	0.106	0.019	0.055	0.439	0.019	0.430	1.000

All variables scaled by average total assets

To obtain the residuals needed to measure the quality of earnings of accused firms the following regression was run:

$$\Delta WC = b_0 + b_1 *CFO_{t-5} + b_2 * CFO_{t-4} + b_3*CFO_{t-3} + b_4 * CFO_{t-2} + b_5 * CFO_{t-1} + b_6 * CFO_{t\,0} + b_7 * CFO_{t+1} + e_t$$

Where $CFO_{t-5,}$ CFO_{t-4}, CFO_{t-3}, CFO_{t-2}, and CFO_{t-1} are cash flows from operations before the companies engaged in fraud. $CFO_{t\,0}$ is the fraud year while CFO_{t+1} is cash flow from operations the year after the fraud was perpetrated. Accruals are negatively related to current cash flows and positive related to past and future cash flows (Dechow & Dichev, 2002), results showed a different sign in the current cash flow ($CFO_{t\,0}$) coefficient (0.051), that may be explained by our sample of accused firms. Results on CFO_{t-1} and CFO_{t+1} are in line with past research.

The standard deviation of the residuals was 0.089 for our sample of accused firms. We compared the results of the accused sample with a sample of 65 companies that have not been accused. The standard deviation of the residuals was .058 showing that control firms have better earnings quality than accused firms. To validate the results of the earnings quality coefficient of our model and evaluate its consistency we compared our results with another widely used model, Francis *et al*. (2005), which utilized the Dechow & Dichev approach augmented with two regressors, the results showed a slightly higher coefficient (0.092) for the Francis *et al.* (2005) model compared to our model (0.089), showing the validity of our outcomes and the strength of the results with respect to a different definition of earnings quality. The main objective of this research was to demonstrate if firms with low-quality earnings are more prone to engage in earnings management or financial reporting irregularities. According to our regression results, a measure of less than 0.058 was to be considered a good quality measure, anything above that would be medium to poor quality. Twenty-six percent of the accused firms showed poor quality, while sixty-four percent exhibited good quality and ten percent medium quality. Just eight percent of the control firms showed poor quality, while eighty-seven percent presented good earnings quality, and five percent medium quality.

Persistence of earnings is measured as the persistence of profitability (i.e., ROA) since the metric used for measuring earnings persistence is earnings deflated by some measure of assets such as average total assets. Table 9 shows descriptive statistics for Return on Assets for our sample of accused companies and control companies.

Table 9: *Descriptive statistics for Return on Assets of Control and Accused Firms*

Descriptive Statistics (Control Firms)	Total Assets	ROA
Mean	4,012	0.005
Median	587	0.046
Std Deviation	9,388	0.202

Descriptive Statistics (Accused Firms)	Total Assets	ROA
Mean	6,136	-0.036
Median	1,101	0.032
Std Deviation	13,675	0.530

The return on assets of the control companies is higher than the return on assets of accused sample. The returns are as high as 0.40 on our control sample compared to a high of 0.155 on the accused sample. This means that in general, control companies are more effective in converting the money they invest in net income.

Table 10 provides results for future earnings on current earnings and Cash Flow from Operations and Accruals for our sample of accused companies. We can interpret these results as follows: for each dollar a company earns, $0.36 cents will persist into next year earnings. For each dollar of CFO, only 1.5 cents will persist into next year's earnings. For each dollar that represents accruals $-0.001 persist into the following year. The negative sign in accruals could be because a very high percentage of the accused companies manipulated earnings. These results show that earnings backed by cash flows are more persistent that earnings that represent accruals. This is consistent with Dechow and Schrand (2004).

Table 10: *Regressions results of Future earnings on current earnings and CFO and accruals (Pool Accused Companies)*

	Earnings $_t$		Accruals and Cash Flows	
Variable	Coefficient	t-statistics	Coefficient	t-statistics
Intercept	-0.014	-0.79	-0.059	-2.537
Earnings	0.359	24.134		
Accruals			-0.001	-2.149
CFO			0.015	2.981

Source: Adapted from Dechow and Schrand (2004). All variables scaled by average total assets.

As shown in table 11, for the pool control companies, $0.59 will persist into next year earnings for every dollar the company earns, representing a 64% increase over the accused companies. Also, for each dollar that represents accruals $0.32 persist into the following year. Consistent with Sloan (1996), Dechow & Dichev (2002), and Dechow & Schrand (2004), results show that earnings performance attributable to the accrual component of earnings exhibits lower persistence than earnings attributable to Cash Flows from Operations.

Table 12 shows the analysis of future earnings and future cash flows from operations on current earnings and current cash flows from operations from our pool sample of accused companies;

the firms were grouped into four quartiles. The first quartile contains firms with high negative accruals and fourth quartile groups companies with high positive accruals. Earnings are less persistence than CFO in the first, third, and fourth quartiles, while in the second quartile earnings are more persistence that CFO. In line with Sloan, (1996), Dechow & Dichev, (2002), and Francis et al. (2005), persistence of earnings is reduced with larger accruals of either sign, which is the case in 1^{st} and 4^{th} quartiles.

Table 11: *Regressions of Future earnings on Current Earnings and CFO and Accruals (Pool Control Companies n=165)*

	Earnings $_t$ Only		Accruals and Cash Flows	
Variable	Coefficient	t-statistics	Coefficient	t-statistics
Intercept	-0.103	-1.278	-0.176	-2.724
Earnings	0.594	13.895		
Accruals t			0.324	5.995
CFO t			2.202	12.801

All variables scaled by average total assets

Table 12*: Persistence of accused companies by quartiles*

	1st quartile	2nd quartile	3rd quartile	4th quartile
Earnings	0.627	1.07	0.635	0.612
Accruals	-0.00072	0.672	-0.569	-0.002
CFO	1.01	0.001	0.668	1.11

The test for the persistence of earnings and cash flow from the data revealed that Cash Flow from Operations is more persistence that earnings in both control and accused companies. These results may be perceived inconsistent with past researches that alleged that earnings are more persistence that cash flows, on average, but large accruals significantly reduce the persistence of earnings. Results show that there is a difference between the persistence of earnings of accused companies and the persistence of control companies, overturning our statement that there is no difference in the persistence of earnings between companies with high or low earnings quality. Cash Flows from Operations of control companies shows a higher coefficient than the accused firms. This result validates our statement that there is a difference in the persistence of cash flows between companies with high or low earnings quality.

CONCLUSIONS

The accounting discipline has changed in the practical and theoretical backgrounds at the beginning of the twenty-first century. The resulting environment raises a host of questions that are of concern to academics, regulators, and practitioners, including the way companies manage earnings in order to present financial health that differs from reality. Managers try to achieve their goals through aggressive accounting procedures to demonstrate the company's financial strength. The external pressures from shareholders, financial analysts, and credit rating agencies influence managers to incur in deceitful practices to inflate earnings and increase ~~stocks~~ prices sometimes for their benefit. The data showed that the years 2007 and 2009-2010, the pre-and post-years of the Credit Crisis present the highest percentage of accusations announcements.

Total accruals were calculated for 61 accused firms and 61 non-accused or control firms. To determine which companies are more prone to earnings management or financial statements fraud, we used the Jones Model (1991) to determine non-discretionary accruals. Control firms showed a higher coefficient for non-discretionary accruals (0.004) compare to a (-0.049) for accused firms; the non-discretionary component reflects the business condition that naturally creates and destroy accruals. We can infer from these results that a high percentage of the accruals created by the control firms were due to their normal operations.

The residuals from the regression reflect the accruals that are not related to cash flows from operations, and the standard deviation of those residuals is a firm-level measure of accrual or earnings quality. A higher standard deviation represents lower quality, and a lower standard deviation denotes higher quality. The residuals from the regression reflect the accruals that are not related to cash flows from operations, and the standard deviation of those residuals is a firm-level measure of accrual or earnings quality. A higher standard deviation represents lower quality, and a lower standard deviation denotes higher quality. The standard deviation of the residuals for our sample of accused firms was 0.089, while for the control firms was 0.058 showing that control firms have better earnings quality than accused firms.

Accused firms showed higher discretionary accruals than control firms; this is consistent with the sample characteristic that 68% of the firms were accused of overstating earnings or assets. Earnings management literature agreed that a higher number of accruals denotes the poor quality of earnings, so we expected accused companies to have a poor quality of earnings while expecting a high percentage of the control companies or non-accused firms to have good earnings quality. The latter was the case when eighty-seven percent of non-accused firms showed good earnings quality, while twenty-six percent of the accused firms showed poor quality.

Our results for future earnings on current earnings and Cash Flow from Operations and Accruals for our sample of accused companies showed that for each dollar a company earns, 0.36 cents would persist into next year earnings. For each dollar of CFO, 1.5 cents will persist into next year's earnings. For each dollar that represents accruals $-0.0001 persist into the following year. The negative sign in accruals could be because a very high percentage of the accused companies manipulated earnings. For future earnings on current earnings and Cash Flow from Operations and Accruals for the sample that for control companies $0.59 will persist into next year earnings for every dollar, the company earns, representing a 64% increase over the accused companies. Also, for each dollar that represents accruals $0.32 persist into the following year. Our findings are consistent with Sloan (1996), Dechow & Dichev (2002), and Dechow & Schrand (2004), and show that earnings performance attributable to the accrual component of earnings exhibits lower persistence than earnings attributable to Cash Flows from Operations.

The difference between the persistence of earnings of accused companies and the persistence of control companies overturns our statement that there is no difference in the persistence of earnings between companies with high or low earnings quality. Cash Flows from Operations shows different coefficients for both samples. The coefficient for control or non-accused firms is higher (1.33) than the coefficient for our accused sample (1.03) validating our statement that there is a difference in the persistence of cash flows between companies with high or low earnings quality.

As general findings, our research showed that the industry with the highest total of accusations was Business Services (SIC code 73) with 16% of the accusations. This finding agrees with prior similar research (i.e., Dechow *et al*. 1996). Also, 64% of the accused companies overstated earnings, being this violation, the most common technique used to deceive investors, taking an average of 6.78 years for the company to be accused. This result also agrees with the *Fraudulent Financial Reporting 1998-2007,* where 61% of the firms (n=347) were accused of some type of improper revenue recognition. Our results indicate that nor the firms (SIC 73) or the technique used to deceive investors (improper revenue recognition) have not changed much in the past 22 years; we can consider this a new finding in the literature.

An important and unexpected finding was the high percentage (100%) of Unqualified Opinions on Financial Statements for the sample of accused firms. We expect at least some opinions to be qualified, or even a disclaimer. One reason to issue a qualify opinion could be that the company did not present its financial records in accordance with GAAP, which was the case in some of these firms. Only nineteen percent (19%) of the accused firms received an Adverse Opinion on internal control due to a material weakness. This should have raised red flags on those companies. The Sarbanes-Oxley Act requires all financial reports to include an Internal Controls Report that shows that a company's financial data are accurate (within 5% variance) and adequate controls are in place to safeguard financial data. Results also show that out of 61 firms accused, 80% of the companies kept operating, 12% were either sold or privatized, and 8% merged.

This study corroborates prior empirical studies and updates the earnings management, earnings quality, and earnings persistence literature. Consistent with previous research (Sloan,1996), Dechow & Dichev (2002), and Dechow & Schrand (2004), our results validate that earnings performance attributable to the accrual component of earnings exhibits lower persistence than earnings attributable to Cash Flows from Operations. Similarly, our findings show that in general, control or non-accused companies are more effective in converting the money they invest in net income than accused firms or firms with low earnings quality, or in other words, control firms have a better Return on Assets percentage than accused firms. Nonetheless, earnings are more persistence for control firms than for accused firms.

Our results also showed firms that have not been accused of any wrongdoing showed a higher coefficient for non-discretionary accruals compare to the accused firms, meaning that a high percentage of the accruals created by the control firms were due to their normal operations. But accused firms have higher discretionary accruals than our control firms signaling some type of manipulation or lower earnings quality.

Finally, a numerical scale is suggested for expressing the earnings quality measure of firms. This scale can be used as a guide to determine good or poor earnings quality in any public firm that a user can be interested in and could serve as a complement to the SEC's Accounting Quality Model to eradicate potential fraud or simply identify aggressive uses of the discretion allowed under GAAP. We used the standard deviation of control firms (0.058) to develop an earnings quality scale. If a data distribution is normal then about 68 percent of the data values are within one standard deviation of the mean (mathematically, $\mu \pm \sigma$, where μ is the arithmetic mean), about 95 percent are within two standard deviations ($\mu \pm 2\sigma$), and about 99.7 percent lie within three standard deviations ($\mu \pm 3\sigma$) because 95 percent is the most commonly used interval we can argue that a good earnings quality can be any measurement below 0.058 based on our results. A measurement of 0.074 or higher can be view as a poor earnings quality.

LIMITATIONS

Our model for earnings quality is based on Dechow & Dichev (2002); the model was modified to include seven years instead of the original three years. As with the original model, the independent variables are measured with error, implying that the regression coefficients are likely to be biased toward 0, and the R^2 will be reduced. Also, the bias is more severe for b_1 and b_3 as compared to b_2 (Dechow & Dichev, 2002). The final sample size might be a limitation too, from an original 303 accused firms, 61 had the required data. However, the results were comparable with prior research.

One constraint common in prior works on persistence models is the assumption that the coefficient on accruals and cash flows is equal in the absence of accrual measurement error (Richardson et al., 2005). That same constraint is present in our persistence model also.

FUTURE RESEARCH

In this research, the focus was on earnings quality, earnings, and cash flow persistence as an indicator of earnings management or accounting fraud. A partially functioning model was developed based on earnings quality but might need to be developed further to increase its precision and accuracy. Also, some non-financial variables such as personality, optimism, aggressiveness, and other related traits can be an interesting addition to the model, since tone at the top sometimes dictates the organizational culture, and we need to understand what drives these executives to report misleading information.

REFERENCES

Artikis, P. G., & Papanastasopoulos, G. A. (2016). Implications of the cash component of earnings for earnings persistence and stock returns. *The British Accounting Review*, *48(2)*, 117-133. https://doi.org/10.1016/j.bar.2016.02.002

Barth, Mary E., Donald P. Cram, Donald P., and Nelson Karen K. (*2001*) Accruals and the Prediction of Future Cash Flows. The Accounting Review: January 2001, Vol. 76, No. 1, pp. 27-58.

DeAngelo, L. E. (1986). Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. *Accounting Review*, 400-420.

Dechow, P. M., Larson, C. R., & Resutek, R. J. (2019). Discretion Lost? Accrual Heterogeneity and Implications for Accrual Quality Inferences. *Accrual Heterogeneity and Implications for Accrual Quality Inferences*

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *Accounting review*, 193-225.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary accounting research*, *13*(1), 1-36.

Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons*, *14*(2), 235-250.

Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The accounting review*, *77(s-1)*, 35-59.

Dechow, P. M., & Schrand, C. M. (2004). Earnings quality.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of accounting and economics*, *50*(2), 344-401

Deo, P. L., & Cathy, Z. S. (2016). All cash is not created equal: Detecting fraudulent cash flow. *Journal of Forensic and Investigative Accounting*, *8*(2).

Dichev, I. D., Graham, J. R., Harvey, C. R., & Rajgopal, S. (2013). Earnings quality: Evidence from the field. *Journal of Accounting and Economics*, *56*(2), 1-33.
Dichev, I. D., & Tang, V. W. (2008). Earnings volatility and earnings predictability. *Journal of Accounting and Economics*, *47*(1-2), 160-181.
Denis, D. J. (2012). The persistent puzzle of corporate capital structure: current challenges and new directions. *Financial Review*, *47*(4), 631-643.
Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of accounting and economics*, *39*(2), 295-327.
Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of accounting and economics*, *7*(1-3), 85-107.
Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, *13*(4), 365-383.
Jiambalvo, J. (1996). Discussion of "Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC". *Contemporary Accounting Research*, *13*(1), 37-47.
Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of accounting research*, 193-228.
Juma'h, A. H. (2019). Behavioral Elements Related to Consideration and Use of the Materiality Concept in Accounting Practices, *Accountancy Business and Public Interest*, 18(1), 144-161.
Juma'h, A. H. (2014). The Materiality Concept: Implications for Managers and Investors, *Revista Finanzas y Política Económica*, 6(1), 159-168
Khan, U., Peddireddy, V., & Rajgopal, S. (2019). Earnings Quality on the Street. *Available at SSRN 3347217*.
Kaplan, R. S., & Roll, R. (1972). Investor evaluation of accounting information: Some empirical evidence. *The Journal of Business*, *45*(2), 225-257.
Keefe, T., *Earnings Quality: Measuring the discretionary portion of accruals*. *Investopedia* Investopedia US, (n.d.). Web. 17 June 2015
Kubota, K., & Takehara, H. (2019). Measuring the Persistency of Earnings Components: Applications of the VAR Model to Long-Run Japanese Data. *Journal of Accounting, Auditing & Finance*, *34*(2), 329-342.
Mummolo, J., & Peterson, E. (2018). Improving the Interpretation of Fixed Effects. *Political Science Research and Methods*, 6(4), 829-835.
Nakashima, M., & Ziebart, D. A. (2016). Tone at the top and shifts in earnings management: Evidence from Japan. *Journal of Forensic & Investigative Accounting*, *8*(2), 288-324.
Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of accounting and economics*, *39*(3), 437-485.
Schipper, K., & Vincent, L. (2003). Earnings quality. *Accounting Horizons*, *17*, 97-110.
Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?. *Accounting review*, 289-315.
Statement of Financial Accounting Concepts No. 2 Qualitative Characteristics of Accounting Information
Sun, L., & Rath, S. (2010). Earnings management research: a review of contemporary research methods. *Global Review of Accounting and Finance*, *1*(1), 121-135.
Warshavsky, M. (2012). Analyzing earnings quality as a financial forensic tool. *Financial Valuation and Litigation Expert Journal*, *39*, 16-20.

APPENDIX

List of variables and descriptive statistics for sample data
The following table shows variables, descriptive statistics, and correlation for our sample of accused firms from 2007 to 2016. Where CFO is cash flow from operations, IBEI is income before extraordinary items, A/R is account receivables, CPLTD is current portion of long-term debt, TA is total assets, PP&E is property, plant, and equipment, TL is total liabilities, Dep&Amor is depreciation and amortization, DIV is dividends, COGS is the cost of goods sold, and GPM is gross profit margin.

Summary of Variables and Descriptive Statistics

VARIABLES	MEAN	MEDIAN	STD DEV	MINIMUM	MAXIMUM
CFO	1,045.32	73.71	5,142.06	-663.20	48,601.00
IBEI	486.86	32.59	2,522.92	-5,467.00	22,468.00
Cash	1,168.64	205.31	5,933.70	0.05	82,563.0
A/R	7,928.90	200.28	47,844.76	0.66	418,777.00
Inventory	947.45	175.67	2,645.29	0.00	23,925.00
Taxes payable	22.92	1.68	52.76	-59.00	286.40
CPLTD	1,223.72	1.97	8,102.79	0.00	101,532.00
Current assets	10,258.12	604.02	55,651.75	5.66	472,049.00
TA	18,710.06	1,324.15	100,431.06	6.15	797,769.00
PP&E	2,099.12	146.38	9,340.66	0.30	78,530.00
TL	15,219.95	649.45	85,128.07	3.18	684,157.00
DEP&AMOR	325.89	35.29	1,317.63	0.00	11,492.00
DIV	228.26	0.00	1,287.04	0.00	12,408.00
Sales	8,935.74	1,487.84	24,843.45	8.44	180,929.00
COGS	6,322.17	996.93	16,173.19	4.67	96,605.00
GPM	30.87	29.53	28.23	-395.92	86.77
Gross profit	2,614.24	349.57	11,474.22	-376.06	107,527.00

Correlations

	CFO	*IBEI*	*Cash*	*A/R*	*Inventory*	*Taxes payable*	*CPLTD*	*Current assets*	*TA*	*PP&E*	*TL*	*DEP&AMOR*	*DIV*	*Sales*	*COGS*	*GPM*	*Gross profit*
CFO	1																
IBEI	0.952	1															
Cash	0.755	0.668	1														
A/R	0.981	0.943	0.804	1													
Inventory	0.892	0.875	0.679	0.889	1												
Taxes payable	-0.009	-0.001	-0.002	-0.055	0.137	1											
CPLTD	0.923	0.847	0.928	0.949	0.821	-0.050	1										
Current assets	0.975	0.930	0.848	0.996	0.892	-0.039	0.969	1									
TA	0.981	0.941	0.817	0.997	0.890	-0.034	0.950	0.996	1								
PP&E	0.978	0.938	0.798	0.991	0.911	0.020	0.939	0.989	0.991	1							
TL	0.981	0.938	0.814	0.996	0.884	-0.040	0.949	0.994	0.999	0.989	1						
DEP&AMOR	0.970	0.919	0.825	0.979	0.897	0.048	0.941	0.982	0.983	0.992	0.982	1					
DIV	0.973	0.950	0.724	0.981	0.886	-0.028	0.888	0.967	0.974	0.976	0.973	0.961	1				
Sales	0.867	0.835	0.704	0.858	0.929	0.094	0.804	0.865	0.868	0.870	0.867	0.874	0.847	1			
COGS	0.652	0.632	0.508	0.626	0.777	0.128	0.570	0.635	0.647	0.642	0.649	0.656	0.621	0.930	1		
GPM	0.104	0.116	0.074	0.097	0.044	0.032	0.097	0.094	0.094	0.095	0.089	0.093	0.104	-0.021	-0.118	1	
Gross profit	0.958	0.918	0.809	0.975	0.916	0.024	0.937	0.978	0.967	0.979	0.962	0.968	0.960	0.855	0.603	0.122	1

DECISION SCIENCES INSTITUTE

The Making of a President Using Data Analytics and Social Media

Ed Lindoo, Ph.D
Regis University
3333 Regis Blvd.
Denver, CO 80221-1099
303-964-6358
elindoo@regis.edu

ABSTRACT

Today, political campaigns rely heavily on analytics to target potential voters. The algorithms used, as well as bias that may be introduced in the process, may skew the results and cost an election, as seen in the Clinton campaign. Additionally, Social Media plays a huge role in political races, as does fake news. More alarming is the new trend for Social Media sites to censor anyone or anything which can also make or break a campaign. In this paper we research events that affected the 2016 Presidential election and present issues that may have an effect on future elections.

KEYWORDS Big Data, Bias, Analytics, Fake News, Social Media

INTRODUCTION

In Data Sciences and Computing in general, bias can be introduced in a variety of ways. Conscious or subconscious prejudice, bias, can be found in the data collection phase, analysis phase, and insight phase. Bias can also be found in the outcome phase, assessment phase and improvement phase. Analysts must be aware of bias and cautious to prevent it in all phases of their analysis. Analytics and ethics are not necessarily enemies but can be if not properly treated. For example, releasing data gained through analytics, whether it is technically correct or not, could result in a company's loss of their reputation, create competitive weakness, or possibly result in legal sanctions (Noyes, 2015). Big data and analytics extend well beyond organizational benefits and Gartner (Noyes, 2015) predicts that by 2020, big data analytics will cause half of all business violation ethics. The use and or misuse of data also holds true in political analytics.

Today, companies world-wide use big data in many meaningful ways, and recently, so has our political system. In fact, during the 2008 and 2012 elections, the Obama campaign relied heavily on analytics to outperform their Republican counterparts. Going into the 2016 elections, both sides had come to realize that there had been a major shift in how people communicated, moving away from landline phones and going more for mobile phones and Social Media. As a result, it came as no surprise that both the Democratic and Republican campaigns would follow what Obama had done and use analytics in a big way for the 2016 election (Enderle, 2016). With so much data, what went wrong for the Clinton campaign and what went right for the Trump campaign?

LITERATURE REVIEW

Introducing Bias

The Book “Shattered” shares a perspective on why Clinton lost the 2016 election. Highlights include Clinton’s campaign manager, Robby Mook, who put a lot of faith into a self-developed, super analytic system called Ada. Mook felt that Ada was the campaigns “secret weapon." As for Mook’s algorithm, when approached, Mook’s team said, “we’re the smartest guys around here so let us do our thing” (Jacknis, 2017). The Ada algorithm played a role in just about every strategic decision made by the campaign, including where and when to send the candidate, her staff, and where to place TV ads. Interestingly, Ada was able to run about 400,000 simulations a day, providing a detailed picture of which battleground states would be the tipping point. (Wagner, 2016). Sometimes too-much is too much. Because it’s humanly impossible to sift through 400,000 simulations every day, staff had to rely on Ada to produce summary results of these simulations.

One of the things that Ada missed was the states of Michigan and Wisconsin. The Clinton campaign had paid little attention to these states, assuming they were in the bag, and it wasn’t until just a week before the election that staff realized they had a problem. It appears that Ada had been programmed to underestimate the power of rural voters. The campaign then tried a last-minute ad campaign and a tour of these states with Clinton and the Obama’s, but it was simply too late. (Wagner, 2016)

Was there bias? According to Aurerbauch (2017), Ada was developed by dozens of researchers who were led by Clinton’s director of analytics, Elan Kriegel, in close consultation with campaign manager Robby Mook. Inputs to Ada included polls and surveys as well as the campaigns data from field workers and it’s possible that bias could have been introduced at these points.

The UPB Phenomena

At the beginning of the 21st century, accounting scandals, among other events, led to the collapse of billion-dollar companies. These scandals only confirmed that unethical acts were being conducted, often flourishing within organizations. Research into unethical behavior within organizations has found assorted reasons as to why employees might engage in unethical acts, for example, personal gain, the gain of the organization, or even to cause harm against the organization. (Umphress, 2010). These acts fall under the title of unethical pro-organizational behavior. UPB includes acts of commission (e.g., “cooking” numbers to boost analyst projections and stock values) and omission (e.g., withholding information about the hazards of a pharmaceutical product) that are typically considered unethical.

UPB is predominantly considered pro-organizational behavior which is neither specified in formal job descriptions nor ordered by superiors, yet, is carried out to benefit or help the organization. Many theorists believe that the more an individual identifies with their organization, the more likely they are to disregard their personal moral standards and engage in acts that favor the organization. Further research has found that the stronger the employees’ organizational identification, the more influence it will have on productive work behavior, such as increased extra-role behaviors, and job performance. Umphress et al., believe that individuals who identify strongly with their organization may be more likely to engage in UPB. (Umphress, 2010).

As a public example, we have recently noted journalists who have been found to be lying or “bending” the truth. In 2015, NBC Nightly News anchor, Brian Williams, was caught lying about riding in a helicopter that was shot down in Iraq in 2003. NBC suspended him for six months without pay and the Peabody Award-winning newsman, became the top story himself. In 2015, Fox News Channel star, Bill O'Reilly, was caught lying about experiencing war involving Argentina and England in 1982. In 2011, MSNBC's Rachel Maddow, was caught lying when she accused Rush Limbaugh of racism. She was later forced to apologize to Limbaugh on the air. Be it public figures, or an analyst within an organization, people tend to get caught up in the moment, the organization, personal gain, or a number of reasons that might fall under UPB. Imagine then, how an analyst could add an extra day of revenue to the monthly or quarterly reports for a company. Depending on their position within the company, unethical behavior may be committed (Umphress, 2010).

What does the UPB phenomenon have to do with political analytics? Probably more than we know! In fact, Aurerbach (2017) determined that Clinton analysts did have bias, and programmed that into Ada, assuming that Michigan, Wisconsin and Pennsylvania were a "lock". As Aurerbauch (2017) points out, "If a piece of code crashes, it's broken, but at least you know it's broken. The most dangerous kind of code—as I learned too many times in my years as a software engineer at Google and Microsoft—is the kind that breaks but appears to keep working. The worst part is that you have only yourself to blame" Aurerbauch (2017). In essence this was Ada's failure……..it went wrong early and no one caught it.

Passive Bias

Enderle (2016) found another type of bias was introduced, passive bias. As the campaigns matured prior to the November elections, things heated up and Clinton began a very negative campaign against Trump, one that would see her call his followers "deplorables". This caused many folks on both sides to either not respond to surveys or not respond truthfully. As a result, this introduced massive pro-Clinton bias, a bias that Ada could not account for. Enderle went on to state that "As an analyst, a big part of the job is identifying and mitigating bias otherwise you are driving the people who pay you to make bad decisions and, given the outcome, that would seem to be the case here" (Enderle, 2016).

If that were not enough, even more bias was introduced by complacency. Because Ada was reporting great numbers to the Clinton campaign, Mrs. Clinton reportedly became complacent and started laying back while Trump who didn't believe any of the numbers pushed even harder. As Enderle (2016) pointed out, when people don't challenge the results they like, it almost always leads to bad results…….this certainly proved out in Clinton's case. In fact, Enderle pointed out the three rules of analytics. First, assure your data source. You must have a strong sampling methodology to obtain accurate results. Secondly, it is a must to identify and eliminate any and all bias; any bias will invalidate the results. Finally, decision makers must challenge the analysis, especially if the analysis is telling you what you want to hear.

Even more bias was introduced via polarization. Smith (2017) found that as campaigns target their supporters with data, there is less effort put forth to persuade these same supporters, causing polarization. Polarization happens because the data puts these potential voters into predetermined buckets which reflects the public's ideology and voting preferences. As Smith (2017) points out, this can be dangerous and may lead to self-induced bias.

The Use of Facebook

While the Trump campaign relied on voter data owned by the Republican party, they also engaged the services of Cambridge Analytica, using their models to help make decisions on advertising and how to better reach donors (Kirchaessner, 2017). Because the Clinton campaign was using a highly sophisticated analytical approach, Trump's campaign knew they needed to combat that with their own analytics. One of the things that helped Cambridge Analytica turn the tide was in their ability to acquire data on people who had voted early, data that it collected from local counties and states. This data was then linked to individual Facebook accounts and then analyzed to determine the political advertisements that they had been subjected to (Kirchaessner, 2017). In his book "Outnumbered" David Sumpter, a professor of applied mathematics at the University of Uppsala in Sweden, analyzed the Cambridge Analytica's models and found they were using a regression "that takes the data we already have about a person and uses them to predict something we don't know about him or her." (Rathi, 2018) However, as Rathi (2018) points out, it is quite difficult to predict a person's personality from their Facebook page using this regression model. Rathi went on to state that a better method is to use Facebooks advertising platform which enables developers to build audiences similar to a manually selected audience.

Although Trumps campaign was investigated for using Cambridge Analytica and their use of Facebook, the Clinton campaign also gathered tons of information from Facebook users by creating and distributing an app that asked users to pair their Facebook friends and families with the smartphone contacts. In this way the Clinton campaign could reach friends of friends (Jacknis, 2017)

The Analytical Focus on Social Media

It used to be that elections would almost certainly be decided by about 20 percent of voters who fell somewhere in the "ideological middle", but that logic has now disappeared as a result of big data analytics. And because of this, Todd (2017) believes that Big Data is hurting American politics, not because it's wrong or that it can't be used for good, it's the cries of "fake news" and "unskew" that are accurate.

Prior to the 1900's, local, national and world news came to us primarily by way of newspapers. In the early 1900's, radio introduced a new medium by which people obtain news. These mediums had editorial boards that scrutinized articles written by trained professionals. These professionals often went through years of college preparatory courses and internships prior to attaining a job in the world of news. Television appeared in the 1950's, and the footsteps of radio were followed when it came to news integrity. When the Internet arrived in the late 90's, so did Blogs, Facebook, Twitter, YouTube, and a host of other Social Media sites. With Social Media, came reporter "wanna-bees", with little or no training in the world of news reporting, and frequently, with little, if any, reporting ethics, presenting information that is often hard to substantiate.

Fake News

It's not known which came first, poor reporting on the Internet, or, poor reporting on the major cable networks such as CNN, MSNBC, FOX, BBC. As witnessed prior to the 2016 presidential election, everyone appears to be scrambling to be the first to report news, be it accurate or not. As a result, we regularly see, companies such as The New York Times and CNN retracting stories upon realizing the stories are incorrect. How much damage is done when a wrong (Fake News) story goes out? Researchers at MIT researched this question, and recently concluded a five-year study on "Fake News on Twitter" (Meyer, 2018).

The study found that a false story reaches 1,500 people six times quicker, on average than a true story does. They also found that false stories outperform the truth on every subject including business, terrorism and war, science and technology, and entertainment, determining that fake news on politics routinely outperform the other categories. Ultimately the researchers found that between 2006 and 2016, about 126,000 tweets had been retweeted more than 4.5 million times. Some of these were linked to "fake" stories hosted on other websites while others started rumors themselves, either in the text of a tweet or in an attached image. For example, a rumor was tweeted in February of 2016, reportedly by a recently deceased elder cousin of Donald Trump, and opposing the presidential bid, stated in his obituary, "As a proud bearer of the Trump name, I implore you all, please don't let that walking mucus bag become president". However, Snopes could not find evidence of the story and rejected it as false. Nonetheless, roughly 38,000 Twitter users re-tweeted the story. It's not known how many might have shared this in the form of an email or other medium (Meyer, 2018)

In a 2017 study of Higher Education Students, Chandra, Surjandy and Ernawaty found that higher education students can become emotional and be misguided by fake news. Alarmingly they can disseminate this fake news faster because of Social Media and instant messaging applications. Because Social Media and instant messaging are used by almost everyone who has a smartphone, to interact and communicate with just about anyone, they can quickly disseminate information, motivations, and promotions, good or bad. Through their research, it was determined that a large proportion of their sample population of Higher Education students actually modify the news and made it look legitimate by including a reliable source! (Yakob, C., Surjandy, Ernawaty, 2017).

Misinformation and Disinformation

False information comes in two forms: misinformation and disinformation. Disinformation is false information that is purposely spread to deceive. Misinformation is simply incorrect information, for example, "I was misinformed about when to meet you for lunch, however I know it was not deliberate." One study, as reported by Kshetri (2017) found that roughly 62% of US adults get their news from Social Media sites, and of those 40%, from Facebook. Kshetri (2017) also determined that in the final three months of the 2016 US presidential campaign, the top performing fake election news stories on Facebook attracted more views than the top stories from major news outlets such as the New York Times, Washington Post, Huffington Post, or NBC News. During that time 8.7 million shares, reactions, and comments were generated on Facebook as compared to just over 7.3 million from 19 major news websites. As we all know now, Russia was involved.

Russia, however, was not the only country spreading fake news. We have recently learned that other fake news creators were known to be operating from countries such as the Republic of Georgia and Macedonia. For example, during the one-year period before the 2016 US presidential election, residents of the Macedonian town of Veles (population 45,000) launched more than 140 US political websites. Most of the domain names looked American, such as WorldPoliticus.com,TrumpVision365.com,USConservativeToday.com,DonaldTrumpNews.com, and USA DailyPolitics.com Kshetri (2017).

The China Syndrome

Realizing that Social Media can sway stock markets and has the potential to overthrow governments, the Chinese central government has placed bans on many Social Media URLs. It is common knowledge that Facebook, Google, YouTube and Twitter are banned, in fact, over 8,000 sites are banned in China (Greatfire.org). Studies (Kuang, 2018) suggest that all media content harmful to the legitimate ruling of the communist state is censored. To protect its rule, the Chinese Communist Party (CCP) has identified a few priorities including the sustaining of economic growth, nationalism, social stability, and rational legal authority and electoral legitimacy. This is also what sustains the current political regime. Therefore, the news most likely to be censored is news that the propaganda authorities believe will have a negative effect on the legitimate rule of the Communist Party state. Further, Chinese law mandates that all Chinese news media must contribute to the enhancement of party ruling (Kuang, 2018).

China, however, is not the only country to ban Social Media sites. Figure 1 is a good representation of the major sites banned by numerous countries around the world (Yasaklanan and Ağlar, 2018). North Korea is not shown as everything is blocked. Until recently, Cuba was very much like North Korea, although presently more and more sites are being allowed by the Cuban government. However, all the Internet access is controlled by the Cuban government and opposition sites are blocked (San Pedro, 2016).

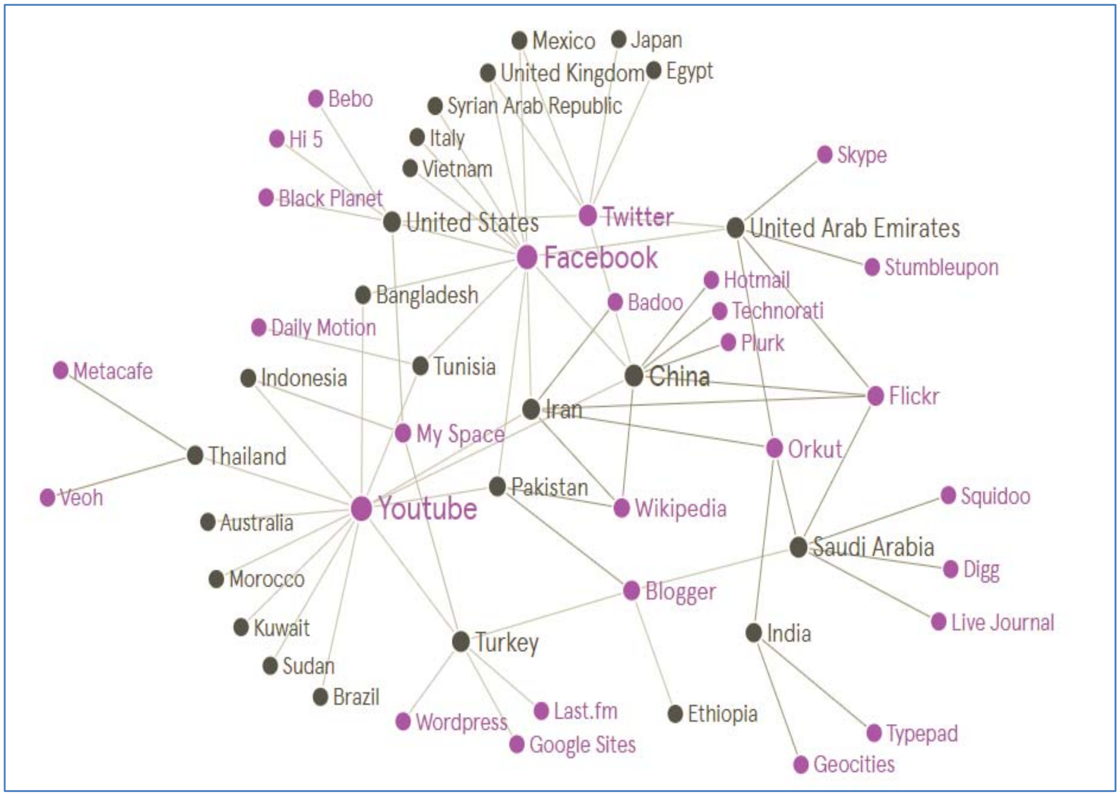


Figure 1 (Yasaklanan, 2018) a high-level view of countries that block various websites

The Push to Curtail Social Media

As we have seen, some governments simply ban Social Media sites while others such as the U.S. are either trying to pass laws to ban certain content or have asked the leaders of these organizations to do their own censoring. After the 2016 election, attention was given to the false, inflammatory and misleading information being presented on Facebook, Twitter, Google and other online services who now find themselves in a position of how to appropriately operate on a global platform (Lima, 2019). In a recent and certainly controversial move, Facebook banned the accounts of far-right extremists Alex Jones, Milo Yiannopoulos and Louis Farrakhan. To most, this seems like the right thing to do, however it begs the question, where do you draw the line. For example, when questioned by Congress, Zuckerberg stated that he that “he'd allow posts denying the Holocaust as long as the people behind them sincerely believed what they were saying” (Lima, 2019).

In a Senate hearing during April 2019, Republicans accused Facebook, Google and Twitter that they censor conservative user’s content online. They went on to threaten regulation of these and other online services. In the hearing, Senator Ted Cruz stated "Not only does big tech have the power to silence voices with which they disagree, but big tech likewise has the power to collate a person's feed so they only receive the news that comports with their own political agenda" (Romm, 2019). During this hearing it became clear that Republicans believe that Facebook, Google and Twitter are biased against the party. Political parties or not, "Every time Facebook makes the choice to remove content, a single company is exercising an unchecked power to silence individuals and remove them from what has become an indispensable platform for the speech of billions” (Lima, 2019). This should put fear into a lot of people because it is possible that just a few Social Media companies have the power to define elections in the United States.

CONCLUSION

As demonstrated in this paper, data analytics is often swayed by bias. Managers of data and analytics need to be aware of this and ensure that the findings they present are accurate and without bias. It remains to be seen if the UPB phenomenon will continue to grow, or if those in charge of data will come to realize the ethical and moral values of accurate reporting. It also remains to be seen if college students, news organizations and others will continue to spin “fake news” across Social Media sites, or if they too will come to realize the ultimate ramifications of putting out un-validated stories.

This paper presented the UPB phenomenon, pointing out that it is quite easy to fabricate data, be it for personal reasons, or to benefit a company or a political campaign. Data can be analyzed and presented in a way that may be incorrect and more importantly, analysts need to be aware of personal bias, or biases that may come from above. Finally, it was determined that Social Media and “Fake News” are presenting and will continue to present dangers to our society, especially through our young, college-level students. This danger was seen not only in the use of Social Media, but also television, radio, print and cable news. As pointed out, Fake New travels at a rate six times faster than real news, yet Fake News is often more believed and therefore opinions formed with no real basis.

With the 2020 Presidential elections closing in, it remains to be seen as to what effect analytics will again have on the election. Will campaigns be able to data mine Facebook users accounts as they did in the 2016 elections? Will they find other ways, legal or not to get to the data that they’re looking for? Will they be better able to target market those swing voters? Most importantly, will Social Media sites introduce bias or cutoff certain messages in an effort to sway the election?

REFERENCES

Auerbach, D. (2017), Confirmation Bias: Did big data sink the Clinton campaign? https://nplusonemag.com/online-only/online-only/confirmation-bias/

Bringing transparency to the Great Firewall of China. (2018). GreatFire.org https://en.greatfire.org

Enderle, R. (2016). How Trump defeated Clinton using analytics. *CIO Magazine,* https://www.cio.com/article/3138972/how-trump-defeated-clinton-using-analytics.html

Jacknis, N. (2017). Campaign Analytics: What Separates The Good From The Bad. https://startupsventurecapital.com/campaign-analytics-what-separates-the-good-from-the-bad-974fb273456a

Kirchgaessner, S. (2017). Cambridge Analytica used data from Facebook and Politico to help Trump. The Guardian,https://www.theguardian.com/technology/2017/oct/26/cambridge-analytica-used-data-from-facebook-and-politico-to-help-trump

Kshetri, N. and Voas, J. (2017). The Economics of “Fake News” IT Pro-IEEE. https://ieeexplore-ieee-org.ezproxylocal.library.nova.edu/stamp/stamp.jsp?tp=&arnumber=8123490

Kuang, X. (2018). Central State vs. Local levels of government: Understanding news media censorship in China. Fudan University and Springer Nature Singapore Pte Ltd. https://link.springer.com/article/10.1007/s41111-018-0091-5

Lima, C. (2019). Facebook wades deeper into censorship debate as it bans 'dangerous' Accounts. *Politico*, https://www.politico.com/story/2019/05/02/facebook-bans-far-right-alex-jones-1299247

Meyer, R. (2018). The grim conclusions of the largest-ever study of Fake News. *The Atlantic,* https://www.theatlantic.com/technology/archive/2018/03/largest-study-ever-fake-news-mit-twitter/555104/

Noyes, K. (2015). Mind your analytics, or you could soon be in ethical hot water. *IDG News* Service. https://www.computerworld.com/article/2993922/mind-your-analytics-or-you-could-soon-be-in-ethical-hot-water.html

Rathi, R. (2018) Effect of Cambridge Analytica's Facebook ads on the 2016 US Presidential election. https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d

Room, T. (2019). Senate Republicans renew their claims that Facebook, Google and Twitter censor conservatives. *The Washington Post*. https://www.washingtonpost.com/technology/2019/04/10/facebook-google-twitter-under-fire-senate-republicans-censoring-conservatives-online/?utm_term=.beef09528527

San Pedro, E. (2016). Cuba Internet access still severely restricted. *BBC* https://www.bbc.com/news/world-latin-america-35865283

Smith, A. (2017). Campaigns know more than ever about you — And it may be breaking our politics. *Business Insider*, https://www.businessinsider.com/campaign-big-data-use-politics-analytics-2017-6

Todd, C. and Dann, C. (2017). How Big Data broke American politics. *NBC News*. https://www.nbcnews.com/politics/elections/how-big-data-broke-american-politics-n732901

Umphress, E., Bingham, J., and Mitchell, M. (2010). Unethical behavior in the name of the company: The moderating effect of organizational identification and positive reciprocity beliefs on unethical pro-organizational behavior. *Journal of Applied Psychology*. Vol. 95, No. 4, 769 –780

Wagner, J. (2016). Clinton's data-driven campaign relied heavily on an algorithm named Ada. What didn't she see? *The Washington Post.* https://www.washingtonpost.com/news/post-politics/wp/2016/11/09/clintons-data-driven-campaign-relied-heavily-on-an-algorithm-named-ada-what-didnt-she-see/?utm_term=.b61a6889691b

Yakob, C., Surjandy, Ernawaty. (2017). Higher Education student behaviors in spreading Fake News on Social Media: A case of LINE Group. *International Conference on Information Management and Technology* (ICIMTech). https://ieeexplore.ieee.org/document/8273511

Yasaklanan, D. and Ağlar, S. (2018). Mapping the banned Social Media by country. https://graphcommons.com/graphs/52eadc97-6bc3-48a2-8526-741f358c9146

DECISION SCIENCES INSTITUTE

Factors Influencing the Success of P2P Microfinancing: A Sentiment Analysis Approach

Supavich (Fone) Pengnate
North Dakota State University
Email: fone.pengnate@ndsu.edu

Frederick J. Riggins
North Dakota State University
Email: fred.riggins@ndsu.edu

ABSTRACT

This study investigates how individual borrowers in the online peer-to-peer (P2P) platform can improve the chance of success for their loans. Specifically, this study examines the relationships between the length and emotions in loan descriptions and loan success. Emotions in the descriptions were extracted using one of the text mining approaches—sentiment analysis. The results reveal valuable new information in the P2P lending context such that, in the absence of market interest rates, borrowers can improve the chance of funding success by incorporating specific emotions in their loan descriptions.

KEYWORDS: P2P Lending, Sentiment Analysis, Emotions, Microfinancing

INTRODUCTION

Ledgerwood (2000) defines microfinance as “the provision of financial services to poor or low-income clients, including consumers and the self-employed.” A 2018 study showed that at the end of 2017 there were more than 900 microfinance institutions (MFIs) providing loans to more than 139 million borrowers (Convergences, 2018). Over the past decade, the provision of loans to microentrepreneurs in this sector has increased due to rapid adoption of the online channel in various parts of the world and the proliferation of peer-to-peer (P2P) applications. In P2P microfinance lending, the lending platform is used to provide microfinance services that match lenders with borrowers without the use of an official financial institution as an intermediary (Investopedia, 2018).

One of the main reasons for the growth of P2P lending is increasing adoption of the Internet, which has provided new opportunities for microentrepreneurs in developing countries to raise funds. While the Internet has provided a new dimension to the establishment of the relationship between P2P borrowers and lenders, questions regarding the motivation for individual lenders to fund loan projects has not been fully addressed. In general, P2P platforms do not pay interest to lenders; therefore, in the absence of the interest, there are several other uncertainties and risks associated with a decision of a lender to fund some loan projects and not others (Zohir & Matin, 2004). Specifically, without interest payments, P2P lending platforms rely on lenders’ trust in the borrowers to promptly repay the loans (van Bastelaer & Leathers, 2006). Consequently, there are increased concerns in P2P lending regarding trustworthiness of borrowers since it is difficult for a lender to ensure whether the borrower intends to repay the loan.

The major objective of this study is to explore characteristics of P2P loans, especially emotions in loan project descriptions, that make them more appealing to lenders. The results of this study contribute to the literature on P2P lending platforms by increasing our understanding of the distortions that can arise from information asymmetries and can provide practical guidance to platform providers and microentrepreneurs seeking to understand ways to improve P2P lending success.

LITERATURE REVIEW AND HYPOTHESES

Loan Descriptions

In the P2P lending context, previous studies have reported that among the various features of loans, loan description appears to be the top feature that influences the loan's success. According to Heller and Badding (2012), loan description was found to be the most important factor that determines the funding success measured by funding time. Therefore, we further investigate components of loan description in influencing the loan success.

Past studies have provided interesting evidence that, in addition to the content of the message, the length of the message can be rationally perceived as an indicator of the quality of the information since it is relevant to completeness of the message (Ferrari, Bouffard, & Rainville, 1998; Ransbotham & Kane, 2011). Such effects can be explained by the elaboration likelihood model (ELM) used by Petty and Cacioppo (1986) who examine the underlying mechanisms or influencing routes individuals rely on when making a decision or evaluation. ELM posits that the two routes (central and peripheral) of influence can cause individuals' attitude changes. These two routes differ in the amount of thoughtful information processing or "elaboration" demanded of the individuals. The central route requires higher cognitive effort in the evaluation, for example, an individual can critically think about the informational message by scrutinizing relative merits prior to forming an informed judgement. In the P2P lending context, an influence route may refer to the evaluation of a loan's features to determine the risk, such as loan purpose, amount, and the borrower's personal information.

On the other hand, the peripheral route requires lower cognitive effort to process information than the central route since it involves examining simple cues. It has been reported that when a high level of cognitive processing is demanded to evaluate a message such that it overloads an individual's working memory, he or she may use the peripheral route (length of text) as a primary route in forming the judgement (Petty & Cacioppo, 1986).

Previous studies have provided evidence of such peripheral route features of textual information on user's evaluation. For example, Korfiatis et al. (2012) reported that text's stylistic elements such as readability and length positively influenced helpfulness of online reviews. In a recent study in microfinance in the European countries, Dorfleitner et al. (2016) confirmed that soft information in loan descriptions especially description length can mitigate information asymmetry between lenders and borrowers and thus can be used as a predictive factor for loan default probability. Consequently, we utilize the insights of previous literature to further investigate how description length drives P2P funding success. Thus, we propose that lenders in the P2P lending may perceive longer loan descriptions as more complete and trustworthy because they provide more information of the loans, which subsequently influences funding success.

H1: The length of loan description is positively correlated with funding success.

Emotions in Loan Descriptions

In this study, we aim to investigate how affect-related factors (especially emotion) in loan descriptions influence funding success. Specific to textual information, previous studies have suggested that emotion can shape readers' attitudes and judgment of credibility of messages (e.g., Chaiken & Maheswaran, 1994). In addition, research in the marketing domain has provided evidence that product description is a crucial factor for sellers to convince and persuade customers of their legitimacy and trustworthiness. For example, MacInnis et al. (1991) found that product description can be considered as a stimulus that provides executional cues to borrowers' emotional responses and behaviors.

Consistent with previous studies, the persuasive message in P2P loan description can evoke emotional states of potential lenders, such as joy, liking, and fear, which can further influence the lender's evaluation of the borrower (Larrimore, Jiang, Larrimore, Markowitz, & Gorski, 2011). This effect can be explained by the concept of "affect-as-information" (Schwarz, 1986) which posits that emotions generated from a loan description are carried over to the lender's evaluation of the borrower's credibility (Hwang & Kim, 2007; Kim & Tadisina, 2010). In the online commerce literature, emotion has been widely studied and its important has been consistently shown as an effective component that drives overall impression of users which subsequently influence several key factors in online commerce success, such as trust (Hwang & Kim, 2007), satisfaction (Hsu, Chang, & Chen, 2012), and evaluations and responses (Éthier, Hadaya, Talbot, & Cadieux, 2006). In this study, we investigate a set of basic emotions such as anger, fear, joy, disgust, and sadness, which are thought to serve distinct functions in individuals' motivation and behavior (Ekman, 1992; Frijda, 1986). In the online P2P lending context, associations between an emotion and lenders' responses should be stronger if the emotion elicits physiological changes that are to a higher degree compatible with the lenders' lending interests and preferences, which consequently influence funding success. For example, a loan with sadness as the main emotion in the description may be more probable than joy to activate a lender's lending behavior. Therefore, we propose the following hypothesis.

H2: Emotions in loan description are associated with funding success.

METHODOLOGY

Data Collection

We conducted this investigation based on an open public access data source retrieved from the P2P lending website www.kiva.org. Kiva is a non-profit organization which provides microfinancing services to people from more than 83 countries (www.kiva.org/about). Kiva's lending model is based on a typical crowdfunding model which any individual can fund a particular loan by contributing to a loan individually or as a part of a lending group. Typical microloans are in increments of $25.

Data Preparation Procedure

Our original dataset was comprised of P2P loans that were collected between November 2011 and December 2016. However, most of the loans were fully funded by field partners and the funds were disbursed within 24 hours after the loans had been listed. Those field partners were local financial intermediaries that sought to general profits from P2P lending by collecting interests or fees. Therefore, those loans were excluded from the analysis since they did not represent the actual P2P lending interactions between individual borrowers and lenders as discussed in this study aiming at assessing the factors influencing funding success in the context of philanthropic giving. As a result, the dataset included 815 loans (240 from Kenya and 575 from the U.S.) that were funded solely by lenders and the funds were directly transferred from the lender's account to the borrower's account without going through an intermediary.

Sentiment Analysis

In this study, sentiment analysis was selected to perform emotion tagging. According to Pang and Lee (2008), sentiment analysis refers to the use of data mining techniques to identify and extract subjective information in source material. Sentiment analysis addresses the polarity of texts—positive, negative, or neutral—at different levels and for various aspects, including emotions, which are investigated in this study.

A proprietary sentiment analysis tool, IBM Watson Natural Language Understanding (NLU) was used to analyze emotions in Kiva loan descriptions. Watson NLU is available on the IBM cloud computing PaaS (Platform-as-a-Service). Watson NLU provides a set of services to build applications for natural language processing (NLP) and emotion recognition. For our study, the detectable emotions include the five basic emotions (anger, disgust, fear, joy, and sadness) suggested by Ekman (1992). A programming script was created in Python to pass the loan descriptions to Watson NLU, which subsequently extracted emotions and emotion score values associated with the descriptions. The emotion score values ranged from 0.0 to 1.0 and represented the prominence of the emotions conveyed in a loan description. A higher value represented more prominent emotion found in a loan description and if a score was above 0.5, the corresponding emotion was considered as the major emotion in the description. In addition to the emotion score values, Watson NLU also evaluated sentiment (positive, negative, or neutral) and provided sentiment score values to be used in our analysis to indicate the significance of sentiment in loan descriptions.

ANALYSIS AND RESULTS

Emotion in Microfinancing Loan Descriptions

The results from the emotion analysis reveal that joy and sadness appear to be the highest and second-highest emotions found in loan descriptions. We then determined the major emotions conveyed in loan descriptions based on two criteria—(1) an emotion that appears to have the highest frequency and (2) the emotion must have a documented emotion score above 0.5 (IBM Watson, 2018).

Dependent and Independent Variables

The dependent variable is loan's success determined by the length of time (in days) a loan takes to be fully funded. A set of independent variables as proposed in the hypotheses was examined. These variables include length, loan repayment term, sentiment score, five emotion scores, and a dummy variable assessing whether the major emotion found in a lender's lending description matches the major emotion of the loan description the lender funded.

Analysis and Hypothesis Testing

A series of multiple regression analyses was carried out to examine the influence of loan features including emotions on the dependent variable, loan's success. Overall, the results suggest that the length of loan descriptions, sentiment score, anger score, and disgust score appear to be statistically significant predictors of the funding at the 0.05 level.

Regarding the hypothesis testing, the results indicate that longer loan descriptions are significantly associated with funding success in the overall dataset (H1, $p < 0.05$). As for the emotion in the loan description (H2), the hypothesis is supported such that Anger and Disgust are found to negatively influence funding success ($p < 0.05$). The results are presented in Table 1.

Table 1: Regression Results for the Relationships between Loan's Success and the Independent Variables

Independent Variable	***β***	**Sig. ($p < 0.05$)**
Length	0.011	Yes
Repayment Term	0.055	No
Sentiment	7.276	Yes
Anger	-20.380	Yes
Disgust	-11.115	Yes
Fear	2.648	No
Joy	13.325	No
Sadness	11.214	No

DISCUSSION AND CONCLUSION

Overall, our results confirm the relationship between loan description and funding success reported in existing studies. For various features in P2P lending, borrowers have struggled to determine which strategies work best in improving the chance of finding success. This study investigates how P2P borrowers can improve the chance of funding success by manipulating their descriptions. The results provide some insights into the linguistic features of successful strategies based on trust-related factors that can be applied to the online persuasion context. Specifically, in accordance with the trust building strategies, the results indicate that increasing word count or incorporating appropriate emotions in loan descriptions can improve funding success.

Regarding research implications, our research indicates that while there are several strategies borrowers can implement to increase the likelihood of P2P loan success, borrowers can reduce uncertainty and establish trust by providing more information to lenders. This can be achieved by simply adding more details to the loan descriptions which can help lenders better understand how the funds are to be spent and repaid. This strategy not only increases the length of the descriptions but also is effective in providing more concrete information to lenders to understand what they can expect after funding the loans.

The findings of this study also suggest a number of avenues for future research. First, a controlled empirical test of how emotions influence funding success in the P2P lending context would shed further light on the actual value of incorporating emotions in loan descriptions. Second, it might be interesting to investigate whether emotional loan description is more preferred over a rational detailed informative loan description. Third, it is important to explore how interactions between borrowers and lenders occur in the online P2P lending environment. Therefore, as the P2P lending market expands, future research may consider investigating matching features of loan projects and characteristics of lenders who are interested in funding the loans.

Overall, the factors related to funding success in the microfinance area have been studied too little by scholars and applied too little by practitioners. This paper articulates the importance of trust-related loan description features on funding success. In the current research, we provide empirical evidence that loan project description is an important factor in online P2P lending by adding to the literature emphasizing emotions extracted from the descriptions. The results provide practical solutions that can be effectively implemented and a theoretical foundation for future research on the success of the growing field of P2P lending.

REFERENCES

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. Journal of Personality and Social Psychology, 66(3), 460-473. doi:10.1037/0022-3514.66.3.460

Convergences. (2018). Microfinance barometer 2018. Retrieved from http://www.convergences.org/wp-content/uploads/2018/09/BMF_2018_EN_VFINALE.pdf

Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I., & Kammler, J. (2016). Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms. Journal of Banking & Finance, 64, 169-187. doi:https://doi.org/10.1016/j.jbankfin.2015.11.009

Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, 6(3-4), 169-200. doi:10.1080/02699939208411068

Éthier, J., Hadaya, P., Talbot, J., & Cadieux, J. (2006). B2C website quality and emotions during online shopping episodes: An empirical study. Information & Management, 43, 627-639.

Ferrari, M., Bouffard, T., & Rainville, L. (1998). What makes a good writer? Differences in good and poor writers' self-regulation of writing. Instructional Science, 26(6), 473-488. doi:10.1023/a:1003202412203

Frijda, N. H. (1986). The emotions. Cambridge: University Press.

Heller, L. R., & Badding, K. D. (2012). For compassion or money? The factors influencing the funding of micro loans. The Journal of Socio-Economics, 41(6), 831-835. doi:https://doi.org/10.1016/j.socec.2012.08.005

Hsu, C.-L., Chang, K.-C., & Chen, M.-C. (2012). The impact of website quality on customer satisfaction and purchase intention: perceived playfulness and perceived flow as mediators. Information Systems and e-Business Management, 10(4), 549-570. doi:10.1007/s10257-011-0181-5

Hwang, Y., & Kim, D. J. (2007). Customer self-service systems: The effects of perceived web quality with service contents on enjoyment, anxiety, and e-trust. Decision Support Systems, 43(3), 746-760. doi:10.1016/j.dss.2006.12.008

IBM Watson. (2018). Alchemylanguage API reference. Retrieved from https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/#emotion_analysis

Investopedia. (2018). What is 'peer-to-peer lending (P2P)'. Retrieved from https://www.investopedia.com/terms/p/peer-to-peer-lending.asp

Kim, E., & Tadisina, S. (2010). A model of customers' initial trust in unknown online retailers: An empirical study. International Journal of Business Information Systems, 6(4), 419-443. doi:10.1504/ijbis.2010.03574

Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. Electronic Commerce Research and Applications, 11(3), 205-217. doi:https://doi.org/10.1016/j.elerap.2011.10.003

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. Journal of Applied Communication Research, 39(1), 19-37. doi:10.1080/00909882.2010.536844

Ledgerwood, J. (2000). Microfinance handbook: An institutional and financial perspective. Washington, DC: The World Bank.

MacInnis, D. J., Moorman, C., & Jaworski, B. J. (1991). Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads. Journal of Marketing, 55(4), 32-53.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Informaiton Retrieval, 2(1-2), 1-135.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. Advances in Experimental Social Psychology, Volume 19, 123-205.

Ransbotham, S., & Kane, G. C. (2011). Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in wikipedia. MIS Quarterly, 35(3), 613-627. doi:10.2307/23042799

Schwarz, N. (1986). Feelings as information: Informational and motivational functions of affective states. In T. Higgins & R. M. Sorrentino (Eds.), Handbook of Motivation and Cognition (Vol. 2, pp. 527-561). New York: The Guilford Press.

van Bastelaer, T., & Leathers, H. (2006). Trust in lending: Social capital and joint liability seed loans in southern zambia. World Development, 34(10), 1788-1807. doi:https://doi.org/10.1016/j.worlddev.2006.02.007

Zohir, S., & Matin, I. (2004). Wider impacts of microfinance institutions: issues and concepts. Journal of International Development, 16(3), 301-330. doi:doi:10.1002/jid.1080.

DECISION SCIENCES INSTITUTE

Abstractive Summarization Through Sentiment Analysis Of User Product Reviews - An RNN Approach

Anjana Arunkumar
Arizona State University
Email: aarunku5@asu.edu

M.K. Sandhya
Meenakshi Sundararajan Engineering College, Anna University
Email: mksans@gmail.com

ABSTRACT

This research study focuses on performing granular sentiment detection, using a recurrent neural network approach, over a dataset of online reviews for Slack, an American cloud-based set of proprietary team collaboration tools and services. A client interface was built, attaching sentiment scores to opinion units, keywords and performance indicators, and generating interactive visualizations of the same. An encoder-decoder model with Attention was also implemented, to generate an abstractive customer feedback summary for the client, as a premise for future product development.

KEYWORDS: Information Science, Data Mining, RNN, Sentiment Analysis, and Summarization

INTRODUCTION

Online platforms, such as Capterra, are increasingly being used as a source of information and opinions about a plethora of issues, concepts and products. Such a user review corpus is a rich data source, as it provides information to both the company's prospective customers and marketing teams about the user sentiment with respect to the product or service considered. People's choices are also dependent to some degree on how others view and evaluate the world.

E-commerce platforms provide companies with a structured data set for analysis. In the real world, tracking data is not as easily obtainable, with the primary sources of intelligence being user feedback and the financial data of firms. The purpose of this research study is to perform effective sentiment detection of user feedback, using an encoder-decoder RNN(recurrent neural network) approach with Attention to process a dataset of online service and product reviews for the Slack software. A user interface was then built to implement a report generation engine for the clients' to present a priority view of areas to be addressed to increase customer satisfaction and maximize product performance and recommendation likelihood.

The first step towards training a classifier with machine learning is feature extraction, which is an approach that transforms each text into a numerical representation in the form of a vector. This is done using one shot encoding in RNN. Then, the RNN algorithm is fed with training data that consists of pairs of feature sets to produce a classification model. To learn new tasks, new tags can be created.

Key phrases can provide highly condensed and valuable information that allows users to quickly acquire the main ideas. The task of automatically extracting them is achieved using an RNN encoder-decoder model to combine key words and context information. It has two hidden layers to discriminate keywords and classify key phrases, and these two sub-objectives are combined into a final objective function.

Automatic summarization is the process of shortening a text document with software so as to create a summary with the major points of the original document. Technologies that can make a coherent summary consider variables such as length, writing style and syntax. There are two general approaches to automatic summarization: extraction and abstraction. Extractive methods work by selecting a subset of existing words, phrases, or sentences in the original text to form the summary. In contrast, abstractive methods build an internal semantic representation and then use natural language generation techniques to create a summary that is closer to what a human might express. Such a summary might include verbal innovations. Attention is used in the RNN model to generate an abstracted summary.

LITERATURE REVIEW

Ghosh & Sanyal (2018) investigated the inability of the widely used feature selection methods (IG, Chi Square, Gini Index) individually as well as their combined approach on four machine learning classification algorithms. The proposed methods were evaluated on three standard datasets viz. IMDb movie review, electronics product review dataset, and kitchen product review dataset. Initially, the feature subsets were selected from three different feature selection methods. Thereafter, the statistical methods of UNION, INTERSECTION and revised UNION method were applied to merge these different feature subsets to obtain all top ranked (including common selected) features. Finally, the classifiers of SMO(sequential minimal optimization), MNB(multinomial naïve bayes), RF(random forest), and LR (logistic regression) were trained with this feature vector for classification of the review data set. The performance of the algorithm was measured by evaluation methods such as precision, recall, F-measure and ROC(receiver operating characteristic) curve. Experimental results showed that the combined method achieved best accuracy of 92.31 with the classifier SMO.

Liu (2010) described a threefold approach to sentiment analysis. The first one, classical approaches, used techniques that predated the statistical revolution, i.e. before natural language processing researchers embraced techniques that speech engineers had already been using successfully. The second, empirical and statistical approaches, covered state-of-the-art data-driven models. The third one, application oriented approaches, catered to the needs of NLP(natural language processing) practitioners and language-engineering professionals in academia as well as in industry. No efficient method to tackle opinion spam was identified. The usage of SVM(support vector machine) was proposed, though the results produced were not directly human understandable. A pure machine learning reliance hence served to limit the cognitive ability of a system.

Santos & Gatti (2014) implemented character-to-sentence mapping in a convolutional neural network. Two convolutional layers were used to extract relevant word/sentence features. The datasets used were SSTb Movie Review, and Tweets (STS). Word and character level embeddings and sentence level scoring interfaces were implemented using CNN(convolutional neural network) to extract wider feature selection. Feed forward neural networks were found to be as effective as RNTN(recursive neural tensor network) usage. A binary accuracy of 85.7% and a fine-grained accuracy of 48.3% were achieved. Their study did not consider the impact of an unsupervised pretraining step using texts from the domain.

Severyn & Moschitti (2015) described a deep learning system for the sentiment analysis of tweets. They developed a model for initializing the parameter weights of the convolutional neural network, which is crucial for training train in an accurate model while avoiding the need to inject any additional features. An unsupervised neural language model, trained initial word embeddings, and these were further tuned by a deep learning model on a distant supervised corpus. The pre-trained parameters of the network were then used to initialize the model. The deep learning model was trained on the supervised training data made available by the official system evaluation campaign on Twitter Sentiment Analysis organized by Semeval-2015. A comparison between the results of this approach and the systems participating in the challenge on the official test sets, suggested that this model could be ranked in the first two positions in both the phrase-level subtask A (among 11 teams) and on the message-level subtask B (among 40 teams).

Habernat, I., Ptacek, T. & Steinberger, J. (2014) examined N gram and POS(part of speech) related features, with emoticons selected using MI(mutual information), Chi, RS(relief based selection). Classifiers used include SVM, on a Czech Social media data set, the IMDb Movie set, and the Amazon Product set. Lemmatizing was done with the PDT toolkit using Java framework with a 10-fold cross validation. Bigram feature had the highest performance, outperforming the baseline in three class classifications with an F-Measure of 0.69. In their study, many negative posts had happy emoticons (misleading features) which couldn't be handled. Hidden irony was also not identified in negative posts.

Pang, B., Lee, L. & Vaithyanathan, S. (2002) used NB, SVM, Maximum Entropy in a bag of features framework on the IMDb Movie Review dataset. The bag of unigram features achieved roughly 90% accuracy. Feature presence was found to be more effective than feature frequency (SVM). Overall accuracy was 77.4-82.9%. The model, however, was not able to achieve accuracy on topic based categorization. It was also not able to classify mixed reviews. Thwarted expectations and rhetorical devices were also not identified.

Extraction

Zhang, Q., Wang, Y., Gong, Y. & Huang, X. (2016), in their study, considered key phrases as capable of providing highly condensed and valuable information that allowed users to quickly acquire the main ideas. They examined the issue of automatically extracting key phrases from tweets. Due to the length limitations of Twitter-like sites, the performances of existing methods usually drop sharply. To overcome this, they sued a novel deep RNN model to combine key words and context information to address this problem. Their experimental results indicated that the RNN method performed significantly better than previous methods.

In their study, Zhang et. al. (2016) extracted explicit product features and opinions by phrase level SA(sentiment analysis) to generate recommendations and dis-recommendations according to specific product features to the user's interests. Intuitional feature level explanations were also generated. However there was no comparison group assigned, with other explanations for the dis-recommendation engine. Cart items were recorded as user purchase behavior approximation, as security prevented the tracking of user purchases, which introduced a significant margin of error. Their study's focus was only on persuasiveness, not on other utilities like transparency, user trust, effectiveness, suitability, etc. which might also factor into SA.

Churn

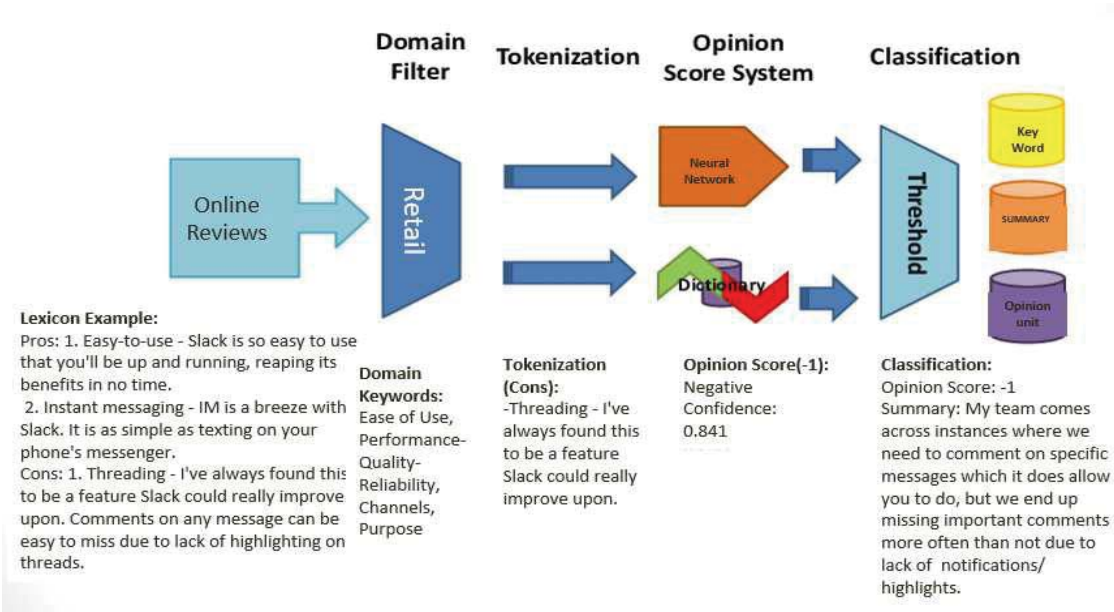
McKnight (2011) presented a reference implementation of a customer churn analysis project that was built by using the Azure Machine Learning Studio. Businesses in the consumer market and in all enterprise sectors have to deal with churn. Sometimes churn is excessive and influences policy decisions.

His study suggests that the common factor is that businesses need to minimize these special customer retention efforts. A natural methodology to use would be to score every customer with the probability of churn and address the top N ones. The top customers might be the most profitable ones. For example, in more complex scenarios, a profit function is employed during the selection of candidates for a special dispensation. However, these considerations are only a part of the complete strategy for dealing with churn. Businesses also must consider risk (and associated risk tolerance), the level and cost of the intervention, and plausible customer segmentation.

His study's findings suggest that the implementation in Azure Machine Learning Studio lags behind SAS in accuracy by about 10-15% (Area Under Curve or AUC). However, the most important metric in churn is the misclassification rate: that is, of the top N churners as predicted by the classifier, which of them did not churn, and yet received special treatment. This was found to be identified with highest accuracy by the model implemented.

MODEL

Figure 1: System Architecture



The system architecture given in Figure 1 explains the overall flow of the model in this research study. The customer reviews are scraped from the feedback site. These are then fed into models for data cleaning and splitting as per categories - pros, cons and sub-ratings. Then, tokenization is done and a keyword/opinion unit lexicon is generated from the training set, along with sentiment weighting at a word level, using an RNN. The test data is then passed through the network, and an encoder-decoder RNN model with Attention is used to abstract summaries and generate an average user feedback report to the client.

The client is provided with a series of tree-maps, a facility to construct pie review, word and letter level sentiment charts, along with an overall review summary. The client's analyst may generate and view the above, as well as additionally view all the extracted raw review sets and reweight the neural network according to the service provided by the client to its customers, in order to dynamically generate summary reports of reviews to the client in real time.

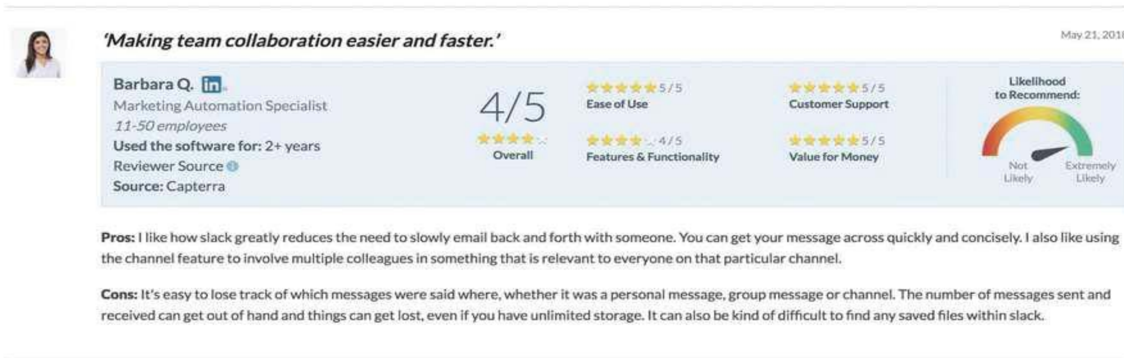
MODULES

The various modules that form part of this research study's methodology are: Data collection, Data cleaning, Lexicon construction and weighting, Scoring, Sentiment Threshold Computation, Summary Generation and Client Interface.

Data Collection

The first step was to collect all of the reviews of Slack that people have left on Capterra. On Capterra, reviews were parceled into chunks including the reviewer's overall rating of the product, the text of their review, and sub-ratings (the ratings of certain aspects of the product, e.g. Customer Support, Ease of Use, and Value for Money). The reviews were then saved in .csv files for future analysis.

Figure 2 : Slack Review on Capterra



Data Cleaning

The following were implemented in the data cleaning module using modified Rattle package functions:

- Getting Rid of Extra Spaces
- Selecting and Treating All Blank Cells(NA)
- Converting Numbers Stored as Text into Numbers
- Removing/Merging Duplicates
- Highlighting Errors
- Changing Text to Proper Case
- Deleting all Formatting

Lexicon Construction

Aspect classifier and extractor lexicon was built to help detect the content of the texts in the reviews (e.g. pricing, UX, customer support, performance, integrations, etc.). A set of tags was defined, that accounted for most of the topics or aspects included in the Slack reviews. By tagging the text data, the RNN algorithm were written so that for a particular input (reviews), a particular output (tags) is expected. The more data were tagged, the better was the model's working. Since most reviews include more than one sentiment, splitting longer texts into smaller bits helped detect different sentiments which were then later aggregated. The reviews were first submitted to the aspect classifier and extractor module, and then the opinion units and keywords obtained were fed as input to the next RNN layer.

Scoring

The text was transformed into a numerical representation, i.e. a vector. Each component of the vector represents the frequency of a word or expression in the predefined dictionary. Some samples were tagged as positive, negative, or neutral. These were used as a basis to classify the aggregate of lexicon scores, considering the context of their usage by first world level and then sentence level weighting, (feature extraction and weighting) to compute a review sentiment score.

Sentiment Threshold Computation

The algorithm used to implement the sentiment threshold computation was an RNN Encoder Decoder Model, with Attention considered, over the set of polarized keywords and opinion units generated in the scoring module. Scores were generated on a scale from -1 to 1, with Laplace smoothing used to weigh review scores as +1 (Positive), -1 (Negative) and 0 (Neutral).

Summary Generation

There were two main types of summarizing text in NLP:

- **Extraction-based summarization:** Involved pulling key phrases from the source document and combining them to make a summary. The extraction was made according to the defined metric without making any changes to the texts.
- **Abstraction-based summarization:** Entailed paraphrasing and shortening parts of the source document. When abstraction was applied for text summarization in deep learning problems, it overcame the grammar inconsistencies of the extractive method.

The merited key phrases were extracted across the review corpus, using the aspect classifier and extractor. A binary classifier was used to make the text summarization, by identifying and organizing pros and cons. Some of the features used include:

- Length of the key phrase
- Frequency of the key phrase
- The most recurring word in the key phrase Number of characters in the key phrase
- Lexicon weighting of the key phrase

Client Interface

The client was provided with a series of tree-maps, color coded according to review sentiments which can be viewed selectively using keywords and opinion unit search. A facility to construct pie charts to discover keywords pertaining to sentiment categories was also provided, along with category wise scoring of reviews. Word and letter level charts were displayed, along with an extracted summary generation for one per ten opinion units. An overall review summary was generated using an abstraction of the previously extracted summaries.

Hardware Description

Developing Kit			
Software	**Processor**	**Ram**	**Disk Space**
R Studio -Rattle -Monkey Learn Api -Rfutilities -Testthat	Intel Computer With A 2.6 Ghz Processor Or Higher	64mb	Minimum 20 Gb
Google Chrome -Selector Gadget	Intel Pentium 4 Processor Or Higher	128mb	Minimum 100mb
Database			
MS Office -Excel	Intel Pentium Processor At 2ghz Or Faster	2gb Ram Recommended	Minimum 4gb
Interface			
Tableau	Intel Pentium 4 Processor Or Higher	Minimum 512gb Physical Memory; 2gb Recommended	Minimum 1.5gb

Table 1: Hardware and Software Specifications

Scoring

The algorithm used to implement scoring was an RNN Encoder Decoder Model, with Attention taken into account, over the set of polarized keywords and opinion units generated in the scoring module. Scores were generated on a scale from -1 to 1, with Laplace smoothing used to weigh review scores as +1 (Positive), -1 (Negative) and 0 (Neutral).

Attention

A problem with the encoder-decoder architecture is that performance is poor on long input or output sequences. This is due to the fixed-sized internal representation used by the encoder. Attention is an extension to the architecture that addresses this limitation. It works by first providing a richer context from the encoder to the decoder and a learning mechanism where the decoder can learn where to pay Attention in the richer encoding when predicting each time step in the output sequence.

To model sentiment behavior, we used Feature Functions, that has multiple input values, such as:

- a sentence s
- the position i of a word in the sentence
- the label li of the current word
- the label li−1 of the previous word

Next, we assigned each feature function fj a weight λj . Given a sentence s, score a labelling l of s by adding up the weighted features over all words in the sentence:

$$score(l|s) = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j f_j(s, i, l_i, l_{i-1})$$

We transformed these scores into probabilities p(l|s) between 0 and 1 by exponentiating and normalizing:

$$p(l|s) = \frac{exp[score(l|s)]}{\sum_{l'} exp[score(l'|s)]} = \frac{exp[\sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j f_j(s,i,l_i,l_{i-1})]}{\sum_{l'} exp[\sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j f_j(s,i,l'_i,l'_{i-1})]}$$

The merited key phrases were extracted across the review corpus, using the aspect classifier and extractor. A binary classifier was used to make the text summarization, by identifying and organizing pros and cons. Some of the features used include length of the key phrase, frequency of the key phrase, the most recurring word in the key phrase, number of characters in the key phrase, and lexicon weighting of the key phrase.

To obtain the above accurately with respect to user ratings, we obtained a summary of the mean sentiment (based off numerical representation of sentiment as -1, 0,1) for opinion units that were classified into each aspect. We split these mean sentiment ratings into three equal parts and assigned those parts valences that describe the mean sentiment for that aspect. We found the tertiles of the mean sentiments by dividing them in three groups as they relate to each other. Then, we used these tertiles as the bounds for assigning valences.

We colored the bars of the bar-plots generated with those valences as arranged by different variables while retaining the measure of sentiment per category. We weighted the aspect-sentiment by the number of times it occurred in an opinion unit. This gave us a better idea of the sentiment in the context of how often it's mentioned. We considered this important because if an aspect has very low sentiment, but it's almost never mentioned, it may be less critical to focus on improving than an only mildly badly rated category with a lot of mentions. We used this to rank the order of abstraction when generating the summary.

The client was provided with a series of tree-maps, color coded according to review sentiments, which can be viewed selectively using keywords and opinion unit search. A facility to construct pie charts to discover keywords pertaining to sentiment categories was also provided, along with category wise scoring of reviews. Word and letter level charts were displayed, along with an extracted summary generation for one per ten opinion units. An overall review summary was generated using an abstraction of the previously extracted summaries.

Figure 3 : Sentiment Weighted Extraction

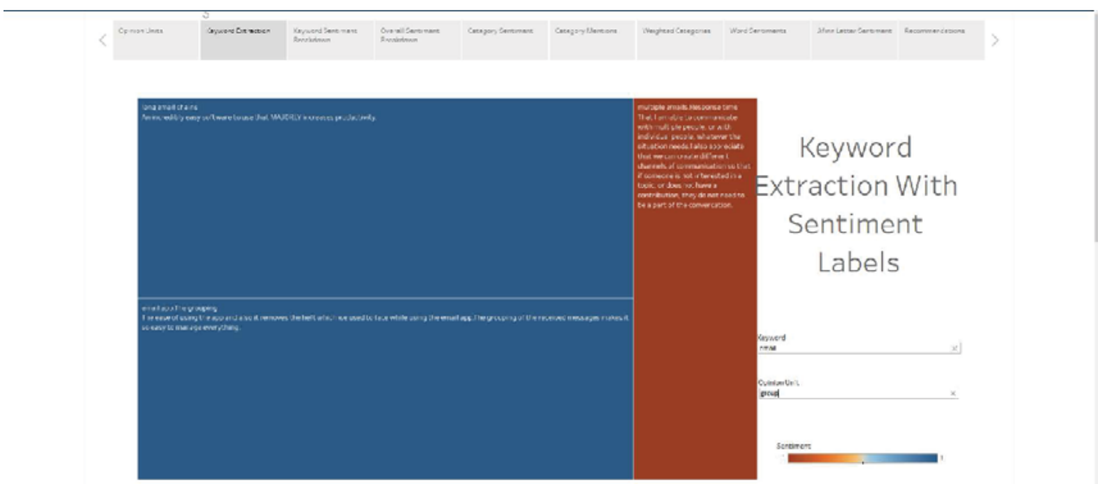


Figure 4 : Category Weighting

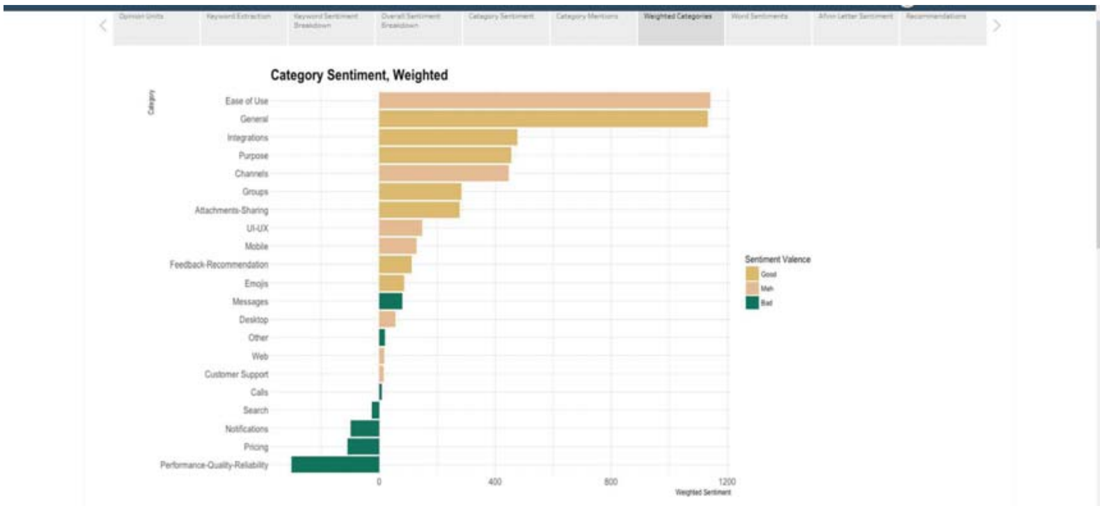
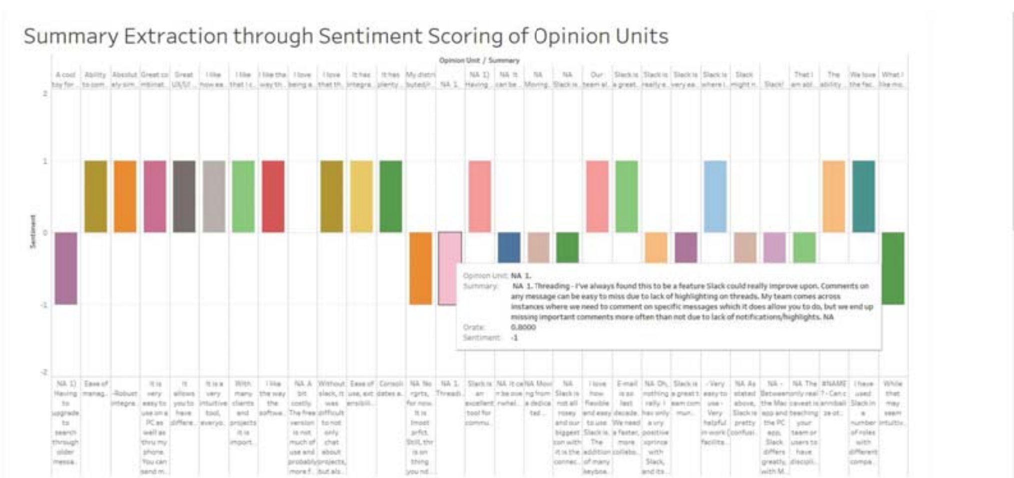


Figure 5 : Summary Generation



RESULTS

Performance Metrics Used

$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

$$\text{recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}$$

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

$$\text{Balanced Accuracy} = \frac{TPR + TNR}{2}$$

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

Test Cases

Figure 6 : Aspect Classifier- Accuracy and F1 score

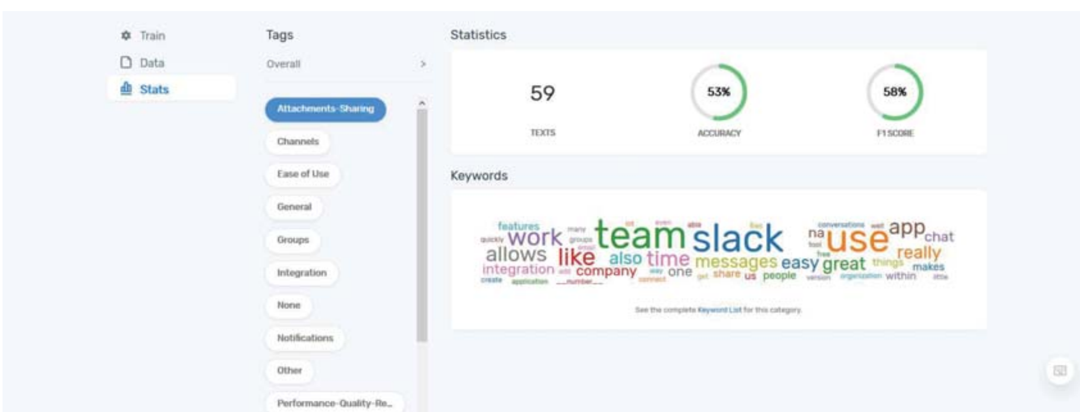


Figure 7 : Aspect Classifier- Precision and Recall

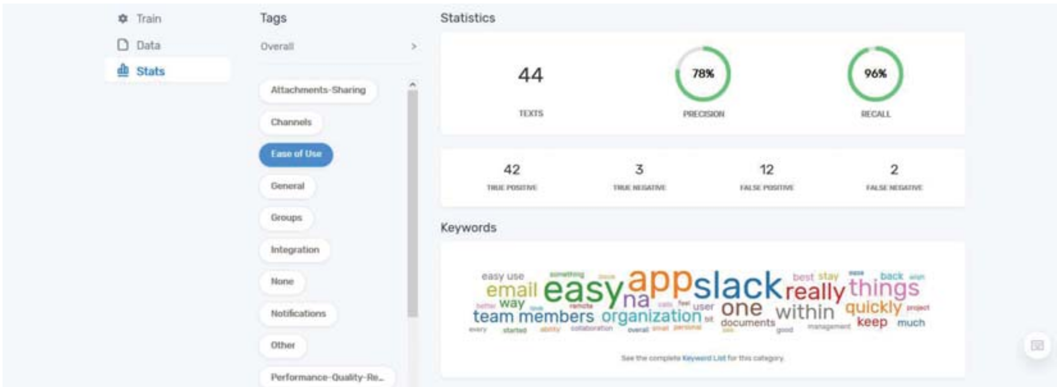


Figure 8 : Aspect Classifier- Aspect Wise Confidence

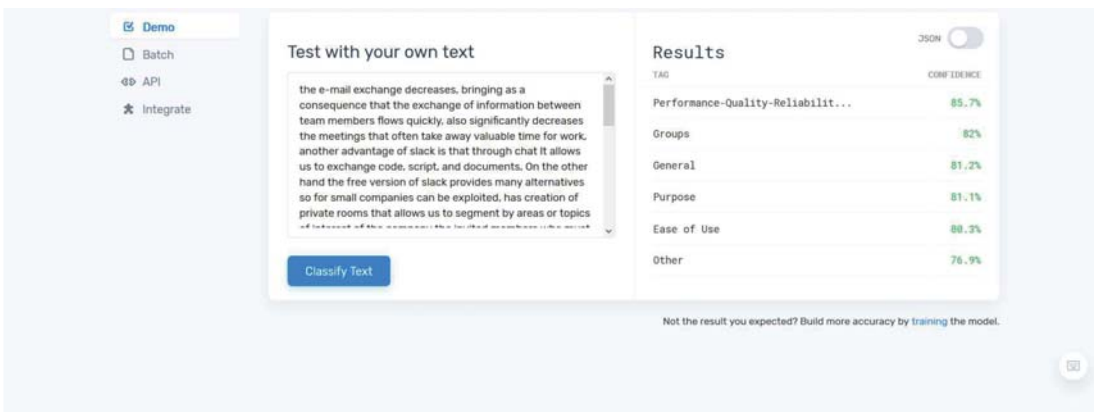


Figure 9 : Extractor- Opinion Unit Performance

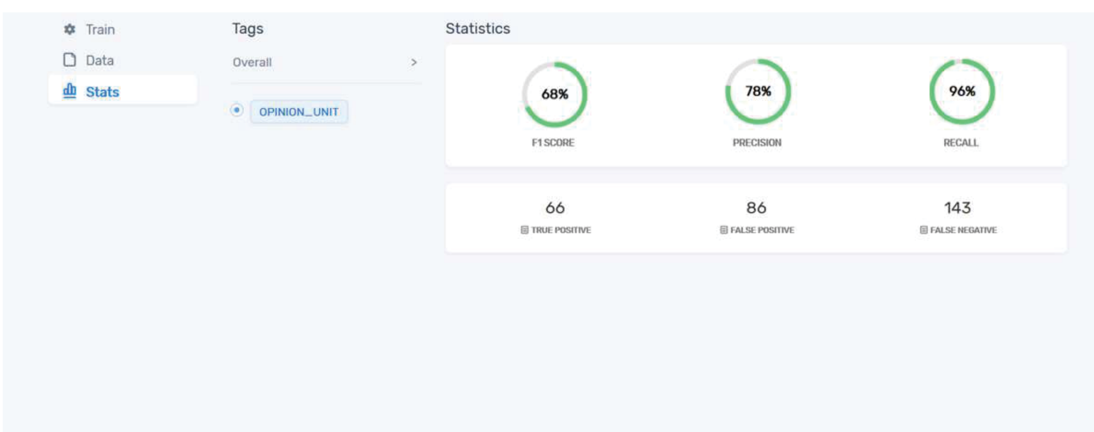


Figure 10 : Extractor- Keyword Performance

Figure 11 : Sentiment Analysis Performance

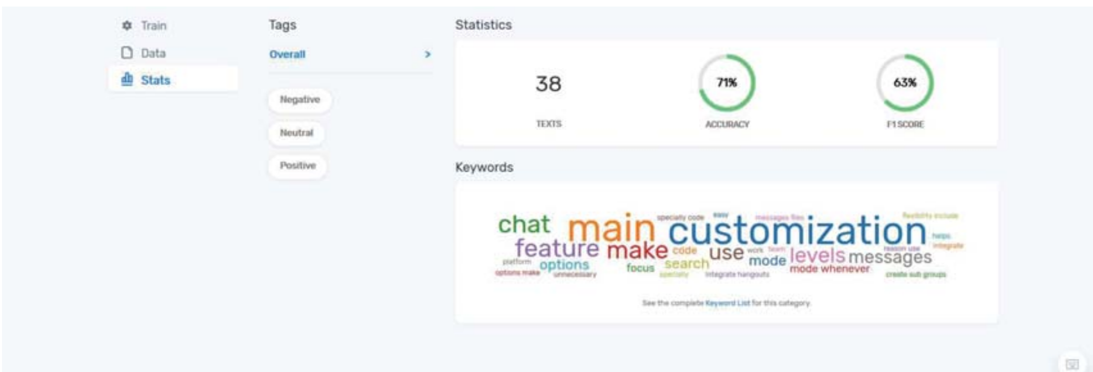


Figure 12 : Sentiment Analysis

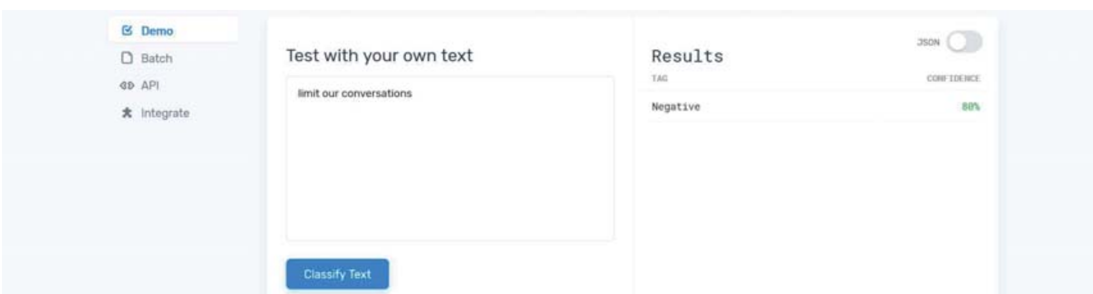


Figure 13 : Summary Abstraction Performance

CONCLUSION

This research study focused on performing more granular sentiment detection, using a recurrent neural network approach, over a dataset of online reviews for Slack, an American cloud-based set of proprietary team collaboration tools and services. A client interface was built, attaching sentiment scores to opinion units, keywords and performance indicators, and generating interactive visualizations of the same. An encoder-decoder model with Attention was also implemented, to generate an abstractive customer feedback summary for the client, as a premise for future product development.

Various test cases on both, real-life and synthetic data, were conducted to and the findings suggest that our approach outperformed traditional methods in sentiment analysis on data and delivered more or less the same performance on sparse data. Hence, our findings suggest that the system we proposed in our research study helps in the efficient handling of report generation based on sentiment analysis.

Future Enhancement

The scope of this research study can be further extended by:

- Developing a generalized keyword and scoring set, to apply across different products/services.
- Implementing sarcasm detection and increased handling of emojis in reviews.
- Integrating an automatic review mining and analysis module instead of periodic manual addition to and processing of the data set.

REFERENCES

Ghosh, M. & Sanyal, G. (2018). *An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning.* In: Journal of Big Data: Springer.

Habernat, I., Ptacek, T. & Steinberger, J. (2014). *Supervised Sentiment Analysis in Czech Social Media using supervised machine learning.* In: Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis.

Liu, B.(2010). *Sentiment analysis and subjectivity.* In: Indurkhya, N. & Damerau, F.J., editors. Invited chapter for the handbook of natural language processing. 2nd ed. England: Taylor & Francis.

McKnight, W. (2011). *Predictive Analytics: Beyond the Predictions.* Information Management, p.18-20.

Pang, B., Lee, L. & Vaithyanathan, S. (2002). *Thumbs up? Sentiment classification using machine learning techniques.* In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia, pp. 79-86.

Santos, C. & Gatti M. (2014). *Deep convolutional neural networks for sentiment analysis of short texts.* In: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, 69-78.

Severyn, A. & Moschitti, A. (2015). *Twitter sentiment analysis with deep convolutional neural networks.* In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, 959-962.

Sutskever, I., Vinyals, O. & Le, Q. (2014). *Sequence to Sequence Learning with Neural Networks.* arXiv:1409.3215 [cs.CL].

Zhang, Q., Wang, Y., Gong, Y. & Huang, X. (2016). *Keyphrase Extraction Using Deep Recurrent Neural Networks on Twitter.* Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 836–845.

DECISION SCIENCES INSTITUTE

Integrating Cloud Computing into the Curriculum: Thai Students' Perspective

Chuleeporn Changchit
Texas A&M University-Corpus Christi
Email: chuleeporn.changchit@tamucc.edu

Robert Cutshall
Texas A&M University-Corpus Christi
Email: robert.cutshall@tamucc.edu

ABSTRACT

This study investigates how cloud computing is perceived by Thai college students and which factors can influence their willingness to accept cloud computing as part of their course curriculum. The study findings should help increasing the chance for the success of the integration of cloud computing into the program.

KEYWORDS: Cloud Computing, Technology Acceptance Model (TAM), Perceived Usefulness, Perceived Ease of Use, Perceived Security, Perceived Speed, Perceived Cost, Thailand

INTRODUCTION

Cloud computing has attracted considerable attention because their services offer numerous benefits (Ya-Ching, 2019). The popularity of the usage is derived by its ability to provide faster on-demand infrastructure, self-service, and independent ability to contribute to and access resources. It is increasingly growing in convenience as a ubiquitous network that requires little interaction between the cloud service provider and the user (Changchit & Chuchuen, 2018). This flexibility and independence can help to save the consumer time, and stress by making the process simple and allowing the account design to be based entirely on individual needs.

In the past few years the concept of "cloud computing" has emerged as a viable and promising solution to the challenges associated with shrinking IT budgets and escalating IT needs. Cloud computing has become increasingly popular among users and businesses around the world. Education is also a driving force for the continuous improvement of cloud computing. Students can gain a lot from this technology as it serves as a convenient mobile storage space (Singh & Veralakshmi, 2012). Cloud computing can bring an increased number of benefits to an educational setting (Behrend et al., 2011). It is not only the cost effectiveness, but also the thirst for technology that college students have today, which allows learning and adopting these new technologies easier for them.

Cloud computing is now considered one of the commonly deployed services due to its relative advantages for organizations, firms, and enterprises (Almubarak, 2017). This technology provides a number of benefits to an educational setting (Behrend et al., 2011). The technology allows students to have their own mobile storage space (Singh & Veralakshmi, 2012). Some of the benefits that students can derive from cloud computing, in an educational setting, include using it for completing assignments, online classes, group projects, creating and editing papers and presentations; as well as for work or entertainment (Changchit, 2015).

With the many benefits generated by the cloud computing model, it is interesting to investigate students' perceptions of this technology. This is key since the success of integrating a course into a curriculum depends a lot on students' attitudes toward such a topic. The purpose of this study is to examine the factors that encourage or discourage students to accept cloud computing as part of their course curriculum.

LITERATURE REVIEW

The high demand for more advanced and efficient technology has contributed in the creation of new advancements in technology such as cloud computing. The popularization of cloud computing by companies like Amazon®, Google®, and Apple® ensure that the usage of the cloud as a storage medium for music, movies, and other media content files, will be ubiquitous in the next 10 years. Cloud computing is a promising prospect for educational institutions, especially during times of budget constraints. Research in cloud computing adoption in educational settings has focused part of its efforts to understand the drivers and constraints that students and schools perceive in the adoption of this computing model. With today's technology, students' learning is no longer confined to the classroom. The educational environment could be improved to allow students to access learning resources anywhere and anytime (Wu, 2013).

One study examined the factors leading to adopting cloud computing as a virtual computing lab for a class (Behrend et al., 2011). The authors in this study found that students' ease of use perception would positively affect intentions for future use, but not for actual use. Students who complete their work faster and in a more practical manner were more likely to recognize cloud computing as an effective service, and use it more if there is no "effort to learn". This study also found that students with anxiety about new technologies had a negative effect on perceived usefulness. Another study also suggested that in order to deal with technology anxiety, it is important for universities to plan hands-on training to help students become more familiar with these new technologies (Blue & Tirotta, 2011).

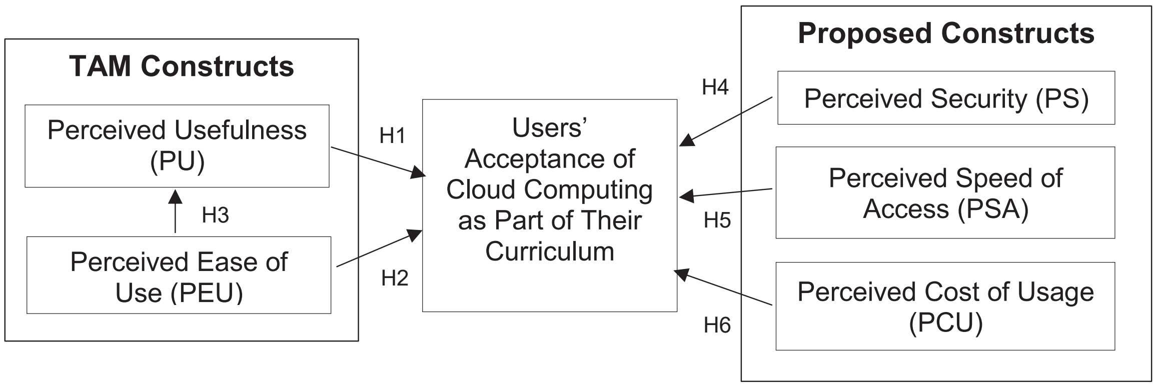
In Thailand, since the year 2011, the use of cloud computing services has negatively impacted the hard disk manufacturing industry as it is considered a hard disk substitution. Cloud computing services in Thailand are in the early adopter stage, with the market projected to grow exponentially in the next 4-5 years (Leesa-Nguansuk, 2018).

With many benefits promised by the use of cloud computing, it is not surprising that the demand for graduates with exposure to Cloud Computing is on the rise (Chen et al., 2012). However, despite the potential for positive change, there are still other factors to consider before cloud computing can be accepted entirely by academic groups. Before integrating such topics into the curriculum, it is crucial to understand which factors can encourage or discourage students from accepting it as a part of their course curriculum. The findings in this study should help programs to focus on the factors that encourage students to like the topic as well as finding ways to minimize the effect of discouraging factors, thus increasing the chance for the success of the integration of cloud computing into the program.

THEORETICAL DEVELOPMENT/MODEL

The Technology Acceptance Model (TAM) developed by Davis in 1986 and was proposed to address the question why users accept or reject information technology. The model was derived from the Theory of Reasoned Action, developed by Fishbein and Ajzen (1975). TAM was a big hit for modeling user acceptance of information systems (Bagozzi, 2007; Davis et al., 1989). This study has modified TAM to include three additional factors: (1) Perceived Security (PS), (2) Perceived Speed of Access (PSA), and (3) Perceived Cost of Usage (PCU). The proposed research model is shown in Figure 1 below.

Figure 1 Research Model



Perceived Usefulness of Cloud Computing (PU)

Perceived usefulness is defined as the prospective users' subjective probability that using a specific application system will increase his or her job performance within an organizational context (Davis et al., 1989). This factor has a significant effect on usage intention (Davis et al., 1989; Venkatesh & Davis, 2000). A prior study also reported that this construct positively influence subjects' intention to adopt cloud computing (Changchit & Chuchuen, 2018). Based on the foregoing, we propose the following hypothesis:

H1: Perceived usefulness (PU) positively affects subjects' acceptance of cloud computing as part of their curriculum.

Perceived Ease of Use of Cloud Computing (PEU)

Perceived ease of use is defined as the degree to which the prospective user expects the target system to be free of effort (Davis et al., 1989). This factor plays a crucial role in understanding individual responses to information technology (Chau & Hu, 2001). Research over the past decade provides evidence of the significant effect perceived ease of use has had on usage intention (Venkatesh & Davis, 2000). We, thus hypothesize that:

H2: Perceived ease of use (PEU) positively affects subjects' acceptance to have cloud computing as part of their curriculum.

In addition, perceived ease of use has been shown to have an impact on perceived usefulness. Davis (1986, p. 26) states "a system which is easier to use will result in increased job performance (i.e., greater usefulness) for the user." We, therefore, posit that:

H3: Perceived ease of use (PEU) positively affects perceived usefulness (PU).

Perceived Security of Cloud Computing (PS)

Security awareness is an important issue for all individuals who are dealing with sensitive data in everyday life (Changchit, 2008). Flavián and Guinalíu, (2006) presented their view of perceived security as a subjective probability with which consumers believe that their personal information (private and monetary) will not be viewed, stored, and manipulated during transit and storage by inappropriate parties in a manner that is inconsistent with their expectations. There is a high likelihood that subjects should be more willing to learn about cloud computing if they find this technology to be secure. We, therefore, posit that:

H4: Perceived security (PS) positively affects subjects' acceptance of cloud computing as part of their curriculum.

Perceived Speed of Access of Cloud Computing (PSA)

The speed of using applications over the Internet can be a factor that has prevented cloud computing from being a viable option for outsourcing IT operations. Users may be unaware that the use of applications via the Internet still allows them to retrieve the data at the same speed as when the data is stored on their personal computers. For cloud computing to be widely accepted, it is crucial that the services must allow users to access data at a reasonable speed. Users' perception on the speed of access should influence their intention to use cloud computing service. Hence, we posit that:

H5: Perceived speed of access (PSA) positively affects subjects' acceptance of cloud computing as part of their curriculum.

Perceived Cost of Usage of Cloud Computing (PCU)

A survey conducted by ComputerWorld magazine (Wood, 2011) of IT professionals revealed that while "saves money" ranked first on the list of cloud computing key benefits, "costs more" ranked third on the list of drawbacks suggesting that the issue of cost in outsourcing information technology resources is a complex issue. In many public and private sector industries, including education, federal and state government, and telecommunications, cloud computing systems are being pilot-tested and implemented to save IT costs and improve performance (Behrend et al., 2011). Based on the foregoing, we hypothesize that:

H6: Perceived cost of usage (PCU) positively affects subjects' acceptance of cloud computing as part of their curriculum.

METHODOLOGY

The questionnaire designed for this study adapted the instrument and scales developed for the TAM with additional constructs added as described in the proposed research model.). The questions used to measure the additional constructs were adapted from prior studies (Pikkarainen et al., 2004; Venkatesh & Davis, 2000, Vijayasarathy, 2004; Wong & Hsu, 2008). (Pikkarainen et al., 2004; Venkatesh & Davis, 2000, Vijayasarathy, 2004; Wong & Hsu, 2008; Venkatesh et al., 2003).

The questionnaire consists of thirty-seven (37) questions. Thirty questions with the five point Likert scale were designed to measure subjects' perceptions on cloud computing and whether they believe it should be integrated into their curriculum. The remaining seven questions were asked to gather some demographic data on the subjects. To validate the clarity of these questions, three professors and three researchers were asked to read through the survey questions. Revisions to the survey were made based on the feedback received.

The surveys were administered to students at a Northern University in Thailand. These students are certainly part of the target group for companies providing cloud computing services. Five hundred and forty-eight (548) subjects participated in this study. All 548 surveys were returned completed so all 548 of the surveys were used in this study. The participants' demographics are shown in Table 1 below.

Table 1: Subjects' demographics (n=548)

	No.	%		No.	%
Gender			**Level of Education**		
Male	110	20.07	Above bachelor degree	26	4.74
Female	438	79.93	Bachelor degree	474	86.50
No answer	0	0.00	Below bachelor degree	33	6.02
			No answer	15	2.74
Age			**Monthly Income (in Thai Baht)**		
15 - 24	460	83.94	Below 15,000	178	32.48
25 - 34	46	8.39	15,000 - 30,000	95	17.34
35 - 44	10	1.82	30,001 - 50,000	22	4.01
above 44	32	5.84	Above 50,000	10	1.82
No answer	0	0.00	No answer	243	44.34
Use Cloud Computing Before			**Employment Status**		
Yes	350	63.87	Un-employed	414	75.55
No	164	29.93	Part-Time	36	6.57
No answer	34	6.20	Full-Time	98	17.88

DATA ANALYSIS AND DISCUSSION

The data analysis for this study was conducted using SPSS 25.0 and AMOS 24.0 statistical software. A reliability test was conducted to examine the internal consistency of the research instrument. The test confirmed the reliability with Cronbach's alpha coefficient of 0.975. In addition, since multicollinearity can have harmful effects (Cenfetelli and Bassellier 2009), multicollinearity was assessed for all of the indicators in the research model. The results revealed that the multicollinearity is not an issue with this data set.

Confirmatory factor analysis with varimax rotation was also conducted to examine the construct validity and to verify the groupings of the survey items adopted from previous studies. The results of the factor analysis confirm that the thirty survey items distributed themselves into six factors (see Table 2). The survey items which recorded a value below the suggested reliability level of 0.5 (Hair et al. 2009) were removed from the data analysis.

Table 2: Factor analysis

Constructs	Component					
	1	2	3	4	5	6
PU1	.259	.283	.271	.243	.215	**.743**
PU2	.258	.261	.286	.236	.202	**.756**
PEU4	.251	.283	.272	.181	**.760**	.166
PEU5	.322	.252	.212	.211	**.744**	.228
PS1	.156	.209	**.797**	.213	.237	.163
PS2	.231	.209	**.805**	.227	.163	.169
PS3	.273	.210	**.776**	.234	.110	.212
PSA3	.277	.265	.288	**.733**	.177	.208
PSA4	.269	.282	.280	**.767**	.177	.157
PSA5	.354	.261	.248	**.734**	.148	.210
PCU2	**.759**	.281	.268	.259	.069	.127
PCU3	**.811**	.224	.200	.196	.200	.152
PCU4	**.772**	.254	.218	.208	.227	.203
PCU5	**.720**	.260	.137	.272	.266	.187
ACC1	.201	**.782**	.142	.210	.139	.157
ACC2	.208	**.776**	.179	.228	.199	.062
ACC4	.254	**.747**	.168	.135	.216	.186
ACC5	.262	**.735**	.253	.196	.085	.243

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.[a]

Structural Equation Model (SEM) and Hypothesis Testing

Properties of the causal paths including standardized path coefficients are presented in Figure 2. The results of hypothesis testing are shown in Table 3.

Figure 2 Structural equation model path analysis

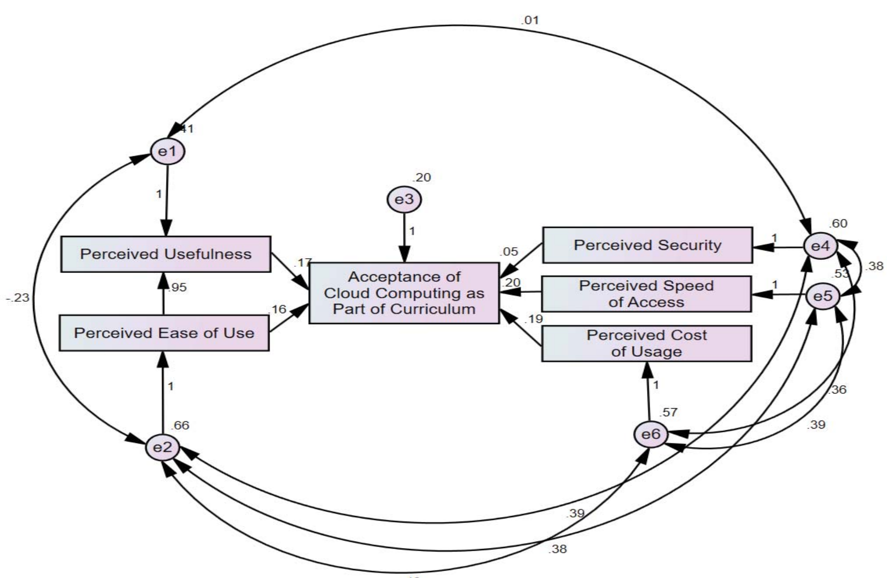


Table 3: Hypothesis testing and results

H#	Hypothesis Testing			Standardized estimate (β)	Critical Ratio	p-value
1	Perceived Usefulness	→	Acceptance of Cloud Computing as Part of Curriculum	0.169	0.039	***
2	Perceived Ease of Use	→	Acceptance of Cloud Computing as Part of Curriculum	0.157	4.453	***
3	Perceived Ease of Use	→	Perceived usefulness	0.948	19.455	***
4	Perceived Security	→	Acceptance of Cloud Computing as Part of Curriculum	0.054	1.467	0.142
5	Perceived Speed of Access	→	Acceptance of Cloud Computing as Part of Curriculum	0.203	4.735	***
6	Perceived Cost of Usage	→	Acceptance of Cloud Computing as Part of Curriculum	0.190	4.702	***

*** indicates significance level < 0.001

Hypothesis 1 examined the relationship between perceived usefulness and subjects' acceptance of cloud computing as part of their curriculum revealed a p-value of 0.000 making it significant at the <.001 level of significance. Thus, the data demonstrates support for H1. The presence of perceived usefulness is positively related to subjects' willingness to accept cloud computing as part of their curriculum. This finding indicates that the higher level of usefulness they perceived in using cloud computing, the more likely they are willing to learn more about this technology.

Regarding Hypothesis 2, the results also show that there is a positive relationship between subjects' perceived ease of use and their willingness to accept cloud computing as part of their curriculum, with the p-value of 0.000 making it significant at the < .001 level of significance. Thus, the data demonstrates support for H2. The result reveals that subjects tend to accept cloud computing as part of their core curriculum if they perceive that the technology of cloud computing is not difficult to use.

Hypothesis H3 examined the relationship between perceived ease of use and perceived usefulness revealed a p-value of 0.000 making it significant at the < .001 level of significance. Thus, the data demonstrates support for H3. This is in line with the connection discovered by Davis et al. (1989), which indicates that the perceived ease of use can positively influence subjects' perceived usefulness.

Hypothesis H4 examined the relationship between perceived security and subjects' acceptance of cloud computing as part of their curriculum revealed a p-value of 0.142 indicating that there is no support for the relationship between these two constructs. This finding is quite interesting. This may be because Thai students usually do not store the confidential data in the cloud storage and thus security is not viewed as an important factor to them. In other words, regardless of whether they believe in the security of cloud computing technology, their belief will not affect their willingness to learn this technology.

Regarding hypothesis H5, the results shows a positive relationship between the constructs perceived speed of usage and subjects' willingness to accept cloud computing as part of their curriculum with p-value of 0.000. It is pretty obvious that the speed of access should be considered a crucial feature of cloud computing. Subjects who view cloud computing as having a good speed of access are more likely to want to learn about this technology and thus willing to accept it as part of their core curriculum.

Hypothesis H6, examined the relationship between perceived cost of usage and subjects' acceptance of cloud computing as part of their curriculum revealed a p-value of 0.000 making it significant at the < .001 level of significance. Thus, the data demonstrates support for H6. This finding indicates that subjects who perceive the cost of the cloud computing as low are willing to learn more about the cloud computing technology and are likely to accept it as part of their core curriculum.

CONCLUSION

The results in this study reveal that four out of the five factors play an important role in encouraging Thai students to accept cloud computing as part of their core curriculum. These factors are perceived usefulness, perceived ease of use, perceive speed of access, and perceived cost of usage. The results reveal that if students perceive cloud computing as a technology that will be beneficial to them, they will be generally open to using it as well as learning it as part of their curriculum. In addition, the students today grew up surrounded by several technologies, they thus will be more likely to accept cloud computing if their "perceived ease of use" is high, which means that less effort is required to learn how to use it.

The result in this study also reveals that the speed of access and the cost of using cloud storage are also crucial factors that can influence Thai students' acceptance of cloud computing as part of their curriculum. Students prefer technology that allows them to get the thing they want in a timely manner. In addition, it does not matter how great a service a new technology can render, it will be useless if students cannot afford the cost of using it. Students will be more likely to accept cloud computing technology if they perceive that it does not cost them much to use the technology.

Interestingly, this study does not find a relationship between students' willingness to learn cloud computing and their perception about its security. This finding may be unique to Thai students as most of them do not store sensitive data in the cloud storage. So, regardless of whether the cloud storage is secure, it does not impact their willingness to learn about this technology.

There is an inherent limitation in this paper. The sample in this research was limited to subjects in one university. Although there was an attempt to gather the data from a variety of courses in the university, future research should be conducted at multiple universities. Further research should also consider expanding demographics to include users in various countries. In addition, a future study conducted could investigate in more details which feature of cloud computing help increase students' effectiveness and efficiency in the classroom.

REFERENCES

Almubarak, S. (2017). Factors influencing the adoption of cloud computing by Saudi university hospitals. *International Journal of Advanced Computer Science and Applications*, 8(1), 41-48.

Bagozzi, R.P. (2007). The legacy of the Technology Acceptance Model and a proposal for a paradigm shift. Journal of the Association for Information Systems, 8, 244-254.

Behrend, T., Wiebe, E.N., London, J., & Johnson, E. (2011). Cloud computing adoption and usage in community colleges. Behavior & Information Technology, 30(2), 231-240.

Blue, E., & Tirotta, R. (2011). The benefits and drawbacks of integrating cloud computing and interactive whiteboards in teacher preparation. *TechTrends*, 55, 31-39.

Cenfetelli, R. T., & G. Bassellier. G. (2009). Interpretation of formative measurement in information systems research. *MIS Quarterly,* 33(4), 689–707.

Changchit, C. (2008). Data protection and privacy issue. *Journal of Information Privacy and Security*, 9(3), 1-2.

Changchit, C. (2015). Cloud Computing - Should it be integrated into the curriculum? *International Journal of Information & Communication Technology Education*, 11(2), 105-117.

Changchit C. & Chuchuen, C. (2018). Cloud computing: An examination of factors impacting users' adoption, *Journal of Computer Information Systems*, 58(1), pp. 1-9.

Chau, P.Y.K., & Hu, P.J.H. (2001). Information technology acceptance by individual professionals: a model comparison approach. Decision Sciences, 32, 699-719.

Chen, L., Liu, Y., Gallagher, M., Pailthorpe, B., Sadiq, S., Shen, H. T., Li, X. (2012). Introducing cloud computing topics in curricula. Journal of Information Systems Education, 23(3), 315-324.

Davis, F.D., Bagozzi, R.P., & Warshaw, P.R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management Science, 35, 982-1003.

Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Addison-Wesley, Reading, MA, USA.

Flavián, C., & Guinalíu, M. (2006). Consumer trust, perceived security and privacy policy: three basic elements of loyalty to a web site. *Industrial Management & Data Systems*, 106, 601-620.

Hair, J., W. Black, B. Babin, and R. Anderson. 2009. *Multivariate data analysis: A global perspective*. 7th ed. Upper Saddle River: Prentice Hall.

Leesa-Nguansuk, S. (2018). AWS: Cloud computing to skyrocket. Retrieved from https://www.bangkokpost.com/tech/local-news/1411854/aws-cloud-computing-to-skyrocket

Pikkarainen, T., Pikkarainen, K., Karjaluoto, H., & Pahnila, S. (2004). Consumer acceptance of online banking: an extension of the Technology Acceptance Model. Internet Research, 14, 224-235.

Singh, S.P, & Veralakshmi, R.S.R. (2012). Cloud computing: A promising economic model for library and information centers. DESIDOC Journal OF Library & Information Technology, 32, 526-532.

Venkatesh, V., & Davis, F.D. (2000). “A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science, 46(2), 186-204.

Venkatesh, V., Morris, M.G., Davis, G.B. & Davis, F.D. (2003). User acceptance of information technology: toward a unified view. MIS Quarterly, 27, 425-478.

Vijayasarathy, L.R. (2004). Predicting consumer intentions to use on-line shopping: the case for an augmented Technology Acceptance Model. Information and Management, 41, 747-762.

Wong, Y.K. and Hsu, C.J., (2008). A confidence-based framework for business to consumer (B2C) mobile commerce adoption. Personal and Ubiquitous Computing, 12(1), 77-84.

Wood, L. (2011). Cloud storage a steep climb. *ComputerWorld*, 20-30.

Wu, C. (2013). Learning attitude and its effect on applying cloud computing service to it education. *International Journal of u- and e- Service, Science and Technology*, 6(1), 39-48.

Ya-Ching, L. (2019). Adoption Intention of cloud computing at the firm level, *Journal of Computer Information Systems*, 59(1), 61-72.

Yenisey, M., Ozok, A., & Salvendy, G. (2005). Perceived security determinants in E-commerce among Turkish university students. *Behavior & Information Technology*, 24, 259-274.

DECISION SCIENCES INSTITUTE

The Use of Decision Support Systems Tools to Improve Efficiency in Supply Chain Management

Aurore J. Kamssu
Tennessee State University
Email: akamssu@tnstate.edu

Jean-Claude Pedjeu
Tennessee State University
Email: jpedjeu@tnstate.edu

Christopher Hopkins
Tennessee State University

Mina Ramzy
Tennessee State University

ABSTRACT

In the past few years, supply chain management has been an important business function creating competitive advantage. Companies worldwide have strived to develop and improve their supply chain networks. Incorporating information technology and Decision Support Systems in Supply Chain Management operations has help companies achieve the desired competitive advantage.

KEYWORDS: supply chain management, decision support systems, information technology tools, Walmart, Amazon, competitive advantage

INTRODUCTION

Companies such as Amazon and Walmart are the leaders on the current competitive online marketplace. It would be beneficial to understand how those companies were able to reach and sustain the competitive advantage. It would be interesting to know if Decision Support Systems (DSS) tools play in a role in supply chain management (SCM) functionality and secure competitive advantage to those companies online.

In 2017, an American research firm specialized in Information System (IS) related insight to IT and business leaders in functional areas like Supply chain, IT, and Marketing, ranked both Walmart and Amazon in the Gartner Top 25. Amazon on the one hand, reaching Masters Status for being in the top five, seven times in the past ten years; Walmart on the other hand secured the 18th place. As information technology became an integral part of businesses focused on communicating and interacting with customers, businesses are able to react dynamically and respond rapidly to both, internal and external changes (Hershman, 2004). As it is common knowledge that companies like Amazon and Walmart are the leaders in today business world, it is difficult to determine to what extend their domination can be attributed to utilization of DSS throughout their SCM strategies. The current paper analyzed the impact of DSS on the SCM of both Amazon and Walmart.

COMPANY OVERVIEW – 2016 YEAR END

The choice of Amazon and Walmart derives from the fact that one company started as a brick-and-mortar storefront (Walmart) and the other as fully an online company (Amazon). Today, both companies are present online and essentially fighting for the same customers. Let us first give a brief summary of each company.

Company	**Industry**	# of EE's	Revenue	Profit Margin
Walmart Stores Inc.	Discount, Variety Stores	2,300,000K	487.51B	2.79%
Amazon.com Inc.	Online Shopping	341,400K	150.12B	1.28

Amazon.com Inc.

Amazon is a Fortune 500 online company established in Seattle, Washington. The company is known for being one of the first large company to sell goods over the Internet. Founded in 1984, Amazon Inc. started out as an online bookstore and added other unique items like DVDs, music, video games, electronics, and apparel. The Company also offers services, such as database options, publishing opportunities, digital content subscriptions, and advertising. It functions through three divisions: North America, International, and Amazon Web Services (AWS).

Amazon has become what one may identify as an online mall where third parties sell goods of virtually any kind. Amazon.com provides customers with the most variety of packaged products online in the marketplace. The company's strategy of continuous improvement has encouraged its leaders to explore different business opportunities, which resulted in the creation of separate business franchises. Amazon can be credited for popularizing online shopping in the world today. The company is very customer-centric. Amazon believes that the company's success depends greatly on customers' satisfaction. According to Schneider (2017), "Not putting the customer first, will certainly lead to total failure".
Business intelligence (BI) is at the core of Amazon success and business strategy. Thanks to BI technology, Amazon can personalize product recommendations and market products, uses BI software tools to make successful logistical business decisions. In-depth data analysis is what enables Amazon's massive supply chain to run smoothly. "From optimizing shipping routes to allocating inventory among warehouses, data and BI tools influence practically every step of Amazon's supply process", reports CCS Technology , an IT Service company (CCS Technology Group , 2018).

Walmart Stores Inc.

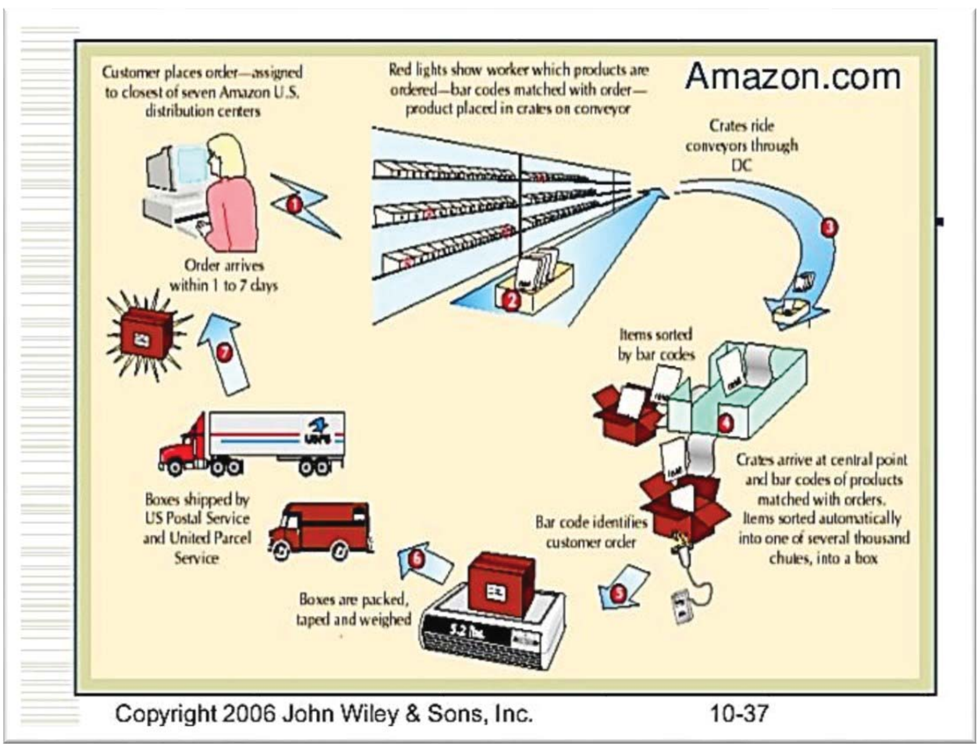
Walmart began as a discount store in Rogers, Arkansas, selling products for less. It rapidly grew in the second half of the century into the largest retail company in the world. Each week, over 260 million customers visit the more than 11,000 stores around the world. The company has approximately 2.3 million employees worldwide and is the number one employer in the US workforce. Wal-Mart sales a wide variety of products in their stores as well as online. From grocery and entertainment to sporting goods and apparel, Wal-Mart carries a good mix of products preferred by customers for its low prices. For many customers, Wal-Mart is a one-stop-shop for basic goods and services such as grocery products, health and beauty, household goods, pharmacy, electronics, cellular service, entertainment, small appliances and many other necessary items.

With a revenue of $485.9 billion in 2016, Walmart generates the most revenue than any other company in the world. With supermarkets built at a rapid pace in urban areas around the world, Walmart was projected to gross over $500 billion dollars in sales. One of the company's key goals since 1962 is to help customers save money so they can live better worldwide. They use their size to provide access to high-quality goods at everyday low price.

SUPPLY CHAIN MANAGEMENT STRATEGIES

SCM - Amazon

Amazon's SCM system is very simple: a customer places an order and the order is received at the warehouse. Their SMC network uses barcode technology to match the orders with the merchandises in the warehouse. Then, Amazon uses conveyors to ship the merchandise to the central point where goods are sorted, customer matched, weighted, and packaged before shipping via USPS, FedEx, or UPS. The illustration below by John Wiley and Sons, Inc. provides visualization for Amazon's SCM process model.



Amazon wants to take advantage of every opportunity that can add business value to the company using the latest technological to stay ahead of competitors. In spite of technology, Amazon remains true to the company's focus, which is "The Consumer-Centric Approach". Over the years, Amazon has invested in minimizing delivery time to improve customers' satisfaction. The goal was to reduce most transactions to same-day delivery and ultimately achieving even less than one-hour-delivery for some products. Same-day delivery and less than one-hour shipping is not realistic for traditional shipping lines like UPS and FedEx to handle. For this reason, Amazon decided to insource some of their distribution by investing in their own fleet of distribution trucks to satisfy same-day deliveries. Today, Amazon in-house delivery service has grown so much that that e-commerce giant relationship with delivery/shipping service providers is in jeopardy. In June of 2019, FedEx decided to cut ties with Amazon. In addition, the US Postal Service (USPS) reported a volume decline in package shipments during its fiscal third quarter, citing "intense competition" from "certain major customers" that have formed in-house last-mile delivery services, which according to USPS's CFO Joe Corbett, is the carrier's first quarterly volume decline for packages in nine years.

According to Burnson (2016), Amazon has over 100 warehouses or fulfillment centers in the United States alone. These warehouses are strategically positioned outside major metropolitan areas to reach customers as quickly as possible. Therefore, reducing the delivery time from 2-days to Same-day for certain products. In addition, in the warehouses/fulfillment centers, Amazon has invested in robotics and automation by acquiring Kiva Systems in 2012, in order to achieve quality and efficiencies in the supply chain (Burnson, 2016). In the fulfillment centers worldwide, Amazon has more than 30,000 robots operating to pick, pack, and sort products with no human intervention. This robots help streamline Amazon's SCM process by cutting costs, improving efficiency and speed and therefore increasing customer satisfaction. Using robotics also reduce manual work for hazardous tasks, improving Amazon's working conditions therefore keeping employees safer.

Amazon has also been working on delivering products to customers by air via Drones called "Prime Air". These drones are intended for delivery within 30 minutes, with does not come as a surprise, given Amazon obsession with customers satisfaction with speedy delivery. However, Prime Air development has faced regulation obstacles in the United States (Burnson, 2016); but in other countries like India, Prime Air has a better chance to launch and thrive globally. Amazon has also been using artificial intelligence (AI) to make it easier for customers to buy what they like based on its recommendation engine.

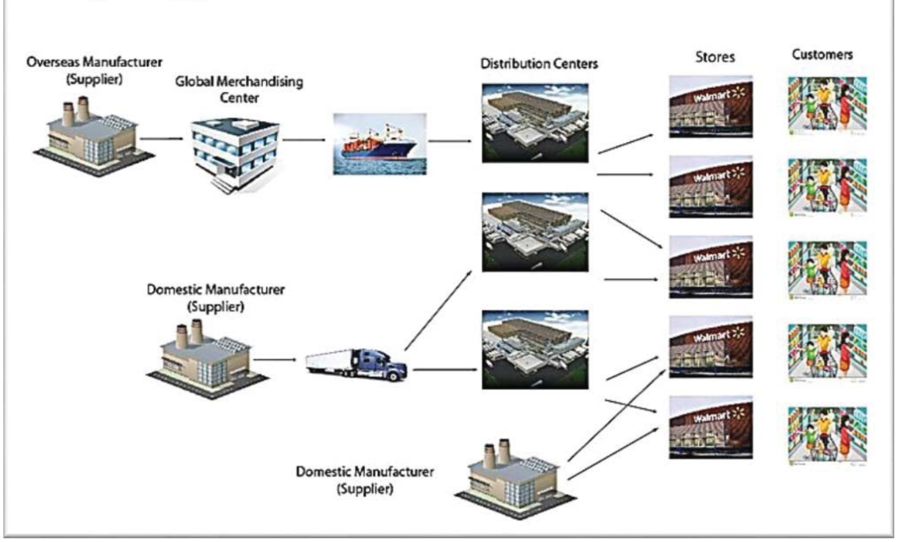
Amazon has innovated its user interface screens to streamline shopping and better serve customers from an overall user experience perspective. They have added the "Dash Button" which enables customers to order products with a click of a button. With just one-click, Amazon made shopping easy and quick. This feature allows Amazon to collect data at real-time and to provide insight to when the demand for a product rises or drops. Therefore, adjust the supply of each product accordingly to inventory levels. This strategy also led Amazon to developing predictive analytics technology. Creating a predictive shopping and purchasing system allows Amazon to deliver goods that customers will likely buy; since Amazon knows what customers are likely to buy. This technological strategy has revolutionized Amazon's supply chain and a brief list of innovative SCM strategies that Amazon incorporates throughout their business operations is presented in the table below.

Innovative Business Strategies: Amazon.com Inc.			
Company	**SCM Strategy**	**Business Impact**	**IS System Involvement**
Amazon	Predictive Shopping and Purchasing	Delivering goods that customers that will likely buy even before they order them.	Yes
Amazon	Army of 30,000 robots	Automated warehouse solutions: Picking, Packing and Sorting Shipments.	Yes
Amazon	Amazon Delivery - Reduce deliver time	Own fleet of delivery vehicles to fulfill same-day deliveries for customers.	Yes
Amazon	Drone Delivery system	Customers who live within 10 miles of a fulfillment center will be able to receive their packages via drones within 30 minutes or less.	Yes
Amazon	Dash Button	Streamlined Shopping: Allows users to simply press a button to order household goods.	Yes

SCM - Walmart

Walmart is listed in Gartner's top 25 companies as having one of the best supply chains in the world. Walmart's SCM strategy including but not limited to product cost reduction, inventory storage cost reduction, and highly competitive pricing. This strategy has enabled Walmart to maintain a sustainable top position against competitors for many decades. A key to Walmart success resides in its focus and investments in innovative technology, especially decision support systems to improve all its business functions including supply chain. Technology is the foundation of Walmart's supply chain strategy.

Figure 2: Supply chain model of Wal-Mart



Inventory storage cost can be so expensive, so incorporating an innovative system to cut inventory storage costs is a tremendous game changer in the retail industry, which proves why Walmart is very profitable because of its successful SCM.

Innovative Business Strategies: Wal-Mart Stores Inc.			
Company	SCM Strategy	Business Impact	IS System Involvement
Wal-Mart	Strategic Vendor Partnerships	Provide potential long-term and high volume purchases in exchange for the lowest possible prices.	Yes
Wal-Mart	Cross Docking	Strategy to replenish inventory efficiently	Yes
Wal-Mart	Predict Inventory Levels	Ability to proactive determine inventory levels to replenish inventory efficiently	Yes
Wal-Mart	RFID Technology	Ability to track pallets of merchandise moving along the supply chain.	Yes
Wal-Mart	Data Harmonization and Data learning	Integrating data from many different sources which will drive new business insights, decreased costs, and improved data accuracy. Also creating algorithms to provide continuous improvement in the supply chain.	Yes

One major SCM business driver for Walmart is the strategic vendor partnership (SVP) system. In this system, Wal-Mart commits with vendors/Suppliers to create strategic sourcing opportunities for products at the lowest possible prices. This can only be achieved by Walmart guaranteeing suppliers a long-term and high capacity purchase in exchange for the lowest price points on the market. In turn Walmart can offer its customers very competitive prices. This SCM partnership also restraints Wal-Mart to streamline purchasing by working with only select suppliers.

Walmart is also a leader in innovative logistic practices. Using the “Cross Docking” tactic, Walmart is able to strategically replenish inventory by direct handover of products directly from sending trucks to receiving trucks therefore, bypassing storage costs. Another cost cutting strategy is Walmart’s cross docking also known as, “centerpiece” strategy, which is an inventory tactic to replenish inventory efficiently (Lu, 2014). This strategy cuts costs, reduces transportation time, and keeps balanced inventory levels.

Walmart also uses radio frequency identification (RFID) technology like no other company in the world. RFID is no only a DSS tool, but it is another innovative technology with the ability to scan numerical codes from a distance to track merchandise. Walmart operates RFID technology to track pallets of merchandise moving across the entire supply chain network. Walmart also mandates that their suppliers/vendors use RFID tagging to control what products are sent through their distribution channels. This innovative SCM tactic improves overall efficiency by tracking products, sorting merchandises, reducing driving, and managing delivery time. Below is an illustration of Walmart’s SCM network.

Walmart is also a forerunner in utilizing big data to drive business insight within its functional areas. With billions of data transactions between Walmart’s distribution centers, stores, suppliers and customers, the company has the capability to collect a vast amount of data on their customer needs a well as the marketplace. This allows Walmart to gain unique knowledge and business insight on the market conditions giving them a competitive edge over their competitors. Walmart integrates data from various sources including suppliers, customers, prices, shipping and many other valuable segments. In order to collect data from all these sources, Walmart has to invest in advance information systems that interface with various data sources to collect and process large amounts of data. Walmart SCM strategy uses data analytics to learn about the entire supply Chain environment. Data analytics help drive additional business insight and make better decision in the marketplace.

SCM Commonalities between the Two Companies

Amazon and Walmart are leaders in the marketplace and have been able to maintain their competitive advantage for decades. Both companies' strengths seem to reside in their SCM strategies. On the one hand, Amazon uses DSS tools to create strategic vendor partnerships; on the other hand, Walmart tracks inventories for competitive advantage using RFID technology. A summary common innovative DSS tools and SCM strategies used by Amazon and Walmart, to improve their supply chain, is in the table below.

Common Innovative SCM Strategies		
Common SCM Strategies	Business Impact	IS System Involvement
Strategic Vendor Partnerships	Provide potential long-term and high volume purchases in exchange for the lowest possible prices.	Yes
Predict Inventory Levels	Ability to proactive determine inventory levels to replenish inventory efficiently	Yes
RFID Technology	Ability to track pallets of merchandise moving along the supply chain.	Yes
Data Harmonization and Data learning	Integrating data from many different sources which will drive new business insights, decreased costs, and improved data accuracy. Also creating algorithms to provide continuous improvement in the supply chain.	Yes

FOCUS GROUP - DSS TOOLS AND SCM ANALYSIS

It is clear that Decision Support Systems tools play an important role in Supply Chain Management of both Amazon and Walmart. On the one hand, the companies could not operate efficiently without automation such as managing supplier transactions and order entries from customers; on the other hand, it would be almost impossible those companies to deliver such amount of products/goods without the information systems to streamline their supply chain management strategy. The Supply chain management systems of both companies therefore gives them gives them more efficiencies in their operations and enable them to expand, grow and become more profitable and competitive. Other innovative DSS and SCM tools that can be used gain competitive advantage are best practices and presented in table below.

The competition between Amazon and Walmart is only going to widen for many reasons. Amazon investment in groceries and gourmet food might have caught many by surprise as did their in-house delivery service. But, just like auto makers, as Walmart.com has become amazon Inc. fiercest competitor, it makes sense that the e-commerce giant brings the competition to where it can hear Walmart hard. The introduction of free shipping by Walmart.com, a service Amazon Inc. offers mainly for its prime members, is one way to gain competitive advantage over Amazon. Since Walmart Labs was originally created to build and test new technology for Walmart.com; they had evolved and focus on getting Walmart.com to a place of operational efficiency: rebuilding the search engine, focusing on social / local / mobile (SoLoMo) and building a new platform.

Participants List Innovative DSS Tools	Amazon and Wal-Mart Current Compatible DSS/SCM Tools	Company
Predictive Inventory	Predictive Inventory Levels	Both
ERP Systems	Strategic Vendor Partnerships	Both
Freights Service Management	Amazon Delivery – Reduce deliver time	Both
Lean Cost Management System	Slash Inventory, Cross Docking, Strategic Vendor Partnerships	Both
Order Management – Predictive Behavior Analysis	Dash Button	Amazon
Automated Warehouse System	Army of 30,000 Robots	Amazon
Inventory Tracking System	RFID Technology, Predict Inventory Levels	Both

One great advantage Walmart has is that it is first and foremost a retail company that recognizes technology as the critical enabler, behind the company's associates, in helping achieve success. Walmart Labs is comprised of: (i) Customer Technology that focuses on customer's shopping experience; (ii) Merchant Technology overseeing our online and in store search (pricing, catalog, content, inventory, replenishment, and fulfillment technology) to provide optimal online shopping experience; (iii) Supply Chain Technology ensuring that customers receive their orders as fast as they want and where they want in a predictable way; (iv) Global Cloud and (v) Global Data Analytics team warrants that Walmart's decisions are data-driven, smarter, faster, and highly-automated. This strategy has enabled Walmart eCommerce and store technology teams to increase operating efficiencies while decreasing enterprise expenses.

As both companies compete at the eCommerce rink, Walmart's tremendous and lengthy successful experience in retail stores is something Amazon is still learning. Amazon name as a store is not as a household name as Walmart is. A 3-year old child knows that his/her toy will be found in a Walmart store or a Target store. Walmart platform also enable customers to shop online and pick up in a nearby Walmart retail store, which Amazon does not possess. While picking a product ordered online, customers are likely to shop in-store. While both companies use similar DSS tools, Walmart competes locally against stores like Publix, Kroger, Targets, Best Buy, to name a few. Its platform is two-fold: traditional mortar and brick store and online.

CONCLUSION

It is evident that leaders like Amazon and Walmart use Decision Support Systems (DSS) tools to improve their in supply chain management to gain and maintain competitive advantage. These companies constantly use DSS tools to manage their daily operations. Both companies have similar innovative SCM strategies, which earned them a spot on the Gartner's Top 25.

REFERENCES

"Amazon.com Inc.". Yahoo.com. Yahoo Finance, 2017 https://finance.yahoo.com/quote/AMZN?p=AMZN 11-13-2017

"Wal-Mart Stores Inc." Yahoo.com. Yahoo Finance, 2017, https://finance.yahoo.com/quote/WMT?p=WMT 11-13-2017

Burnson, Forrest. "6 Ways Amazon Is Changing Supply Chain Management in 2016". www.softwareadvice.com. Software Advice, 2016, http://www.softwareadvice.com/resources/amazon-supply-chain-management/ 10-26-2017

Hershman, Jack "Robots seen as Companions for elderly." AI Magazine Summer 2004, Vol. 25. Proquest.

Lu, Clara. "Walmart's successful supply chain management". www.tradegecko.com. tradegecko, May 8, 2014 https://www.tradegecko.com/blog/incredibly-successful-supply-chain-management-walmart. 10-15-2017

Markman, Jon. "Amazon Using AI, Big Data to Accelerate Profits". Forbes, June 5, 2017 https://www.forbes.com/sites/jonmarkman/2017/06/05/amazon-using-ai-big-data-to-accelerate-profits/#61e286dd6d55

Ovenden, James. "Amazon's Supply Chain Process | Articles | Chief Supply Chain Officer." Articles Chief Supply Chain Officer | Innovation Enterprise. N.p., 24 Feb. 2017. Web. 08 Nov. 2017.

Paolo Taticchi, Patrizia Garengo, Sai S. Nudurupati, Flavio Tonelli & Roberto Pasqualino "A review of decision-support tools and performance measurement and sustainable supply chain management." Int. J. of Production Res., Vol. 53, Issue 21, 2015.

Simpson, Paul. "The Secrets behind Amazon's Success." Supply Management. N.p., 21 Jan. 2016. Web. 08 Nov. 2017.

Soskis, Benjamin. "Man and the Machines." Legal Affairs. Jan/Feb. 2011. LegalAffairs.org. 3 Nov. 2017.

Waller, Matt. "How Sharing Data Drives Supply Chain Innovation". www.industryweek.com. Industry Week, August 12, 2013 http://www.industryweek.com/supplier-relationships/how-sharing-data-drives-supply-chain-innovation

DECISION SCIENCES INSTITUTE

Cost-benefit analysis of patient's choice in precision medicine

Yucheng Chen
University of Connecticut
2100 Hillside Rd, Storrs, CT 06268
Email: yucheng.chen@uconn.edu
(860) 486-6485

Manuel Nunez
University of Connecticut
2100 Hillside Rd, Storrs, CT 06268
Email: manuel.nunez@uconn.edu
(860) 486-1333

ABSTRACT

Precision medicine is a high-cost complex medical procedure that provides targeted treatments according to the individual characteristics of each patient. We present a preliminary cost-benefit dynamic model to assess learning curve effects in cost reduction through interaction with a centralized database repository. We also develop a simulation model to study the dynamics of the precision medicine approach and discuss insights derived from this model.

KEYWORDS: Health Care, Bioinformatics and Biological Data modeling, Cost Benefit Analysis, Knowledge Acquisition, Database Design for Big Data and Analytics Applications, Simulation

INTRODUCTION

Precision medicine is the use of personalized data to provide, integrate, and interpret information about an individual's health and better manage it (Snyder, 2016). Individual data are compiled from sources such as sequencing of the DNA (genome), measuring of biomolecules in the body, sensors to continuously monitor physiology and activity, and studying the microbial community that resides in the body. The cost of sequencing genomes has decreased dramatically in recent years and has had a large impact on personal genomics. In particular, one area of highest impact is cancer, which will strike 40% of people in their lifetime (Snyder, 2016). Traditionally, when a patient is diagnosed with cancer, he/she will be treated by a general procedure targeting the symptoms of the disease. Since the body cannot afford too many trials, only a limited number of treatments can be applied to the patient. The treatment and the drugs used by the physicians must be proven to be effective and be legally approved and registered. For many cancer patients, treatments that were proven effective in the past do not work well. This is because, genetically every person is different, even for the same pathogenesis a drug or a general treatment cannot be effective for all the patients. Further, even if the drug or treatment has proven to be effective for the given pathogenesis in the past, the treatment can still be ineffective because of the dosage or the combination of adjuvants that does not fit well the individual characteristics of the patient. Thus, the effectiveness of traditional methods depends heavily on the similarity between the personal characteristics of the tested group during the general treatment development and approval phase.

On the other hand, precision medicine provides an alternative to treating patients. It classifies individuals into subpopulations whose members (and their tumors) are genetically similar, and probably can be treated by the same therapies (Timmerman, 2013). Precision medicine for the treatment of cancer is based on the context of a patient's genetic content (Lu et al., 2014) or other molecular or cellular analysis. It usually consists of a series of activities including gene sequencing, pathogenesis identification, disease database matching, customized medicine development experiments, database updating, and so on. Development of a personalized treatment requires the use of intelligent-platform experiment-based techniques involving the use of patient's cell cultures for multiple screening runs of different depth (i.e., dosage, toxicity, functional test, molecular targets, etc.) and outcome analysis guided by machine-automation and machine learning to improve the efficiency (Tang-Schomer, 2017). Therefore, personalized treatments are more expensive than general traditional treatments and cost-benefit analysis is a major concern.

A less expensive approach is using a proven successful treatment of a similar case that matches the patient's characteristics. This experience-based approach relies in the use of a data repository of historical treatments generated as a by-product of previous experiment-based findings. One main factor to decide whether a patient should use a newly experiment-based personalized treatment as described before or an experience-based database treatment is the size of the data repository. As more experiment-based personalized treatments are performed and more data points are added to the database, finding a match between a current patient and a previously treated patient becomes more likely. Thus, as the database grows the number of experiment-based personalized treatments, and hence, the overall cost of the precision medicine approach is expected to decrease.

In this paper, we propose a tentative abstract model of the decision process of a patient, and his/her supporting healthcare providers, who is undergoing a precision medicine treatment and must decide whether to use a (more expensive) experiment-based treatment or a (less expensive) experience-based database treatment. In the model, the main driver of the decision is a cost-benefit analysis of the patient's situation that incorporates the expected cost and the likelihood of success of the treatments. This leads to finding an economic breakeven point to decide whether to match the patient's profile to an existing case, or to develop a new personalized treatment via experimentation. By considering information gathered from patients who choose to develop an experiment-driven personalized treatment into a database, we also hope to understand how information technology dynamics (as a learning process) affect the overall cost of this precision medicine approach. To do so, we develop a simulation model to study the evolution dynamics of the database and the overall cost of the precision medicine treatments as the number of patients grows through time. We then apply our simulation model to different scenarios, where we consider the likelihood of success in finding a matching treatment and, if there is a match, the likelihood of success of the treatment. Finally, based on our numerical analysis from the simulation, we provide insights about the learning process associated with precision medicine and some implications to public policy.

LITERATURE REVIEW

In recent years, precision medicine (PM) has become a popular term in the new medical treatment development frontier. Although people have widely discussed the advantages of the PM approach in the literature, adopting its prevention and treatment strategies in real medical practice is still at a very early stage (NRC, 2011). Until now, the medical industry has applied precision medicine to treat limited genetic related illnesses, especially some types of cancer such as lung cancer (Jamal-Hanjani et al., 2014) and metastatic breast cancer (Arnedos et al., 2015). The public recognizes the success of precision medicine, but big challenges are still hindering precision medicine from incorporating into our existing medical system. One challenge is whether people are able to afford the huge cost of precision medicine, because personalized treatments are always more expensive. Whether it is worthwhile to apply precision medicine treatments instead of traditional medicine treatments really depends on how people value the utility of the treatment based on his or her own condition. Finding a general measurement for the value of precision medicine is of great importance (Grosse et al., 2008). On the other hand, to target a disease caused by a certain defective gene(s) needs sophisticated decision algorithms, because a gene segment could be the source of several diseases (Jameson et al., 2015). Thus, whether we can extend precision medicine into the medical system in the future depends on whether we can make precision medicine affordable by enlarging biological databases, which entails identifying problematic genes and genetically classifying patients efficiently (Collins and Varmus, 2015).

Generally, there are two major aspects to consider when valuing the economic benefit of applying medical procedures: health outcomes and treatment cost. Correspondingly, in the literature researchers have used cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA) focusing on different aspects of the procedures (Grosse et al., 2008). Typically, CEA focuses on the cost per unit of "natural" health outcomes, such as the cost of a 1% increase in the five-year survival rate. CEA is suitable for a single target disease since it compares only one aspect of the outcomes at a time. One type of CEA is based on a very popular measure called the *quality-adjusted life year index*, which maps the life duration and quality into a comparable number (Klarman and Rosenthal, 1968). However, since it is very hard to define the state of "quality" of life, the method cannot be widely applied and is under debate (Prieto, 2003; Mortimer and Segal, 2008; Dolan, 2008). Another popular CEA is the *incremental cost-effectiveness ratio* (ICER), which is the ratio of incremental costs to incremental outcomes and it captures how an additional input affects the health outcomes. A patient can decide whether to undergo further treatment based on an ICER value. It is also controversial since it may limit the availability of treatments to patients because of healthcare rationing. In contrast, CBA is more comprehensive and can be applied in a wider range of contexts. It considers more than only the health outcomes, but also many non-health outcomes. The idea is to map all the benefits and losses into monetary units, and then to calculate the net benefit. One method for the mapping is to estimate the indirect value of regain/loss of health, such as future value of economic production when with and without health. A second approach of CBA is more customer-driven and is favored by most economists; it is called *willingness-to-pay* (WTP), which is the minimum amount of money that a person is willing to pay for the benefit from a treatment. The implementation of WTP is very helpful for decision makers to have criteria for choosing a plan or not. However, WTP depends on a subjective contingent perceived valuation of each person (Donaldson et al., 2002; Gafni, 1990; O'Brien and Viramontes, 1994; Jarrett and Mugford, 2006). All CBA results are expressed in monetary units, which makes it a good indicator for decision makers to compare different treatment scenarios (Krupnick, 2004), while CEA compares direct results, making it easy for healthcare system to see the outcomes (Oliver et al., 2002; Grosse et al., 2007).

When it comes to precision medicine, traditional CEA and CBA approaches as discussed above need to be revised and re-defined in some cases because of the personalized nature of the procedure. In particular, to the best of our knowledge, there have been very few attempts to do so. For instance, Grosse et al. (2008) introduce preliminary extensions of CEA and CBA for genetic testing; their paper also contains references to prior attempts of economic evaluations of genetic testing. Unfortunately, those approaches mostly focus on static information and do not consider the evolving nature of precision medicine and reducing costs from learning curves. In our research, we propose an original combination of the CEA and CBA approaches. To do so, we consider the monetary value of the precision medicine experiment and experience-based procedures (as in a CBA approach), as well as the likelihood of success of the procedures, to compute an economic break-even point (or indifference point) between procedures. From this, we derive a formula that compares the cost ratio of the procedures to the probability of success to determine (as in a CEA approach) which procedure to perform for each patient. Our approach also considers the learning dynamics as reflected by an increasing database of successfully treated patients using personalized procedures.

GENOME AS INFORMATION

Like many researchers (e.g., Calude and Paun, 2000), our approach to the structure of the genome is targeted to our goal of a cost-benefit analysis of a precision medicine procedure. Hence, we ignore a lot of the biochemical information not necessary for our analysis and instead, we present our models using a simplified version of the DNA. A DNA molecule is a polymer that serves as the instruction manual to guide the development process from a single cell to a complex individual. It is made up of four constituent parts (nucleotide bases). These bases are adenine (A), cytosine (C), guanine (G), and thymine (T), and they are paired as A - T and C - G. A human DNA molecule contains about six billion base pairs into 46 chromosomes. Hence, it can be seen as a finite string variable consisting of about six billion characters, each character taking values in {AT, TA, CG, GC}. In particular, assignment of values to this variable constitutes an individual's genome. Genomes differ by “variants”; a variant is a DNA sequence change. Approximately 3.8-4 million variants (one in every 1,200 bases) are single letter or base changes (called single nucleotide variant: SNV). There are 50,000-850,000 small insertions and deletions of 1-100 bases ("indels"). There are thousands of large insertions, deletions, inversions, and other types of chromosome rearrangements (some several hundred kilobases, 1 kilobase = 1,000 bases) and are called structural variants (Hartl, 2000; Snyder 2016).

The human genome sequence is a composite of several individuals yielding what is called the “reference” genome. The reference genome in its more finished form was completed in 2003 and it was prepared using DNA pooled from several individuals (thus it does not represent a single individual’s genome) (Snyder 2016). It consists of approximately 3 billion base pairs, and includes each of the 22 chromosomes plus the X and Y sex chromosomes. State of the art technology allows sequencing an individual’s genome in a few days. Variants are mapped when comparing millions of short sequenced fragments (100-150 bases in length) of an individual to the reference genomes. By comparing to known fragments, it can be deduced whether there is deletion, insertion, or inversion in the sequenced regions (Snyder 2016).

In bioinformatics, it is customary to quantify DNA similarity across individuals by modeling DNA as strings of characters and using an “edit” or Levenshtein distance to compute the similarity between strings (Gusfield, 1997). Concretely, we denote by V the set of all the different DNA molecules (strings) in the human population. For each v in V, we denote by $l(v)$ the length of v, that is, the number of characters in v. Hence, $v=(v_1,v_2,\ldots,v_{l(v)})$ is a string such that each character v_i is in {AT, TA, CG, GC}. The Levenshtein distance between two strings is the minimum number of single-character edits (insertions, deletions or substitutions) required to change one string into the other (Gusfield, 1997).

A key notion in our approach is that every genome can be matched to a number in the [0, 1] interval. Concretely, if we (arbitrarily) map the values AT, TA, CG, and GC to 0, 1, 2, and 3, respectively via a function (say) h, we can then create a map ϕ:V -> [0,1], that assigns to each string v in V a unique number of the form:

$$\phi(v) := \sum_{i=1}^{l(v)} h(v_i) 4^{-i}. \tag{1}$$

in the interval [0, 1]. Notice that ϕ is a bijection between V and $\phi(V)$; we call $\phi(V)$ the set of *numerical* genomes. This way, if there is a finite subset of E of V consisting of the genomic information of a group of patients, each individual genome in E can be represented as a point in the interval [0, 1]. Furthermore, if T denotes a set of personalized treatments that have been successful in treating the patients in E, then a genomic-treatment database D based on the individuals in E can be modeled as Cartesian product $D := \phi(E) \times T$, where for every pair (v, r), with v in E and r in T, we have ($\phi(v)$, r) in D.

The Levenshtein distance induces a distance in the set of numerical genomes $\phi(V)$, namely,

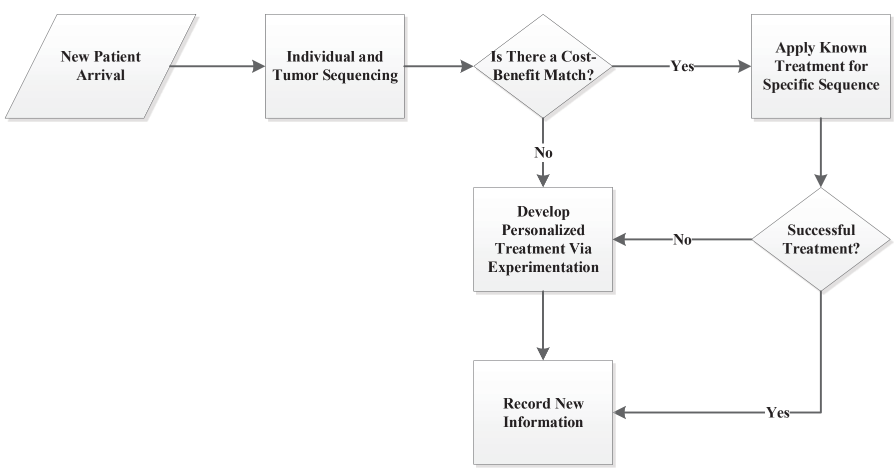
$$dist(x, y) := dist_L(\phi^{-1}(x), \phi^{-1}(y)), \tag{2}$$

for x, y in $\phi(V)$, where $dist_L$ is the Levenshtein distance. Therefore, we can determine the closeness between two individuals in a genomic database, or between one individual not in the database and one in the database, by using this induced distance.

COST-BENEFIT MODEL

We model the process of a patient's decision through a workflow as illustrated in Figure 1.

Figure 1: Precision medicine treatment workflow.



In this model, we assume that patients have been already diagnosed with a cancer (cancerous tumor) and they can be successfully treated if we are able to find an adequate (personalized) treatment. To start the process, the patient’s genomes, as well as the tumor’s, are sequenced for each new arrival. Next, the treatment database is searched for possible genome matches. The matching database string can be an exact or partial match based on the induced Levenshtein distance *dist* from (2). If there is a close enough database match, then the patient will decide whether to use this matching case based on a cost-benefit analysis that incorporates the likelihood of the success of the matching treatment. If the there is no database match or there is one but is not cost effective, then a personalized treatment will be developed from scratch via experimentation. If the database treatment is not successful, then the experiment-based approach will also be used. If the patient is treated using a matching experience-based case and this treatment is successful, or if the patient is treated by an experiment-based treatment, then the patient's genomic data and the details of his/her treatment are recorded into the database for future use.

Concretely, we denote by C_e the expected cost of developing an experimental treatment from scratch, whereas we denote by C_d the cost of applying a known-treatment from a matching database case, where we assume $C_e > C_d$. A patient chooses an experiment-based or experience-based treatment based on a cost-benefit analysis as follows. Let D_t denote the current genomic-treatment database at the time of arrival of the patient number t. From the discussion in the previous section, D_t consists of a finite number of pairs (x, r), where x in [0, 1] is a numerical representation of the genome of a previously treated patient and r denotes that patient's treatment. We denote by $X_1, X_2, \ldots, X_t, \ldots$ the numerical genomic values of patients in order of arrival to the system. We assume that the X_t are independent and identical distributed random variables in the interval [0, 1], and denote by f their corresponding common probability density function.

Given $\delta \geq 0$, we define the probability of a δ–matching of patient as the probability that there is at least one individual in database D_t whose induced Levenshtein distance is less than or equal than δ. Formally, let $B(x, \delta)$ be the subset of [0, 1] consisting of y such that $dist(x, y) \leq \delta$. For a given database D_t, define (abusing notation) the set of points in [0, 1] within δ distance from at least one element in D_t as

$$I_t(\delta) := \left(\bigcup_{y \in D_t} B(y, \delta) \right) \bigcap [0,1]. \tag{3}$$

Then, for x in [0, 1], the δ–matching probability is defined as

$$\pi_t(x, \delta) := \int_{I_t(\delta)} f(z)dz. \tag{4}$$

We denote the probability of a successful database treatment by p. Therefore, from a cost-benefit analysis point of view, a patient should first select an experience-based (database) treatment whenever we have

$$C_e \geq C_d + \left(1 - p\pi_t(x, \delta)\right) C_e, \tag{5}$$

given that $X_t = x$. In other words, from inequality (5), it follows that the experience-based treatment should be first selected when the ratio of the costs is below the probability of success of the database treatment:

$$\frac{C_d}{C_e} \leq p\pi_t(x,\delta). \tag{6}$$

Notice that the δ–matching probability is an increasing function of the size of the database because the more points added to it, the greater the integral region $I_t(\delta)$. Since the cost ratio on the left-hand side of (6) is less than 1, we conclude that, as long as new patients are added to the database, there will be a breakeven point (indifference point), where equality is attained in (6). After that point in time, it will be more cost-effective to perform all treatments using the database approach.

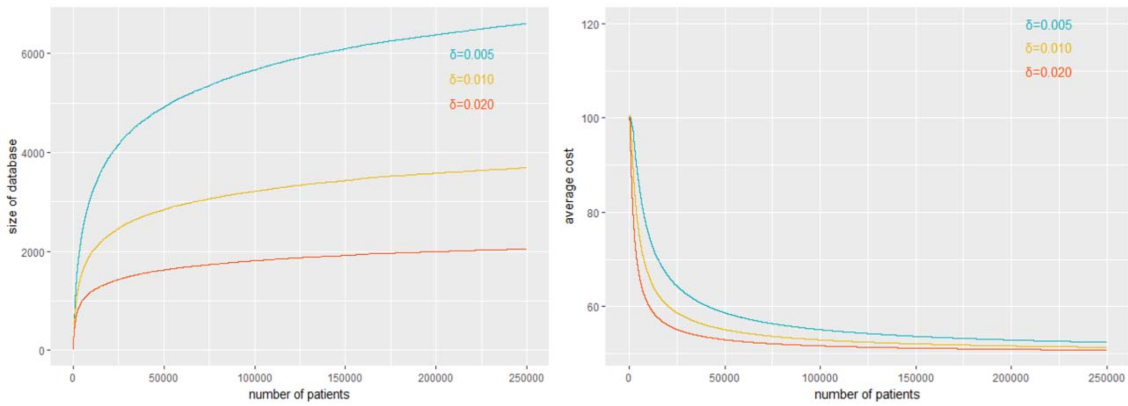
SIMULATION MODEL

We use a Monte-Carlo simulation model to study the dynamics of our proposed approach. In the simulation, for each scenario under consideration we generate a sample of 250,000 patients to go through the workflow from Figure 1. To simulate the matching process, in our experiments we assume that the density *f* (as described in the previous section) corresponds to the density of a uniform distribution. At each iteration of the algorithm, we randomly generate a point in the interval [0, 1] to represent the location of the new patient on the database.
The closeness of a match is determined by a proxy for the induced Levenshtein distance; namely, we use a simple absolute deviation difference between the randomly generated point corresponding to the new patient and the already existing points corresponding to existing database entries. In particular, a database point is considered a close match if the absolute difference between the new point and the database point is less than or equal to predefined value of δ. This is without loss of generality because if there are *k* tumor-causing variants and they are ordered such they represent the first *k* characters of the patient's genome string, then by choosing $\delta = 10^{-k}$ the two matching strings would agree in their first *k* characters (bases). Moreover, if an exact match is required, then we can always set $\delta = 0$ in the simulation.

NUMERICAL RESULTS

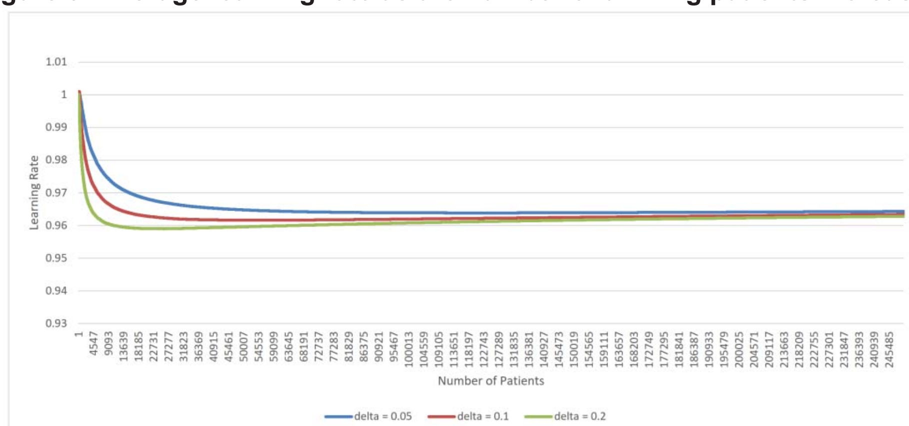
Using our simulation model as described in the previous section, we ran several scenarios. In particular, we considered three values of δ (i.e., 0.005, 0.10, and 0.20) to run our simulations. Figure 2 shows two graphs, the graph on the left corresponds to the growth of the database as the number of patients increases, whereas the graph on the right corresponds to the decrease in average cost per patient as the number of patients increases. The average cost per patient at time *t* is computed as the total accumulated cost up to time *t* divided by the number of arriving patients up to time *t*.

Figure 2: Evolution of the average database size and the average cost per patient as the number of arriving patients increases.



We can readily see that when δ is doubled, the size of database is almost halved. In addition, as δ increases, the average cost per patient decreases.

Figure 3: Average learning rate as the number of arriving patients increases.



With respect to learning curve effects, our results show an exponential the decrease of the average cost per patient at a constant speed exponent (see Figure 3). Moreover, the learning rate increases with δ. In our experiments we also tried other probability distributions (e.g., normal distribution) and found see that learning curve is not significant affected by how gene sequences are distributed, and it is more affected by the cost ratio between database and experimental treatments. When the ratio is low, the learning curve has a faster decrease speed rate.

DISCUSSION AND CONCLUSIONS

We have developed an original economic model of the precision medicine process that incorporates both cost-benefit and cost-effectiveness analysis. Our results show that as long as new patients continue being added to the genomic-treatment database, eventually the database will reach a size that will (1) provide a matching treatment for most patients and (2) the expected cost for patients will be more affordable than the cost from developing a personalized treatment from scratch (inequality (5)). Further, the convergence to this state of the database is exponentially fast in the number of treated patients.

Our formulation provides a decision criterion (inequality (6)) that is statistically computable in practice and it is useful for patients to decide the best approach when undergoing a precision medicine treatment. From a public policy point of view, a comprehensive database does not need a high amount of records, so that for the society benefit it makes sense to implement a subsidy policy for less wealthy patients with the objective of boosting the database growth.

REFERENCES

Arnedos, M., Vicier, C., Loi, S., Lefebvre, C., Michiels, S., Bonnefoi, H., & Andre, F. (2015). Precision medicine for metastatic breast cancer—limitations and solutions. *Nature reviews Clinical Oncology*, *12*(12), 693.

Calude, C., & Paun, G. (2000). *Computing with cells and atoms: an introduction to quantum, DNA and membrane computing*. CRC Press.

Cerrato, P., & Halamka, J. (2017). *Realizing the Promise of Precision Medicine: The Role of Patient Data, Mobile Technology, and Consumer Engagement*. Academic Press.

Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine*, *372*(9), 793-795.

Dhawan, D. (2017). Clinical next-generation sequencing: enabling precision medicine. In *Progress and Challenges in Precision Medicine* (pp. 35-54). Academic Press.

Ding, H., Takigawa, I., Mamitsuka, H., & Zhu, S. (2013). Similarity-based machine learning methods for predicting drug–target interactions: a brief review. *Briefings in Bioinformatics*, *15*(5), 734-747.

Dolan, P. (2008). Developing methods that really do value the 'Q'in the QALY. *Health Economics, Policy and Law*, *3*(1), 69-77.

Donaldson, C., Currie, G., & Mitton, C. (2002). Cost effectiveness analysis in health care: contraindications. *BMJ*, *325*(7369), 891-894.

Dove, A. (1999). Proteomics: translating genomics into products? *Nature Biotechnology*, *17*(3), 233.

Gafni, A. (1990). *Using Willingness-to-pay as a Measure of Benefits: What is the Relevant Question to Ask in the Context of Public Decision Making about Health Care Programs?* (No. 1990-15). Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada.

Grosse, S. D., Teutsch, S. M., & Haddix, A. C. (2007). Lessons from cost-effectiveness research for United States public health policy. *Annu. Rev. Public Health*, *28*, 365-391.

Grosse, S. D., Wordsworth, S., & Payne, K. (2008). Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genetics in Medicine*, *10*(9), 648.

Grosse, S. D., Rogowski, W. H., Ross, L. F., Cornel, M. C., Dondorp, W. J., & Khoury, M. J. (2010). Population screening for genetic disorders in the 21st century: evidence, economics, and ethics. *Public Health Genomics*, *13*(2), 106-115.

Gusfield, D. (1997). *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press.

Hartl, D. L. (1988). *A Primer of Population Genetics*. Sinauer Associates, Inc.

Jamal-Hanjani, M., Hackshaw, A., Ngai, Y., Shaw, J., Dive, C., Quezada, S., ... & Falzon, M. (2014). Tracking genomic cancer evolution for precision medicine: the lung TRACERx study. *PLoS Biology*, *12*(7), e1001906.

Jameson, J. L., & Longo, D. L. (2015). Precision medicine—personalized, problematic, and promising. *Obstetrical & Gynecological Survey*, *70*(10), 612-614.

Jarrett, J., & Mugford, M. (2006). Genetic health technology and economic evaluation. *Applied Health Economics and Health Policy*, *5*(1), 27-35.

Klarman, H. E., & Rosenthal, G. D. (1968). Cost effectiveness analysis applied to the treatment of chronic renal disease. *Medical Care*, *6*(1), 48-54.

Krupnick, A. J. (2004). *Valuing health outcomes: policy choices and technical issues*. Resources for the Future.

Lu, Y. F., Goldstein, D. B., Angrist, M., & Cavalleri, G. (2014). Personalized medicine and human genetic diversity. *Cold Spring Harbor Perspectives in Medicine*, *4*(9), a008581.

Mills, S., O'Sullivan, O. R. L. A., Hill, C., Fitzgerald, G., & Ross, R. P. (2010). The changing face of dairy starter culture research: from genomics to economics. *International Journal of Dairy Technology*, *63*(2), 149-170.

Mirnezami, R., Nicholson, J., & Darzi, A. (2012). Preparing for precision medicine. *New England Journal of Medicine*, *366*(6), 489-491.

Mortimer, D., & Segal, L. (2008). Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. *Medical Decision Making*, *28*(1), 66-89.

National Research Council. (2011). *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*. National Academies Press.

O'Brien, B., & Viramontes, J. L. (1994). Willingness to pay: a valid and reliable measure of health state preference? *Medical Decision Making*, *14*(3), 289-297.

Oliver, A., Healey, A., & Donaldson, C. (2002). Choosing the method to match the perspective: economic assessment and its implications for health-services efficiency. *The Lancet*, *359*(9319), 1771-1774.

Ginsburg, G. S., & Willard, H. F. (Eds.). (2016). *Genomic and precision medicine: foundations, translation, and implementation*. Academic Press.

Perlman, L., Gottlieb, A., Atias, N., Ruppin, E., & Sharan, R. (2011). Combining drug and gene similarity measures for drug-target elucidation. *Journal of Computational Biology*, *18*(2), 133-145.

Prieto, L., & Sacristán, J. A. (2003). Problems and solutions in calculating quality-adjusted life years (QALYs). *Health and Quality of Life Outcomes*, *1*(1), 80.

Shen, B., & Hwang, J. (2010). The clinical utility of precision medicine: properly assessing the value of emerging diagnostic tests. *Clinical Pharmacology & Therapeutics*, *88*(6), 754-756.

Snyder, M. (2016). *Genomics and Personalized Medicine: What Everyone Needs to Know*. Oxford University Press.

Tang-Schomer, M. 2017. *Cost Analysis of Drug Screening for Precision Medicine*. Presentation notes, Department of Pediatrics, Connecticut Children's Medical Center and UConn Health.

Timmerman, L. (2013). What's in a Name? A Lot, When It Comes to 'Precision Medicine'. Xconomy. Available at http://www.xconomy.com/national/2013/02/04/whats-in-a-name-a-lot-when-it-comes-to-precision-medicine/

DECISION SCIENCES INSTITUTE

Analysis and evaluation of e-government in Mexico

Regina Leal
Universidad Autonoma Metropolitana
Email: rlg@xanum.uam.mx
Salvador T. Porras
Universidad Autonoma Metropolitana
Email: stp@xanum.uam.mx

ABSTRACT

The use of Information and communication technologies (IT) has allowed the development of a digital version of public administration through the implementation of electronic government (e-government). This paper aims to analyze the web pages of Mexican government Ministries in order to evaluate the level of performance based on three main characteristics: accessibility, usability and service and contents. The results will allow to propose some improvements in those key elements related to a better performance of Mexican electronic government.

KEYWORDS: Information and communication technologies, Electronic government, Mexico, Accessibility, Usability, Service and contents.

INTRODUCTION

The technological development influences governments to look for new ways of innovation transforming it in order to face new challenges or social environment and new complex situations for public administration. In the same sense, the intensive use of information systems and telecommunication networks has caused society to develop more and new demands on information (Gascó, Rodríguez & Mármol, 2014).

In the last decades the information and communication technologies diffusion changed people's way of life as well as government's performance. As a response, governments have to introduce new policies to improve services through a digital version of public administration, known as electronic administration in Europe, while in Latin America is called electronic government (Ruíz, Morales y Gómez & Contreras, 2014).

The main reason for emerging e-government and the transformation of the relationships among citizens and public institutions is to focus on technological innovation for offering government services (Pérez, et al, 2015). As a result, IT use improve all the papers to make, and allow a quickly, efficient, effective and secure way of access to public information (Aguilera, 2008).

In Mexico, e-government presents different advance levels, citizens can access some online information and services, for example the payment of taxes, downloading the format for driving licenses and passports, to search for a job or to make a medical appointment (Gil-Garcia, Mariscal & Ramirez, 2008). Nevertheless, most of the procedures and services in federal Ministries are not online.

LITERATURE REVIEW

The CINVESTAV (2004) and OECD (2011) define electronic government as the use of Internet and IT to reach a better management of government through transparency and public access to information, reinforces the basic association between public sector and citizens.

In the last years, the development process of electronic government has changed, and it continues changing from evolve non-sequential stages to an overlap and connected building blocks (ONU, 2016).

In this sense, there are several proposals, which establish key requirement elements to develop an e-government. For example, CINVESTAV (2004) points some actions that governments have to implement: working places based on knowledge; change of management; IT infrastructure update which imply an independent but networked, flexible and robust infrastructure; the possibility to do financial transactions, and to provide electronic services, hot spots, sites, and getaways of the state. ONU (2012) defines some evolve stages for e-government implementation: emerge presence, extended presence, interactive presence, transactional presence and total integration. Yu-Che and Gant (2011) include some characteristics for e-government development: incorporate information and services on line, government agencies interoperability, and relationship between governance and citizen for decision-making. Criado and Gil Garcia (2013) suggest some areas to consider in order to create a strategy focus for e-government: e-government national agenda, sites to provide electronic public services, interoperability initiatives, social networks and public administration networks, and open government and transparency.

In México the reason for electronic government is to make easier and to increase the relationship between the State and citizens through IT use, being Internet the key communication tool. Since 1995 the different federal governments have worked in the development of a technological infrastructure for communicate with people, increasing the offer of better services and procedures on line, as well as designing a legal framework for IT use (Luna, Gil-Garcia & Sandoval, 2015). The Mexican electronic government strategy is reflected in the legal and institutional structure built by several projects (Gil-Garcia, Mariscal & Ramirez, 2008). The central axis to manage electronic government for the last government (2012-2018), was the National Digital Strategy (EDN, 2013).

In particular, according to Welp (2008) who analyzed the electronic government web pages of 18 Latin American Countries the Mexican electronic government was considered as “no-functional”, Mexico was ranked 11^{th} compared to other countries in the region. Regarding the countries’ development electronic government functionality, The Global Information Technology Report 2016. Innovating in a digital economy of the World Economic Forum (2016), presents the results of an evaluation conducted in 139 countries. According to such report, Mexico was ranked 76^{th} scoring 4.0 out of 7 (Islas, 2016).

This paper aims to analyze the web sites of Mexican Federal Ministries that offer IT procedures and services, and to review three main characteristics in order to evaluate if the corrections of them would drive to a general improvement of Mexican electronic government: accessibility level, usability level and contents and services. The research uses a quantitative approach through the analysis and evaluation of 21 web sites of Mexican Federal Ministries, including the President’s office, the Attorney General Office and the Legal Advice Office of the Federal Executive (See Appendix 1.).

Based on the previous objective the research question was established as: ¿The improvements of Ministries' web site in terms of accessibility level, usability level, contents and services would increase the low level of Mexican electronic government functionality?

METHODS

A content and qualitative analysis of the Ministries' web sites was conducted. We decided to study the federal Ministries' web sites because they are the governmental main offices in charge of the national public administration. Therefore they represent the first channel for citizen to access procedures and services provided by government. Currently, the Mexican public administration comprises 18 Ministries, the President's Office, the Attorney General Office (PGJ) and the Legal Advice of the Federal Executive (CJEF). A questionnaire was used to analysis each Ministries' web sites. It was designed including 50 questions organized by three subjects: accessibility level, usability level and contents and services.

The following research variables were defined: three independent variables were created: accessibility level, usability level and contents and services. The dependent variable was defined as the level of government site performance; it is the result of the sum of all the independent variables. Based on the dependence variable and establishing a percentage range, the performance level of web sites was defined.

The first independent variable, accessibility level, refers to a set of technologies, rules and design for making easier the use of web sites. The technical basis for developing an accessible web site is to allow the access for information without limits for any reason deficiencies, disabilities, or technology used (CTIC, 2017; DISCAPNET, 2017). For example: hearing problems, sight, movements; problems for reading or cognitive comprehension; inability to use the keyword or the mouse, a small screen, or a slow connection.

We used the rules established by DISCAPNET (2017) on measurable parameters: noticeable, feasible, understandable and robust for accessibility level. The previous characteristics are technical features often included in the web sites that are not easily observed. In order to determine the presence of the previous characteristics we reviewed each one of them at the web sites under study by using two free web analyzers. The first one is part of the TAW Project (2018) and the second one is WAVE (2018), which was developed and offer as a free community service for WevAIM. Those two webs analyzers automatically review the technical features associated with the accessibility level and produce a report of the web site indicating the presence or the absence of diverse accessible characteristics. We considered two values: 1 and 0 to indicate the presence or absence of each characteristic. The total amount of all items produces the score for this first variable; the 30% of the total weighting was assigned.

The second independent variable, level of usability, is utilized to determine the IT products easiness of use. According to Ortiz (2012), this word is used very often among software and hardware products to explain if the use of this technology makes a better and easier experience for user's satisfaction. Following Guia Digital (2017), usability includes several characteristics including corporation identity, web site utility, browsing, visibility of system's status, integrity and compliance with the standards, attention to errors, esthetic and design, help with errors and feedback. This variable was measured using the TAW analyzer, as well as a manual revision of WAVE analyzer results, on the presence or absence of items related to this variable. The characteristics of the level of usability had to values, 1 (presence of the characteristic) and 0 (absence of the characteristic). The total amount of those items produces the score for this second variable; the 30% of weighting was assigned.

The last variable, contents and services, refers to the main activities and procedures provided by the institution to the citizens. It includes for example a guide of procedures, which provides key information for people needing to attend to the Ministry, or information about activities on line (Guia Digital, 2017). We assessed the web sites contents and services that include (CTIC, 2017; DISCAPNET, 2017; Guía Digital, 2017): functional activities to make in the web site; list of applications for mobile devices; links to social networks (YouTube, Facebook, Twitter). Finally to evaluate this variable, we reviewed the presence or absence of the characteristics included in the questionnaire, values of 1 and 0 were used. We also tested the functioning of the government procedures and services offered by the web site. The total amount of general characteristics that produces the score was obtained as well. The weighting of this variable was assigned to be 40% of the total.

Data collected from the web sites was coded and processed using JMP software. In every web site the presence or absence of the previous variables were counted assigning 100% if all the items of each independent variables were included; the percentage of weighting for each variable was assigned according with their characteristics to obtain a final score. This total amount represents the value of the dependent variable as the level of government site performance. The level of government site performance was classified in three categories: non-functional (0-35%), average functional (36%-70%) and functional (71%-100%).

RESULTS

Once the web sites of the Ministries was analyzed, the evaluation of the characteristics of accessibility, usability and contents and services was conducted for obtaining the value of the dependent variable and to establish the level of function of e-government (See Table 1).

Table 1. Federal Ministries Web Sites' level for e-government performance

Ministry	Accessibility level	Usability level	Contents and Services	Total	E-Government level*
President's Office	17.1%	18.9%	21.5%	57.6%	AF
Government Ministry (SEGOB)	16.3%	18.9%	33.8%	69.0%	AF
Ministry of Foreign Affairs (SRE)	18.0%	20.0%	30.8%	68.8%	AF
Ministry of National Defense (SEDENA)	18.0%	20.0%	33.8%	71.8%	F
Admiralty (SEMAR)	18.0%	20.0%	30.8%	68.8%	AF
Treasury (SHCP)	18.0%	18.9%	30.8%	67.7%	AF
Ministry of Social Development (SEDESOL)	18.0%	18.9%	30.8%	67.7%	AF
Ministry of Environment and Natural Resources (SEMARNAT)	18.0%	20.0%	30.8%	68.8%	AF
Ministry of Energy (SENER)	18.0%	18.9%	30.8%	67.7%	AF
Ministry of Economy (SE)	18.0%	20.0%	30.8%	67.9%	AF
Ministry of Public	18.0%	20.0%	33.8%	71.0%	F

Education (SEP)					
Ministry of Agriculture, Livestock, Rural Development, Fishing and Food (SAGARPA)	18.0%	20.0%	33.8%	71.0%	F
Ministry of Communications and Transport (SCT)	18.0%	18.9%	33.8%	69.9%	AF
Ministry of the Public Service (SFP)	18.0%	18.9%	30.8%	66.8%	AF
Ministry of Health (SALUD)	18.0%	18.9%	36.9%	73.0%	F
Ministry of Labor and Social Security (STPS)	18.0%	20.0%	33.8%	71.0%	F
Ministry of Agrarian, Territorial, and Urban Development (SEDATU)	18.0%	18.9%	33.8%	69.9%	AF
Ministry of Tourism (SECTUR)	18.0%	18.9%	30.8%	66.8%	AF
Ministry of Culture (CULTURA)	18.0%	20.0%	36.9%	74.1%	F
Attorney General Office (PGJ)	18.0%	18.9%	30.8%	66.8%	AF
Legal Advice of the Federal Executive (CJEF)	18.0%	18.9%	27.7%	63.7%	AF
*E-Government level: NF: non-functional; AF: average functionality; F: functional.					

Basically, all web sites present the same result for the first independent variable (18%), with the exception of two: the President's Office (17%) and SEGOB (16.3%). The principal reason explaining this result is because web sites are located in the same digital platform following the same web accessibility presenting only few differences among them. The President´s office (2018) and SEGOB (2018) web sites present problems in the links to some documents and downloads, several paths are not accessible to find information or links (PRESIDENCIA 2018; SEGOB, 2018).

The following common characteristics included in web sites were found: consistent browsing, headers everywhere, the possibility to use keyword for browsing, multiple paths to find information, Spanish language as standard, color used for highlights elements, text resizing and images resizing. However, sites do not have hearing description, neither the characteristic of robustness related to compatibility of software and assist devices. As a result, web sites allow a proper interaction with the content and the user can work with them, however, people having special requirements different from sight have problems to work with the platform.

In the same way of the first variable, the result of the level of usability has a similar perform in all Ministries, with the exception of one item, the area for frequent questions. This feature is totally different in all web sites, some of them include a special area for questions, other web sites have a division according to a set of topics, and another does not have any section for questions. The characteristics included in all web sites are: corporative identity, visual hierarchy by sections, procedure catalogs and special information about them, institutions activities, an available searcher, and the Ministries contact details. Although the web site sends an error message, the web site does not send instructions how to resolve or avoid it. There is not a page or section related to the page structure of page functioning; there is not either information about the author or contact of the web site. When the page has an error, it is not possible to go back to the previous page. Even though web sites are useful providing information, there is not attention at errors; this is a key element for digital platforms when users are interacting with them and try to solve some problems. There is not option neither to provide feedback in those characteristics that need to be improved by the people in charge of the web sites.

Results for the last independent variable present a variety of differences among the web sites. Each Ministry has specific procedures and activities, and each of them has a particular offer in their web site. Based on the National Digital Strategy (2013) Ministries have to offer educational activities using IT, such as, courses, training, and on line education. There are also several applications for mobile devices in each institution, but only 7 Ministries offer one application in their web site.

The results on the assessment of the e-government level for the Mexican Ministries indicate that 15 web sites present average functionality scoring from 57.6% to 69.9%, while the remaining 7 Ministries' web sites scores from 71% to 74.1%, meaning web sites are functional.

DISCUSSION AND CONCLUSIONS

According to the research results, the Mexican e-government is using IT to establish relationships with different internal and external agents, generating a change in the way information is given and governmental services are provide, as well as how the interaction between citizens and government institutions is taking place.

However, a characteristic mentioned by various authors (Yu-Che & Gant, 2011; Criado & Gil-Garcia 2013; ONU 2012) relates to how the procedure to develop an electronic government is not always consecutive; the procedure is described in stages, phases or strategies. In that sense, Yu-Che and Gant (2011) posit three stages of electronic government, two of them were confirmed in this research: to add information and services on line, a stage that we found in the contents of all Ministries' web sites. And the second, government agencies' interoperability (Yu-Che & Gant, 2011), as the Federal government decided to centralize its offer in one electronic address for all Ministries: www.Gob.mx, which allows citizens to have access to all services and procedures offered by the Federal government. The third element mentioned by Yu-Che and Gant (2011) refers to governance and citizen for decision-making, situation that was not assessed in this research but probably in a future some improvements in Mexican e-government could be included.

According to strategies for e-government development mentioned by CINVESTAV (2011), the web site Gob.mx follows the strategy related to IT infrastructure updating, which imply the independence on IT use, interconnection and converge. In this sense each Ministry has its own web site to share information and multimedia content, but there is also a specific space named "Unique-window" for procedures and services. This single-window is the result of all procedures and services reviewed to date having a standard digital version; the idea is to continue adding more procedures and services in the following years to save time and to improve the numbers on web sites users'. However the strategy linked to having a flexible and robust IT infrastructure (CINVESTAV, 2011), should to be improved, because based on the research results there are some features on accessibility and usability that need to be developed in order to make all Ministries' web sites functional. Additionally, another important characteristic that Mexican e-government does not include but should be incorporated is the possibility to make financial transactions (CINVESTAV, 2011). The problem was the need of a legal framework allowing the equivalence between a digital with a handwritten signatures. The only web site that accepts this issue is the Tributary Administration Service (SAT for its acronym in Spanish), a decentralized office of the Treasury which main objective is to apply the tax and customs legislation. The SAT's web site is one of the first developed within e-government having advanced and secures procedures and services compared to the rest of the Ministries' web sites. Nevertheless it is currently included as part of the web site Gob.mx, the site was not included in this research because it is an independent platform.

Based on the evolution framework proposed by UN, e-government initiative is classified by its level of maturity considering five evolutionary levels (Gastón & Naser, 2011). Research results let us confirm that Mexican e-government is in the transition process between interactive presence and transactional presence (third and fourth stages each in order). This process shows the presence of several Ministries in the web that offer interactive services and information, as well as complete and secure transactions. While in the majority of web sites there are few digital procedures, it is possible to send and receive official documents such as detail information, registration receipt, pre-registration for a program or service through the web site and by electronic mail; there are also complete and secure on line procedures that include payment of taxes for example.

For 2012, Latin-American countries presented similar process about e-government development. Defined by Criado and Gil-Garcia (2013), the development characteristics include national agendas for e-government that in Mexico has been created by the Digital National Strategy (2013). The second characteristic is the development of web site to provide electronic services, the authors also establish that the web site should offer information and digital procedures as well, considering the idea that the utility of a transactional site is online are the procedures, its statistics and its functionality. In the case of Mexican Ministries', web sites include information about the institution, their projects, results and newsletters, but the main objective is not their on line procedures. The third characteristic relates to interoperability, the possibility of sharing information between public departments and projects developed together (Criado & Gil-Garcia, 2013), we found in the web site of SNIIV a project developed by SEDATU and the National Housing Commission. The following characteristic is the use of social networks (Criado & Gil-Garcia, 2013); all the Ministries' web sites have an official account for popular social networks such as Facebook, Twitter, YouTube and Instagram; in some cases the institutions also use them as a diffusion tool. The last characteristic proposed by Criado and Gil-Garcia (2013) is about an open and transparent government, which implies a way to open and to present their accountability through digital media; we found some information in some web sites of Mexican e-government, but this is a task that has to be improved.

Research results related to e-government phases or classification confirm that Mexican electronic government has a superior average level for transactional phase, because of the result of the majority of Ministries (MF) and only 4 out of 22 had a functional level. Therefore the research hypothesis "The level of Mexican electronic performance is low because of the Ministries web sites' development, which has lacks in characteristics such as accessibility level, usability level and contents and services" was rejected.

However, some of the characteristics associated to accessibility level and usability level can be improved in order to increase the general level of Mexican e-government functionality. As it was mentioned before it is important to review the "error an suggestions" section in order to make the web sites more user friendly and easy to use. Another feature to improve is to make the web sites browsing easier, adding the option to return to previews o next pages or jump directly to a specific section without the need to return each time to the original path. The web sites have to include less frequent tools, for example the audio-description and assistance tools adaption that will make the site Gob.mx a real accessible virtual platform and inclusive for a major number of citizens.

Nowadays due to smartphone and tablet's diffusion, it is important the design and implementation of applications for them. The development of these new technological tools represents a key challenge to increase communication, access to information and services and improve the relationship between citizen and government. This is a much needed further step for e-government web sites.

In the same sense, Mexican e-government should work focused on the inclusion of all the characteristics of the transactional phase, to start with the addition of a major number of digital procedures, even though their information could be completed, the idea is to give citizens the option to complete all procedures on line. Increasing on line procedures will increase also the number of web site's visitors and the utility of the platform.

Finally, another suggestion is the inclusion of characteristics for citizens' participation by e-government use. Through this tool, citizens can be involved in decision making processes, become activities and public resources transparent and establish and better channel of communication for an efficient and effective government.

In order to extend the broad and diversity of studies related to Mexican e-government it would be very useful to conduct research in diverse Federal government institutions, as well as in State and local departments.

REFERENCES

Aguilera Izaguirre, G. (2008). 2017: El gobierno electrónico en México, *CienciaUAT*, 3, (Julio-Septiembre), 20-23. Retrieved from: www.redalyc.org:9081/articulo.oa?id=441942913008, November.

CINVESTAV (2004). 2017: Gobierno Electrónico. Instituto Politécnico Nacional IPN, México, D.F. Retrieved from: http://administracion.cinvestav.mx, November.

Criado, I. J. & Gil García, J. R. (2013). 2018: Gobierno electrónico y políticas públicas: Estado actual y tendencias futuras en América Latina, *Gestión y Política Pública*, Centro de Investigación y Docencias Económicas. A.C. México, 3-48. Retrieved from: www.redalyc.org/pdf/133/13331692001_2.pdf, February.

CTIC (2017). Fundación Centro Tecnológico de la Información y la Comunicación. Retrieved from: www.tawdis.net, November.

DISCAPNET (2017). El Portal de las Personas con Discapacidad. Fundación ONCE. Retrieved from: www.discapnet.es, November.

EDN (2013). 2018: Estrategia Digital Nacional. Secretaría de Gobierno. Retrieved from: https://framework-gb.cdn.gob.mx/data/institutos/edn/Estrategia_Digital_Nacional.pdf, January.

Gascó M., Rodríguez, L. & Mármol, H. (2014). 2017: Guía de buenas prácticas en gobierno abierto. Observatorio de Sociedad, Gobierno y Tecnologías de Información, Bogotá:

Universidad Externado de Colombia. Retrieved from: www.gobiernolocal.gov.ar/sites/default/files/guiadebuenaspracticasengobiernoabierto-140923062754-phpapp02%281%29.pdf, November.

Gastón C. & Naser A (2011). 2017: El gobierno electrónico en la gestión pública. *Serie gestión pública,* ILPES, Abril, Santiago de Chile. Retrieved from: www.cepal.org/ilpes/noticias/paginas/3/54303/El_Gobierno_electronico_en_la_gestion_Publica.pdf, November.

Gil-García, J. R., Mariscal, J. & Ramírez, F (2008). 2018: Gobierno electrónico en México. *CIDE,* Noviembre, Núm. 214, México. Retrieved from: www.libreriacide.com/librospdf/DTAP-214.pdf, February.

Guía Digital (2017). Gobierno de Chile. Retrieved from: www.guiadigital.gob.cl/articulo/definicion-de-contenidos-del-sitio, November.

Islas Carmona, O. (2016). 2018: ¿Por qué la Estrategia Digital Nacional terminará en un rotundo fracaso? . *Comhumanitas: Revista Científica de Comunicación*, Núm. 7, Quito, Ecuador, pp. 60-76. Retrieved from: file:///F:/Seminario/Dialnet-PorQueLaEstrategiaDigitalNacionalTerminaraEnUnRotu-5896236.pdf, February.

Luna Reyes, L. F., Gil-García, J. R. & Sandoval Almazán, R. (2015). 2018: Avances y Retos del Gobierno Digital en México*. Instituto de Administración Pública del Estado de México*, Toluca de Lerdo, México. Retrieved from: www.iapem.mx/Libros/Avances-y-Retos-del-Gobierno-Digital-en-Mexico.pdf, January.

OCDE (2011). 2017: Hacia una gestión pública más efectiva y dinámica en México, Estudios de la OCDE sobre gobernanza pública*.* Organización para la Cooperación y el Desarrollo Económicos, OECD Publishing. Retrieved from: www.oecd.org/gov/regulatory-policy/48808023.pdf, November.

ONU (2016). 2018: Gobierno electrónico en apoyo al desarrollo sostenible, Administración y Gestión del Desarrollo. Organización de las Naciones Unidas, Departamento de Asuntos Económicos y Sociales, Nueva York. Retrieved from: http://workspace.unpan.org/sites/Internet/Documents/UNPAN96407.pdf, February.

ONU (2012). 2018: Estudio de las Naciones Unidas sobre el Gobierno Electrónico, 2012, Gobierno Electrónico para el pueblo. Organización de las Naciones Unidas, Departamento de economía y asuntos sociales, Nueva York. Retrieved from: https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2012-Survey/Complete-Survey-Spanish-2012.pdf, January.

Ortíz Arreola, P. (2012). Usabilidad y el beneficio de su aplicación: Auditoria de un sistema informático de educación a distanci". Universidad Autónoma del Estado de México UAEM, Toluca, Estado de México.

Pérez Zúñiga, R., Camacho Castillo, O., Mena Hernández, E. & Arroyo Cervantes G. (2015). 2017: Análisis general del gobierno electrónico en México. *Paakat: Revista de Tecnología y Sociedad* (México, Guadalajara) Universidad de Guadalajara, 9 (septiembre-febrero). Retrieved from: www.redalyc.org/articulo.oa?id=499051500010, November.

PRESIDENCIA (2018). 2017: Retrieved from: https://www.gob.mx/presidencia, November.

Ruíz Alanís, L., Morales y Gómez, J. M. & Contreras Orozco, L. (2014). 2017: Perspectivas del gobierno electrónico local en México. Universidad Autónoma del Estado de México UAEM, Toluca, Estado de México. Retrieved from: http://iapem.mx/Libros/2014%20160%20Perspectiva%20del%20Gob%20Elec%20Loc.pdf, November.

SEGOB (2018). Retrieved from: www.gob.mx/segob, November.

TAW (2018). WCAG 2.0. Retrieved from: www.w3.org/tr/wcag20, January.

Yu-Che Ch. & Gant, J. (2001). 2017: Transforming Local e-government services: the use of application service providers. *Governement Information Quaterly.* 18(4), pp. 343-355. Retrieved from: www.sciencedirect.com/science/article/pii/S0740624X01000909, November.

WAVE (2018). Retrieved from: http://wave.webaim.org, April.

Welp, Y. (2008). 2017: América Latina en la era del gobierno electrónico. Análisis de la introducción de nuevas tecnologías para la mejora de la democracia y el gobierno. *Revista del CLAD Reforma y Democracia,* Núm. 41, Junio, pp. 173-192. Retrieved from: www.redalyc.org/html/3575/357533672007/, November.

APPENDIX 1

Federal Ministries Web Site's addresses	
Ministry	Web Site's addresses
President's Office	www.gob.mx/presidencia
Government Ministry (SEGOB)	www.gob.mx/segob
Ministry of Foreign Affairs (SRE)	www.gob.mx/sre
Ministry of National Defense (SEDENA)	www.gob.mx/sedena
Admiralty (SEMAR)	www.gob.mx/semar
Treasury (SHCP)	www.gob.mx/shcp
Ministry of Social Development (SEDESOL)	www.gob.mx/sedesol
Ministry of Environment and Natural Resources (SEMARNAT)	www.gob.mx/semarnat
Ministry of Energy (SENER)	www.gob.mx/sener
Ministry of Economy (SE)	www.gob.mx/se
Ministry of Public Education (SEP)	www.gob.mx/sep
Ministry of Agriculture, Livestock, Rural Development, Fishing and Food (SAGARPA)	www.gob.mx/sagarpa
Ministry of Communications and Transport (SCT)	www.gob.mx/sct
Ministry of the Public Service (SFP)	www.gob.mx/sfp
Ministry of Health (SALUD)	www.gob.mx/salud
Ministry of Labor and Social Security (STPS)	www.gob.mx/stps
Ministry of Agrarian, Territorial, and Urban Development (SEDATU)	www.gob.mx/sedatu
Ministry of Tourism (SECTUR)	www.gob.mx/sectur
Ministry of Culture (CULTURA)	www.gob.mx/cultura
Attorney General Office (PGJ)	www.gob.mx/pgj
Legal Advice of the Federal Executive (CJEF)	www.gob.mx/cjef

DECISION SCIENCES INSTITUTE

Electric Vehicles: Are they an Alternative to Weaken the Fossil Fuels' Industry Power?

Jesus A. Cardenas
University of Texas at El Paso
Email: jacardenas3@utep.edu

Leopoldo Gemoets
University of Texas at El Paso
Email: lgemoets@utep.edu

ABSTRACT

The increasing cost of gasoline in the world has been pushing auto makers to the alternative of making electric vehicles, some of them hybrids and others fully electrical. This study tries to identify the factors that might be having a major impact on the sales of electric vehicles in the United States, as well as our prognosis for the near future.

KEYWORDS: Regression, electric, alternative fuels, plug-in vehicles

INTRODUCTION

In a research project, we asked some academics about their Smart Grids' impressions and we found out that for most of them the Electric Vehicle is hardly considered as part of the Smart Grid general concept; they see it as a necessary investment when the time is right. Consumers seem to be expecting the prices to drop, and to see how the charging of the vehicle into their homes will have an impact on their monthly electricity bills. Because there are so few vehicles at this time, their impact on the consumption of electricity is not noticeable yet (Heuer et al., 2011). Khattak et al. (2012) consider that the electric vehicle could hold the largest potential within the smart grid to deliver carbon savings, so the implementation process needs to be included into the SGD projects. One major area of research is the future problem of millions of electric vehicles charging batteries at the same time and connected to the general grid, this phenomenon is called V2G (vehicle to grid connection).

An aspect that is being explored more nowadays is the role of the electric vehicle as a consumer of energy, but also as a mobile energy source (Couillet et al., 2012). This flexibility of the electric vehicles reinforces the modern concept of "prosumers," wherein the consumer also becomes a producer of energy and sells the excess to the utility company for a just price. In the past years, there has been some confusion and fear about electric vehicles catching fire after an accident, even hours after the incident, so the National Fire Protection Association (NFPA) is issuing warnings and recommendations for vehicle owners (Lekach, 2019).

IMPLEMENTATION PROGRESS

It is important to differentiate between three types of vehicles that are grouped under this category: Hybrid Vehicles Plug-In Hybrid Electric Vehicle (PHEV) and Plug-In Electric Vehicle (PEV). PHEV is a vehicle that can operate with either electricity or an alternative energy source. PEV uses only electricity to operate and is plugged-in to recharge the battery in order to be prepared for further use. PEVs sales in the US began in 2010, while hybrids have been in the market since 1999.

There are several factors that have been influencing the erratic behavior of the PHEV growth, as there was a frankly increasing period from 1999 to 2007, but then the trend reversed to decrease until 2011—when it reaches bottom and begin growing again. Form 2011 grew to 2013 and the declined back, but the gains have been larger than the losses. We are going to analyze some of the factors that may be the major contributors for this behavior.

Hypotheses

Based on our consumers' experience, we expect that electric vehicle growth is strongly influenced by the price of gas or oil in general, government incentives, the cost of the vehicle itself, availability of charging stations, wattage of the battery, and miles to run under a normal battery charge.

To evaluate these factors, we develop the following hypotheses:

H1. EV sales are influenced by the price of gasoline

H2. EV sales are influenced by the governmental incentives to consumers

H3. EV sales are influenced by the cost of the vehicle

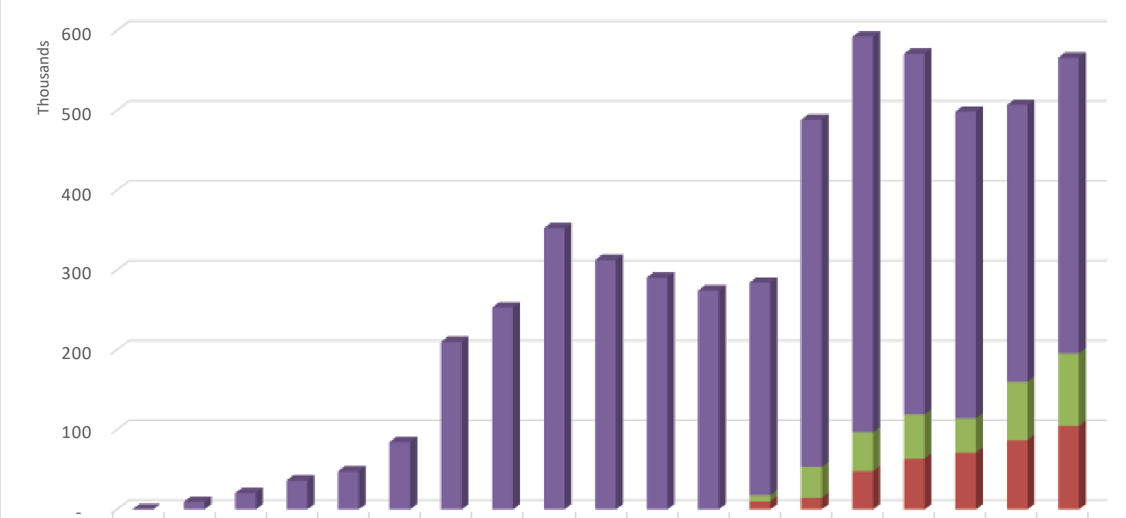
H4. EV sales are influenced by the availability of charging stations in the area

H5. EV sales are influenced by the wattage of the battery in the vehicle

H6. PHEV sales are influenced by the number of miles run in a single battery charge

Deployment of Electric Vehicles in the US

The number of vehicles sold in the US is shown in Figure 1. The chart shows hybrids, PEV and PHEV together, although we are going to focus on PEV and PHEV for the analysis, as there is a longer period of time available to come up with conclusions. Figure 2 shows the average price of gasoline in the US. Although there are some statistics per region, we are using the national average to compare with national sales of Electric and hybrid vehicles.



	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Hyb	17	9,350	20,28	36,03	47,60	84,19	209,7	252,6	352,2	312,3	290,2	274,2	266,3	434,8	495,5	452,1	384,4	346,9	370,6
PHEV													7,671	38,58	49,00	55,35	42,87	72,83	90,77
PEV													10,06	14,65	47,69	63,41	71,04	86,73	104,4

Figure 1: Electric vehicles produced in the U.S.

Table 1: Regression results for gasoline price vs PHEV sales

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.765[a]	.585	.561	138717.99180

a. Predictors: (Constant), Gas
b. Dependent Variable: EV

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
	(Constant)	-198619.204	103518.578		-1.919	.072
	Gas	200168.748	40859.895	.765	4.899	.000

ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	461810315771.318	1	461810315771.318	23.999	.00
Residual	327125581251.103	17	19242681250.065		
Total	788935897022.421	18			

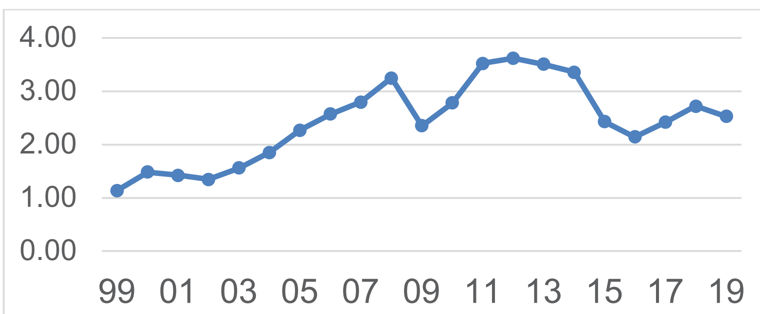


Figure 2: Average Gasoline Prices (Source: eia.gov)

Conducting a linear regression analysis for number of hybrid and electric vehicles sold in the US, and the average price of gasoline in the year, we find that a R2 of .585 and the significance is under 1%. With these results we can support hypothesis H1 regarding the cost of gasoline correlating to the number of hybrid and electric units sold in the US.

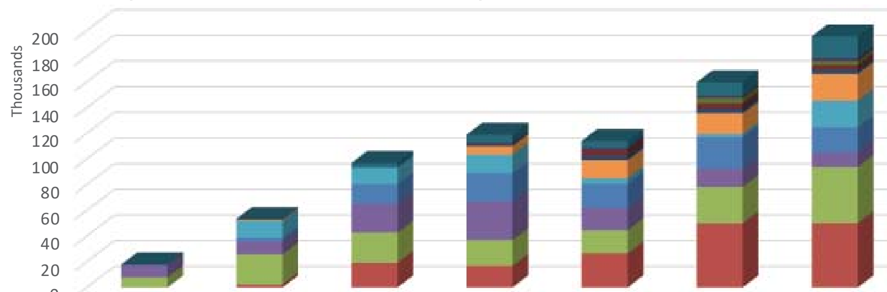
The government has been giving tax credits to consumers who buy PHEVs in order to promote their diffusion. Checking the analysis done by David Diamond (2009), he discovers a weak relationship between government incentives and PHEVs adoptions. For our research, we looked for federal tax credits per manufacturer and calculated the average for all vehicles, then, based on the number of sold units, we tried to find non-linear regressions, but the best one was the exponential distribution with an R^2 of 0.356 and significance of 12%. Based on these results, we do not support the hypotheses H2, as we agree with Diamond that the federal government incentives do not have a relationship with the number of PHEVs sold. Some companies do not for federal tax credits per manufacturer and calculated the average for all vehicles, then, based on the number of sold units, we tried to find non-linear regressions, but the best one was the

Table 2: Average tax credits vs. PHEVs sold in the US

Vehicles	Cost	Rebates	Units sold
Audi A3 Plug In	$ 39,500	$ 6,802	7,157
BMW 3 Series Plug In	$ 45,600	$ 4,801	5,046
BMW 5 Series Plug In	$ 53,400	$ 6,968	3,759
BMW 7 Series Plug In	$ 91,250	$ 5,378	730
BMW Active E	$ 21,726	$ 7,500	965
BMW i3	$ 44,450	$ 10,800	30,997
BMW i8	$ 147,500	$ 6,469	4,902
BMW X5	$ 63,750	$ 7,468	12,118
Cadillac CT6	$ 75,095	$ 8,300	205
Cadillac ELR	$ 74,045	$ 8,300	2,891
Chevrolet Bolt	$ 36,620	$ 7,050	23,876
Chevrolet Spark	$ 17,720	$ 13,220	7,392
Chevrolet Volt	$ 33,530	$ 6,060	133,512
Chrysler Pacifica	$ 39,995	$ 9,800	2,764
Fiat 500E	$ 32,995	$ 10,800	12,313
Ford C-MAX Energi	$ 23,988	$ 4,007	41,649
Ford Focus EV	$ 29,120	$ 10,800	8,685
Ford Fusion Energi	$ 34,595	$ 6,885	52,959
Honda Accord	$ 39,780	$ 8,300	1,038
Honda Clarity BEV	$ 36,620	$ 8,300	1,111
Honda Clarity Plug-in	$ 33,400	$ 9,800	898
Honda Fit EV	$ 36,525	$ 9,800	1,071
Hyundai Ioniq EV	$ 30,315	$ 10,800	427
Hyundai Sonata Plug-in	$ 33,400	$ 7,302	5,254
Kia Optima Plug-in	$ 35,210	$ 4,024	1,542
Kia Soul EV	$ 33,950	$ 10,800	5,150
Mercedes B-Class Electric	$ 41,350	$ 4,134	4,056
Mercedes C350We Plug-in Hybrid	$ 47,900	$ 4,134	981
Mercedes GLE 550e Hybrid	$ 66,700	$ 5,385	642
Mercedes S550 Plug In	$ 93,135	$ 5,385	1,251
Mini Cooper Countryman Plug In	$ 36,900	$ 4,968	475
Mitsubishi i-MiEV	$ 20,160	$ 7,500	2,104
Nissan LEAF	$ 29,990	$ 14,300	114,808
Porsche Cayenne S E-Hybrid	$ 79,900	$ 7,512	4,960
Porsche Panamera S E-Hybrid	$ 99,600	$ 7,470	1,748
Smart ED	$ 28,100	$ 6,750	312
Smart for Two EV	$ 23,900	$ 10,800	6,241
Tesla Model 3	$ 48,500	$ 7,050	1,770
Tesla Model S	$ 85,000	$ 7,050	120,821
Tesla Model X	$ 88,000	$ 7,050	41,508
Toyota Prius Prime	$ 27,350	$ 6,802	65,702
Toyota RAV4 EV	$ 49,800	$ 17,500	2,399
Volvo S90 Plug In	$ 63,650	$ 7,469	112
Volvo XC60	$ 52,900	$ 7,469	518
Volvo XC90	$ 64,950	$ 7,469	4,297
VW e-Golf	$ 30,495	$ 10,800	12,060

exponential distribution with an R^2 of 0.356 and significance of 12%. Based on these results, we do not support the hypotheses H2, as we agree with Diamond that the federal government incentives do not have a relationship with the number of PHEVs sold. Some companies do not show tax credit incentives as shown on Table 2, so they were not used for the comparison. The price of the PHEVs varies based on the vehicle's features, so companies like General Motors are pushing low price PHEVs through Saturn, and big trucks such as the Silverado, Yukon and Tahoe with higher prices. Porsche is selling sport luxury cars that can be as costly as the Panamera with a manufacturer suggested retail price (MSRP) of almost 100,000 dollars.

In 2007, PHEV's prices dropped in order to overcome the fall of sales, but the average price still went down for the first time, and in 2011 the number of sold vehicles was reduced, while the average cost of sold PHEVs went up considerably. Figure 3 shows the breakdown of PHEVs by manufacturer and year. In Figure 4 we can also see how, following 2011, the average prices balanced down with a positive more realistic trend upward.



	2011	2012	2013	2014	2015	2016	2017
Others	386	820	3104	6197	5262	10532	16817
Mercedes				774	2024	1565	2567
Audi						4280	2877
Volkswagen				357	4232	3937	3534
Fiat			260	1503	3477	3737	3336
BMW		965		6647	14097	16088	20720
Toyota		12941	13093	14448	4209	2474	20936
Ford		3057	14981	21947	18923	24796	19589
Nissan	9674	9819	22610	30200	17269	14006	11230
Chevrolet	7671	23461	23654	19950	18022	28353	43669
Tesla		2171	19000	16750	26408	49800	49970

Figure 3: PEV and PHEV vehicles breakdown by brand

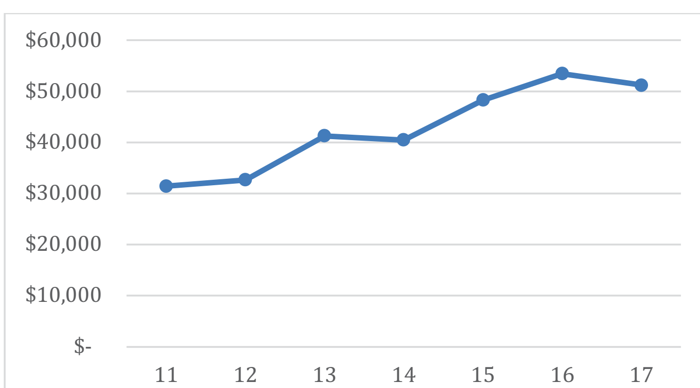


Figure 4: Number of PHEVs sold versus average price

The number of hybrid vehicles sold in the US in the period from 1999 through 2017 presents a positive trend, as compared to the average price of all sold PHEVs. Because the amount of expensive hybrid vehicles is small, the averages do not grow that much on a yearly basis, although there are very expensive PEVs that are sold in smaller amounts. Doing a linear regression analysis between these two variables we are finding a favorable relationship.

With the R^2 of 0.847, we are seeing a significant relationship between the average price of the PHEV and the sales. If the demand and offer of PHEVs in the market react as expected, the higher the prices would imply the lower the sales, but as in the gasoline vehicles the prices mostly increase on a yearly basis, seeing increases on PHEVs may have the same impact. Based on the above shown results, we can support hypothesis H3, stating that the price of PHEV's has an impact on their sales.

Table 3: Regression results for PHEV average price vs yearly sales

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.920[a]	.847	.816	25792.82974

a. Predictors: (Constant), Avg_Price
b. Dependent Variable: PHEV_Sales

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18371092701.851	1	18371092701.851	27.614	.003[b]
	Residual	3326350329.578	5	665270065.916		
	Total	21697443031.429	6			

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-164116.447	52670.550		-3.116	.026
	Avg_Price	6.373	1.213	.920	5.255	.003

To discover the relationship between available PHEVs to the number of charging stations, we have to research the number of outlets and stations as well as the number of registered vehicles per state—in order to determine if one has an impact on the other. The more vehicles, the more charging stations, but if the number of charging stations is limited, sales of EVs may be affected negatively.

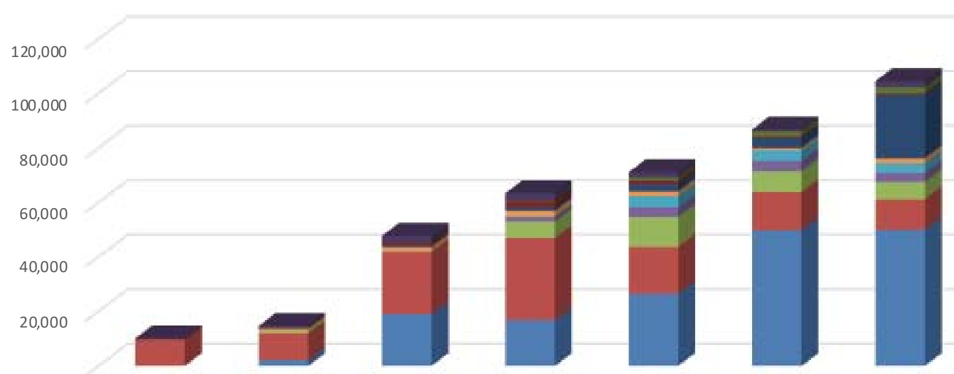
Two noteworthy states from Table 3 are California, which has 19,687 outlets for electrical charging stations for over 500k of electric vehicles. That is, there are 25.73 vehicles per charging station, which is the highest in the United States. On the other side of the spectrum, Missouri es the state with less electric vehicles per outlet or charging station with 3.88, which is telling us that there are almost 4 vehicles per outlet.

The amount of stations and the number of registered vehicles make sense in most cases, so doing a regression analysis we find a linear relation with a R2 of 0.992 and significance below 1%. With these results we can support hypothesis H4, that there is a regression between the number of loading stations and the number of EVs. As the number of EVs increases, the number of charging stations will have to grow accordingly otherwise, it can become a reason for the PHEV and PEV deployment.

Table 3: Registered PHEVs vs. charging stations per state

St	Outlets	EV Stock	EV Stock/ Outlet
CA	19,687	506,608	25.73
TX	3,109	34,239	11.01
FL	3,010	40,548	13.47
NY	2,828	46,397	16.41
WA	2,383	41,459	17.40
GA	2,335	33,947	14.54
CO	1,857	19,738	10.63
MA	1,758	22,824	12.98
MS	1,720	6,676	3.88
MD	1,598	17,900	11.20
OR	1,462	21,433	14.66
VA	1,356	16,505	12.17
NC	1,331	13,054	9.81
IL	1,255	22,475	17.91
AR	1,223	18,129	14.82
MI	1,112	18,434	16.58
TN	1,054	6,684	6.34

To analyze the battery size and wattage before relating these elements to sales, we need to go by model and company. We can see in Figure 4 the distribution of PEVs separated by their manufacturing company. This breakdown is compared to the average battery size by company in order to test the relationship among sales and battery size.



	2011	2012	2013	2014	2015	2016	2017
Others	76	873	2,603	2,561	2,041	726	2,288
Kia				250	1,015	1,728	2,157
Smart	310	139	923	2,594	1,387	657	543
Chevrolet			560	1,145	2,629	3,614	23,320
Ford		683	1,738	1,964	1,582	901	1,817
Volkswagen				357	4,232	3,937	3,534
Fiat			260	1,503	3,477	3,737	3,336
BMW		965		6,092	11,004	7,625	6,276
Nissan	9,674	9,819	22,610	30,200	17,269	14,006	11,230
Tesla		2,171	19,000	16,750	26,408	49,800	49,970

Figure 4: Plug-in Electric vehicles breakdown in the U.S.

Table 4: PEV Vehicles Sales in 2017, Battery Size and Miles Run in One Charge

PE Vehicle	Sales in 2017	Battery Size KWh	Distance in one charge (Miles)
Tesla 3	1,770	75.0	310
Tesla S 60	26,500	60.0	170
BMW i3 (2019)	4,141	42.0	115
Nissan Leaf	11,230	30.0	100
Fiat 500e	3,336	24.0	85
Mercedes B	744	28.0	85
GM Spark	23	21.0	75
Ford Focus	1,817	23.0	75
Smart for Two	543	17.6	68
Mitsubishi MiEV	6	16.0	55
Chevrolet Volt	20,349	16.0	40
Cadillac ELR	17	16.5	35
Ford C-Max	8,140	7.6	21
Ford Fusion	9,632	7.6	21
Porsche Panamera SE	1,574	9.4	20
Toyota Prius Plug In	20,936	4.4	15

Based on the statistical analysis of the regressions conducted in this project, we are finding half of the hypothesis that can be supported, as they nicely correlate the variable of sales with the cost of gas, the cost of the vehicles and the availability of charging stations.

Table 5: Hypotheses Results

Hypotheses	Results	Comments
H1. EV sales are influenced by the price of gasoline	Supported	R^2 of 0.585 with significance under 1% for the linear regression
H2. EV sales are influenced by the governmental incentives to consumers	Not Supported	R^2 of 0.356 with significance above 12% for the exponential regression
H3. EV sales are influenced by the cost of the vehicle	Supported	R^2 of 0.847 with significance under 1% for the linear regression
H4. EV sales are influenced by the availability of charging stations in the area	Supported	R^2 of 0.920 with significance under 3% for the linear regression
H5. EV sales are influenced by the wattage of the battery in the vehicle	Not supported	R^2 of 0.006 for a non-significant linear regression
H6. PHEV sales are influenced by the number of miles run in a single battery charge	Not Supported	R^2 of 0.001 for a non-significant linear regression

CONCLUSIONS

There is an important diffusion of electric vehicles in the US, as the manufacturers and the state and federal governments are addressing some of the major concerns. One of the major concerns about electric vehicles has been the battery durability and reliability. This issue has been addressed as more plug-in electric cars are being sold now with larger batteries, but there is still the concern about the cost of the replacement battery and what is going to be done with the old one. As new and more efficient batteries are being developed, the cost of the PEVs is going down and new lower price versions are appearing.

We are finding that with the ever-increasing cost of gas, the trends of hybrid and plug-in vehicles are going up, but until the cost of gas becomes prohibitive, the larger proportion of vehicles is not going to be electric. Government incentives are not having the expected major effect, although some publications mention that consumers have been taking all the credits and incentives from the governmental programs. A proposed action for the government is to increase taxes on gas, and that will definitively have a major impact on the diffusion of electric cars.

More charging stations are being installed throughout the country, with a growth of 30% since 2013 (DoE, 2014). There are expectations that the growth will become exponential to satisfy the large demand of electric vehicles; but the battery sizes and miles run with a single battery charge do not seem to have a big impact on sales, probably because batteries are still evolving.

REFERENCES

Battery University Group (2019). BU-1003: Electric Vehicle (EV). *Batteryuniversity.com.* April 18, 2019 https://batteryuniversity.com/learn/article/electric_vehicle_ev

Couillet, R., Perlaza, S. M., Tembine, H., Debbah, M. (2012). Electrical Vehicles in the Smart Grid: A Mean Field Game Analysis. *IEEE Journal on Selected Areas in Communications*, 30(6), 1086–1096. doi:10.1109/JSAC.2012.120707

Energy, US Department of (2014) Alternative Fueling Station Counts by State. *Today in Energy*. December 10, 2014 https://www.eia.gov/todayinenergy/detail.php?id=19131

Energy, US Department of (2019) Petroleum and Other Liquids. *Today in Energy*. May 20, 2019 https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPMR_PTE_NUS_DPG&f=W

Energy, US Department of (2019) Alternative Fueling Station Locator. *Alternative Fuels Data Center.* May 20, 2019 AFDC Alterntative Fueling Station Locator Data

EVADoption (2019) Charging Stations By State. *EVAdoption.com*. April 5, 2019. https://evadoption.com/ev-charging-stations-statistics/charging-stations-by-state/

EVADoption (2018) US EV Sales Percentages of Total Vehicle Sales By Brand. *EVAdoption.com*. 2018. https://evadoption.com/ev-sales/evs-percent-of-vehicle-sales-by-brand/

Heuer, J., Komarnicki, P., Styczynski, Z. A. (2011). Integration of Electrical Vehicles into the Smart Grid in the Harz.EE-mobility Research Project. *IEEE Power & Energy Society General Meeting* (pp. 1–6). Ieee. doi:10.1109/PES.2011.6039147

Khattak, A. R., Mahmud, S. A., Khan, G. M. (2012). The Power to Deliver. *IEEE Power and Energy Magazine,* (august), 56–64.

Lekach, S. (2019). Tesla Electric Vehicle Car Crash Fire Reignite. *Mashable.com*, Feb. 25, 2019 https://mashable.com/article/tesla-electric-vehicle-car-crash-fire-reignite/

Loveday, S. (2018) “Monthly Plug-in EV Sales Scorecard: Historical Charts”. *Insideevs.com.* December 3, 2018. https://insideevs.com/news/344007/monthly-plug-in-ev-sales-scorecard-historical-charts/

DECISION SCIENCES INSTITUTE

Teaching Tips: Cultivating Essential Soft Skills in an Undergraduate Business Analytics

Benyawarath Nithithanatchinnapat
University of North Carolina at Charlotte
Email: bnithith@uncc.edu

ABSTRACT

The presentation illustrates how soft skills can be developed in an undergraduate business analytics course. An activity design guideline and one activity were developed. The activity was used once with satisfactory results, and the next step is to develop a complete set of exercises with assessment tools.

KEYWORDS: Business Analytics, Soft Skills, Data Exploring

INTRODUCTION

Employers are looking for new college graduates who possess essential soft skills (Chamorro-Premuzic & Frankiewicz, 2019; Dishman, 2016; Dutton, 2012). However, based on the review of the undergraduate business analytics course outlines and syllabus that are available online, the activity or topic that help students cultivate essential soft skills seem to be missing. Of course, the focus of an undergraduate business analytics course should be quantitative and qualitative methods, but making sure that students can also develop important soft skills employers look for help students become more competitive in the job market (Chamorro-Premuzic & Frankiewicz, 2019). Soft skills that students in the analytics field should possess are, for instance, critical thinking, problem-solving, curiosity, opportunity finding, and storytelling (McPherson, 2016).

Measuring, assessing, and cultivating soft skills in undergraduate students and new hires are essential in the academic realm as well. In the scholarly literature, there are two streams of research about soft skills, one is skills assessment, and another is about cultivating them. To the analytics field, there is still a need for more teaching tips, cases, and approaches in soft skills development since this field is new. Thus, this study aims to offer a guideline in designing activities that business analytics instructors can use to help students develop their essential soft skills. This study is built on a framework presented in Dubey and Gunasekaran (2015) and in Blaszczynski and Green (2012) and Green and Blaszczynski (2012).

In addition to the soft skill issue, the author also found that the ability to make sense of a large volume of data is critical in today's digital age (Kerner, 2019; Mela & Moorman, 2018). So, to fill the gap and to help new graduates become more competitive in the job market, the author developed an activity for an undergraduate Introduction to Business Analytics course in a public AACSB accredited institution. Students must learn critical thinking, opportunity finding, and storytelling in a real-world context. Using this activity, the author shows how students can analyze historical unintegrated data, come up with interesting questions and storytelling plan, and more importantly, students had fun doing the exercises and saw the importance of developing essential soft skills. Most students said that this exercise was much more challenging than learning how to use software and how to write basic SQL. However, a soft skill development activity or a teaching case that is ready to be used in an undergraduate business analytics course is still missing, to the best of the author's knowledge.

This study is to be continued. In the next phase, the author will develop an assessment method (Beard, Schwieger, & Surendran, 2008) such as having peer-evaluation as well as instructor-evaluation and having pre- and post-exercise questionnaires, as well as a complete set of exercises for this course. The rest of the paper is organized as follows. The next section is a review of the literature. Following that is the design and the detail of the exercise and ending with the instructors' guideline and conclusion.

LITERATURE REVIEW

Cultivating soft skills in current employees and students are of interest to both scholars (Beard et al., 2008; Ellis, Kisling, & Hackworth, 2014; Gibb, 2014; Hagmann, Almekinders, Bukenya, Guevara, & Halemichael, 2003; Henry & Venkatraman, 2015; Liberatore & Luo, 2013; Petter & Randolph, 2009; Wats & Wats, 2009) and practitioners (Briggs, 2015; White, 2014). Soft skills are essential in most, if not all, professions, such as in economics (Suciu & Lacatus, 2014), software development (Ahmed, Capretz, Bouktif, & Campbell, 2015), and dental education (Gonzalez, Kasim, & Naimie, 2013).

In academic literature, one stream of "soft skills" studies focuses on the assessment part (Abbas, Abdul Kadir, & Ghani Azmie, 2013; Gibb, 2014), another on cultivating them (Blaszczynski & Green, 2012; Hagmann et al., 2003; Wats & Wats, 2009), and some studies offer both (Beard et al., 2008). After the review of the literature about "soft skills" development and assessment in scholarly business and management journals, the gap is found, which is the lack of soft skills development activities that are suited for analytics classes, such as undergraduate business analytics. This study is built on three previous work, one is Dubey and Gunasekaran (2015), which offered an education and training framework for a successful career in big data and business analytics. The data was collected from ten heads of business analytics for companies situated in Mumbai and the Pune region of Maharashtra, India. Since India has many major IT hubs, the author believes that the finding applies to other countries as well. Another two are Blaszczynski and Green (2012) and Green and Blaszczynski (2012), which offered effective strategies and activities for developing soft skills. They offered soft skills-building activities, such as pyramid building activity, a role-play, a reflection journal, a listening activity, a "What's on a Penny?" activity, a line-dancing activity, and a story-in-the-bag activity. Adding on to this list is the new data exploring activity, which directly relate to data analytics and is easy to implement.

GUIDELINE IN DESIGNING ACTIVITIES

Table 1 shows a list of soft skills from Dubey and Gunasekaran (2015) and McPherson (2016). To get a holistic picture, the author used both lists and came up with plausible things that the instructors can do and have students do in a business analytics class.

Table 1: Soft Skills and Activities for Analytics Class		
Soft Skills Important for Analytics Graduates	**Activity Designing Plausible Things to Do**	
	Have students do	**What the Instructors can do to facilitate**
Ability to learn and Curiosity	Learn new tools that have not been covered in detail in class	Make sure the instructions are not detailed step-by-step, and instead generate curiosity
Ability to focus on key issues	Solve a given business problem	Sit with students, one group at a time, and ask them how they are doing and whether they are clear on what problem they are solving
Communication, Listening Skills, and Storytelling	Communicate with other group members Briefly summarize the plan to the instructor Prepare a written submission Make sure the selected business problems and design of the presentation is coherent	Listen to students' plan Give feedback to students Guide students to the understanding of the importance of storytelling with data
Critical Thinking	Propose rational recommendation based on scientific data	Ask students back what they think, if they ask whether their business questions sound good
Leadership ability	Determine the project manager	Ask students how they are doing, whether there is an informal team leader, have students try to determine one
Opportunity Finding	Explore Big Data and determine interesting questions or business problems that are likely to improve business performance if answered	Give students unintegrated data and ask students to make sense of it
Patience	Brainstorm with team members	Remind students that the right solution may need a few rounds of discussion
Positive attitude and Passion	Learn from others to build their confidence Answer themselves why analytics is essential and what can they benefit from the exercise	Encourage students that they can come up with interesting business problems by recalling all the relevant subject they have learned such as economics, finance, and marketing
Problem Solving	Solve a given business problem, come up with the right approach for solving a problem	Guide students through the steps of solving a problem if needed

Table 1: Soft Skills and Activities for Analytics Class		
Soft Skills Important for Analytics Graduates	**Activity Designing Plausible Things to Do**	
	Have students do	**What the Instructors can do to facilitate**
	Select the appropriate variables and determine the right methods after the business problems are defined	Advise students on how to find necessary resources
Team skills	Work in a small group setting	Have students team up in a group of three to four, and help them find a group if needed

The author has considered both the principle of business analytics, business problem and opportunity, and the current needs in the job market in designing this soft skill development activity. Based on the recent survey showing that up to 73% of company data goes unused (Barrett, 2018), the storytelling and opportunity finding skills are selected as the focus of the activity. To help students develop these two skills, the author asked the students to explore and try to make sense of the data provided. The author put to together publicly available data sources and shortlist to three sources. Following are some publicly available data sources:

International / government organizations:

- UN Data http://data.un.org/
- OECD https://data.oecd.org/
- US Census https://www.census.gov/
- US Public Data Visualization https://datausa.io/
- US Bureau of Economic Analysis https://www.bea.gov/
- World Economic Forum Reports http://reports.weforum.org/

Businesses:

- Tableau's Global Superstore https://community.tableau.com/docs/DOC-1236
- Wells Fargo's Campus Analytics Challenge https://www.mindsumo.com/contests/building-better
- Amazon Web Services public datasets http://aws.amazon.com/datasets

Three datasets were selected. For problem one, students use Tableau's Global Superstore and World Economic Forum's Global Competitive Index (i.e., unintegrated datasets); for another problem, the dataset used is Wells Fargo's. The author explained the objective of this exercise and asked the students to explore the data (e.g., variables, entities-relationship), list potential interesting questions that could add to a value of a company, list of the potential values added, and draft the outline of the presentation with dummy data.

This exercise is designed for students to work together in a group setting for one hour in class and about one to two hours outside of class time. This is a group assignment, and all students receive the same grade based on the following rubric. The students upload a Word file(s) to Canvas as a group submission and are also able to submit images of their presentation design.

- Questions and selected variables are appropriate
- Questions and potential value-added make sense
    - based on an instructor's experience and opinion
    - based on students' justification
- Questions and the output designs are appropriate

While students were working in class, the author walked around and talked to each group, and gave any advice where appropriate. Most groups could not complete this exercise in an hour, so they had to work outside of class for one to two hours on average.

INSTRUCTORS' GUIDE

Before having students do the above exercise, the instructors should have already taught the following: an overview of business analytics, data analytics tools in the market, data management, hands-on problem in combining data from multiple sources (with a business problem and analytics-ready datasets given), this could be exercise #1. After showing the hands-on problem, ask students how difficult it seems. Most students will say that it is challenging because it is their first time exposed to a selected software package, whether it is RStudio, SAS Enterprise Miner, or Tableau Prep.

Before having students start the exercise above, tell them that in the real world of Big Data, they might be asked to make senses of data and come up with interesting questions that would add value to a company if answered. In exercise #1, the business problem and precise requirement of the output file are provided, the students prepared the output file based given conditions. Exercise #2 is intended to be a soft skill development exercise assessing their storytelling and opportunity finding skills. Present the following form to them so they can self-assess their soft skills without the instructor mentioning the name of the skills.

	Specify how did you do in completing the exercise 1. I came up with one immediately. 2. I needed some time to come up with one. 3. I came up with one but not sure if it is good. 4. I had no idea, so I did not offer any to the group.
Figuring out an interesting business problem	
Figuring out the variables needed in the analyses	
Finding opportunities – potential improvement in business performance	
Designing a persuasive presentation	

DISCUSSION AND CONCLUSIONS

Lastly, soft skills development should be part of all exercises used in class and each focus on different skills. The exercise above does not cover all the skills listed in Table 1, but for certain covers at least opportunity finding, critical thinking, storytelling, and teamwork. Not all soft skills are equally important. To determine which ones are more critical to analytics, the author considered the principle of the analytics and the demand in the analytics market. In the next phase of this study, the author plans to develop a complete set of exercises or activities for an undergraduate business analytics course. With that, an assessment of both hard skills and soft skills for each exercise will also be developed and tested.

ACKNOWLEDGEMENT

Thank you, the reviewers, for the comments on how to improve the paper.

REFERENCES

Abbas, R., Abdul Kadir, F., & Ghani Azmie, I. (2013). Integrating soft skills assessment through soft skills workshop program for engineering students at University of Pahang: an analysis. *International Journal of Research In Social Science*, 2(1), 33-46.

Ahmed, F., Capretz, L. F., Bouktif, S., & Campbell, P. (2015). Soft skills and software development: A reflection from the software industry. *arXiv preprint arXiv:1507.06873*.

Barrett, J. (2018). 2019: Up to 73 Percent of Company Data Goes Unused for Analytics. Here's How to Put It to Work. Retrieved from www.inc.com/jeff-barrett/misusing-data-could-be-costing-your-business-heres-how.html, May 24.

Beard, D., Schwieger, D., & Surendran, K. (2008). Integrating soft skills assessment through university, college, and programmatic efforts at an AACSB accredited institution. *Journal of Information Systems Education*, 19(2), 229.

Blaszczynski, C., & Green, D. J. (2012). Effective Strategies and Activities for Developing Soft Skills, Part 1. *Journal of applied research for Business Instruction*, 10(1).

Briggs, S. (2015). 2019: 30 Tips to Cultivate Soft Skills in Your Students. Retrieved from www.opencolleges.edu.au/informed/features/30-ways-to-cultivate-soft-skills-in-your-students/, May 24.

Chamorro-Premuzic, T., & Frankiewicz, B. (2019). 2019: Does Higher Education Still Prepare People for Jobs?. Retrieved from hbr.org/2019/01/does-higher-education-still-prepare-people-for-jobs, May 9.

Dishman, L. (2016). 2019: These Are The Biggest Skills That New Graduates Lack. Retrieved from www.fastcompany.com/3059940/these-are-the-biggest-skills-that-new-graduates-lack, May 19.

Dubey, R., & Gunasekaran, A. (2015). Education and training for successful career in big data and business analytics. *Industrial and Commercial Training*, 47(4), 174-181.

Dutton, G. (2012). Taking Soft Skills for Granted? *Training*, 49(5), 48-50.

Ellis, M., Kisling, E., & Hackworth, R. G. (2014). Teaching soft skills employers need. *Community College Journal of Research and Practice*, 38(5), 433-453.

Gibb, S. (2014). Soft skills assessment: Theory development and the research agenda. *International journal of lifelong education*, 33(4), 455-471.

Gonzalez, M., Kasim, N. A., & Naimie, Z. (2013). Soft skills and dental education. *European Journal of Dental Education*, 17(2), 73-82.

Green, D. J., & Blaszczynski, C. (2012). Effective Strategies and Activities for Developing Soft Skills, Part 2. *Journal of Applied Research for Business Instruction*, 10(2).

Hagmann, J., Almekinders, C., Bukenya, C., Guevara, F., & Halemichael, A. (2003). Developing'soft skills' in higher education. *PLA notes*, 48, 21-25.

Henry, R., & Venkatraman, S. (2015). Big Data Analytics the Next Big Learning Opportunity. *Academy of Information & Management Sciences Journal*, 18(2).

Kerner, S. M. (2019). 2019: 8 Top Big Data Analytics Tools. Retrieved from www.datamation.com/big-data/top-big-data-analytics-tools.html, May 15.

Liberatore, M., & Luo, W. (2013). ASP, the art and science of practice: a comparison of technical and soft skill requirements for analytics and OR professionals. *Interfaces*, 43(2), 194-197.

McPherson, D. (2016). 2019: 10 Important Ways Analytics Skills Boost Your Resume. Retrieved from edx.org/10-ways-analytical-skills-boost-your-resume, May 15.

Mela, C. F., & Moorman, C. (2018). 2019: Why Marketing Analytics Hasn’t Lived Up to Its Promise. Retrieved from hbr.org/2018/05/why-marketing-analytics-hasnt-lived-up-to-its-promise, May 9.

Petter, S., & Randolph, A. B. (2009). Developing soft skills to manage user expectations in IT projects: Knowledge reuse among IT project managers. *Project Management Journal*, 40(4), 45-59.

Suciu, M. C., & Lacatus, M. L. (2014). Soft skills and economic education. *Polish journal of management studies*, 10(1), 161--168.

Wats, M., & Wats, R. K. (2009). Developing Soft Skills in Students. *International Journal of Learning*, 15(12).

White, D. (2014). 2019: 6 Steps To Improving Your Current Employees’ Soft Skills. Retrieved from www.fastcompany.com/3036793/6-steps-to-improving-your-current-employeess-soft-skills, May 9.

DECISION SCIENCES INSTITUTE

Analyzing the Profit-linked Multifactor Productivity Measurement Models

Mohan Rao
Texas A&M University-Corpus Christi
Email: mohan.rao@tamucc.edu

ABSTRACT

This paper analyzes the results from two different profit-linked multifactor productivity measurement models. Theoretical analysis was made in a 1989 *Management Science* article. Now the analysis is based on a real-world case from micro-irrigation industry. The APC (American Productivity Center) model was introduced by the American Productivity and Quality Center, and the PPP (Profitability = Productivity + Price Recovery) model was introduced by David Miller in a *Harvard Business Review* article in 1984.

KEYWORDS: performance evaluation; performance measurement; productivity; multi-factor productivity; chained-dollar; chain-type index; micro irrigation

INTRODUCTION

The desire to accurately determine firm performance led to the creation of several performance measurement systems. Among two of the most widely used approaches include American Productivity Center's total factor model (APC) and Ethyl Corporation's (PPP) model in which Profitability = Productivity + Price Recovery (Singh, Motwani, and Kumar, 2000). These two models link productivity performance to profitability by considering the issue of price adjustment.

In this paper, we introduce a revised comparison between APC and PPP its consequences and, ultimately, to gain a greater insight into the strengths and weaknesses of each model. This method builds on Miller and Rao (1989) by utilizing actual operational data. Its starting point is the actualization of performance, productivity, and price data and the application of deflation based on each model's requirements. To demonstrate how the analysis can be applied to produce insights on a company's performance measurement, we apply operational data from Finolex an Indian micro-irrigation firm.

This article contributes to two key areas. First, this study compares two models that on the surface appear to be similar, but because deflation is applied differently, it can provide management strikingly different assessments of the company's performance measurements. Hence, we fill a gap in the prior research that is scant on comparing similar models, especially the APC and PPP versions. Second, we complement our model comparison by analyzing actual organizational data from Finolex. The content of irrigation is essential to the world, as less than 1% of the Earth's water is available for irrigation (Whitmee et al., 2015). Therefore, it is vital for a manager to optimize the irrigation system to maximize output to assist farmers, households, etc.

The outline of the paper is straightforward. We start with a literature review outlining the APC and PPP model differences. This review is followed by detailed steps of the application of each model. We proceed with a comparison between the APC and PPP models and provide some managerial applications. The conclusions and limitations are presented at the end.

LITERATURE REVIEW

APC and PPP models correlate the relationship between productivity and profitability and have been previously used to measure profitability, productivity, and price recovery performance (Rao et al., 2018; Rao 2015). The application of both models extends into multiple contexts. For instance, researchers applied APC to a pharmaceutical production facility (Phusavat and Photaranon, 2006), wastewater plant (Rao and Phusavat, 2015), and micro-irrigation (Rao et al., 2018). Similarly, PPP has been applied to a wastewater treatment plant (Rao 2015) and the airline industry (Oum and Yu, 1998). The American Productivity and Quality Center developed APC in 1977, while Miller (1984) created the PPP model.

The APC model jointly measures the relationship and impact of productivity and price recovery with operational results and profitability (Phusavat and Photaranon, 2006), while PPP mainly focuses on a production unit with tangible outputs and inputs (Rao, 2013). Both models originate from similar accounting data and present the appropriate productivity, and price recovery impact in a dollar context rather than ratios or indexes. Moreover, the models utilize data from specific base periods to compare performance to future periods. The only difference between the models is how deflating techniques are applied (Rao, 2015). Essentially, the PPP model utilizes an aggregate deflation method and controls dollar values during its analysis. Conversely, the APC model contrasts period to period deflation to arrive at unit-less ratios (Singh, Motwani, and Kumar, 2000).

The PPP model's deflating methodology has its roots in chain-type annual-weighted index. Chain-type indexes strive to address one of the most fundamental obstacles in measuring output and prices of choosing the comparison base period (Landefeld and Parker 1997). To address this issue, the Bureau of Economic Analysis (BEA) introduced the Fisher index which uses the weights of the first period along with the Paasche index that uses the weights of the second period (Parker and Tripplet, 1996).

With this in mind, Miller (1986) conceptualized profitability = productivity + price recovery to address some of APC's shortfalls, including the deflating methodology. As Singh, Motwani, and Kumar (2000) indicate "the PPP approach provides a more suitable tool for examining productivity changes over a trend of at least three or more periods. On the other hand, since the APC approach utilizes only the base period prices, its price weights are independent and separable and more suitable for a two-period comparison" (p. 235).

Prior research is scant in comparing APC and PPP in a real operational setting. It appears Rao and Miller (1989) are the first to attempt such a comparison between the two models. Their findings, based on fictional data, discovered when multiple items are evaluated in multiple time periods, the results differ for each model. More specifically, their research found a nearly 10% deviance between the two models. However, for the past three decades, the comparison between APC and PPP is non-existent. A point, this article serves to resolve. We utilize the micro-irrigation industry in our article, as it has served as a suitable source for data for similar studies (Rao et al, 2018).

Micro-Irrigation Industry

The need to maximize the earth’s water supply is of utmost importance as municipalities and companies examine the best options to maximize water usage for the earth’s 10 billion inhabitant. A number that is expected to grow by 2 billion in the next 50 years (Brinegar and

Ward, 2009). Since much of the world food production depends on irrigation (Howell, 2001) and the need to increase food production to support the growing world population, improving the technical and economic performance of irrigation is critical to long term success (Brinegar and Ward, 2009). Therefore, many countries have started to focus on maximizing water conservation efforts to maximize water usage. A major challenge that many entities face is how to better irrigation water. For instance, the earth's surface consists of 71% water, while only 1% of those water sources is used for irrigation, drinking, and household demands (Rao, et al., 2018). To alleviate this problem, many different tactics have been applied. For instance, in the United States, drip irrigation has been applied to provide for crops, at the same time reduce water usage (Fereres and Soriano, 2006).

Since India is the most populous country in the world, it has taken great steps to maximize micro-irrigation. Hence, India has in excess of 150 drip and sprinkler irrigation companies, with an annual turnover of ~$750 million and growing at 25% annually. The market is dominated by 15 companies that have over 70% of the market share. With three companies controlling nearly 56% of the overall market – Jain Irrigation Systems, Ltd. 36%, Netafilm Irrigation Lrd. 14%, and Finolex Plasson Industries Pvt. Ltd. 5.5%. Our data focus centers around Finolex operational data. To better understand Finolex, it is important to know some of its key operational factors.

Measurement Models

We focus on comparing the two measurement models of APC and PPP. As explained earlier, both are similar yet different. Both models require at least two periods with data from revenue generating organizations. The data needs to include both outputs and inputs, with at least two of the following: quantities, prices, and values (revenues and expenses). The model demands fit our data perfectly as we secured two years-8 quarters of Finolex's operational data.

COMPARING THE MODELS

Conceptually it is understood the models differ on in the application of its deflating techniques. However, it is important to mathematically understand the differences between the two models. With this mind, the following section takes a closer look at both models, including the applicable equations.

APC and PPP Computations

The APC model considers many variables, but at its core, it calculates the profitability contribution by subtracting each input value change from the total output change ration and multiplying the resulting number the inputs value in the first period (Miller and Rao, 1989). To be more specific, APC model computation steps are:.

Step 1 Enter quantities and prices
Step 2 Calculate costs and values
Step 3 Calculate deflated values
Step 4 Calculate change ratios for quantities, price, and values
Step 5 Calculate performance indexes
Step 6 Calculate performance contribution for productivity, price recovery, and profitability

The PPP approach differentiates from the APC approach because its actual gross profit realized in a period is compared with the gross profit that would have been realized had the company's profit margins remained consistent. The PPP model computational steps are:

Step 1 Costs and revenues
Step 2 Implied and composite deflator for each input for all periods
Step 3 Input and output profit margins and deflated profit margins
Step 4 Profitability, productivity, and price recovery indices
Step 5 Productivity, price recovery, and profitability contributions

The implied deflator is the main difference between the two model. The implied deflator (ID_t) is the ratio of the total current period value to the value in previous period prices (deflated value).

$$ID_t = \frac{\prod_{i=2}^{t}\left(Q_{Xi}P_{X(i-1)} + Q_{Yi}P_{Y(i-1)}\right)}{\prod_{i=2}^{t-1}\left(Q_{X(i-1)}P_{X(i-1)} + Q_{Y(i-1)}P_{Y(i-1)}\right)}$$

Hence the applicable equations, are as follow.

Profitability	$= S_t\,(\text{Margin}_t - \text{Margin}_B)$
	$= S_t\,[((S_t - C_t)/S_t) - ((S_B - C_B)/S_B)]$
	$= (S_t C_B - S_B C_t)/S_B$
where Margin_B	= Base period gross profit margin
Margin_t	= Period t gross profit margin.
Productivity	$= S_{tD}\,(\text{Margin}_{tD} - \text{Margin}_B)$
	$= S_{tD}[((S_{tD}-C_{tD})/S_{tD}) - ((S_B-C)/S_B)]$
	$= (S_{tD}C_B - S_B C_{tD})/S_B$
where Margin_B	= Base period gross profit margin
Margin_{tD}	= Deflated gross profit margin in period t.
Price recovery	$= \text{Sales}_{tPR}\,(\text{Margin}_{tPR} - \text{Margin}_B)$

APC VS. PPP PERFORMANCE RESULTS

It is important to contrast the two models to highlight the key differences between the two perspectives. Our approach in this endeavor will visually compare both models. The following illustrations will focus on paralleling the overall performance, material, cost by category, employee, sales distribution, supplies, and sales/deflated sales.

Overall Performance

Figures 1 and 2 illustrate the overall performance calculated by each model. On the surface a manager would compare the two items and determine the profitability is similar under both models. The challenge for any manager is to take the information a step further to understand what is truly driving results. According to the APC model, after starting to decline after period four, price recovery bottomed out in period six, rising significantly periods seen and eight. In contrast, the dip after the fourth period in the PPP model is by far more drastic, dipping into the negative, before becoming positive in period eight. Similarly, both models differ on the evaluation of productivity. Where, PPP has a more drastic improvement after period six, than the APC model. It is obvious the deflating values have a more severe influence on the PPP model's evaluation of productivity and price recovery.

Figure 1 – APC Overall Performance

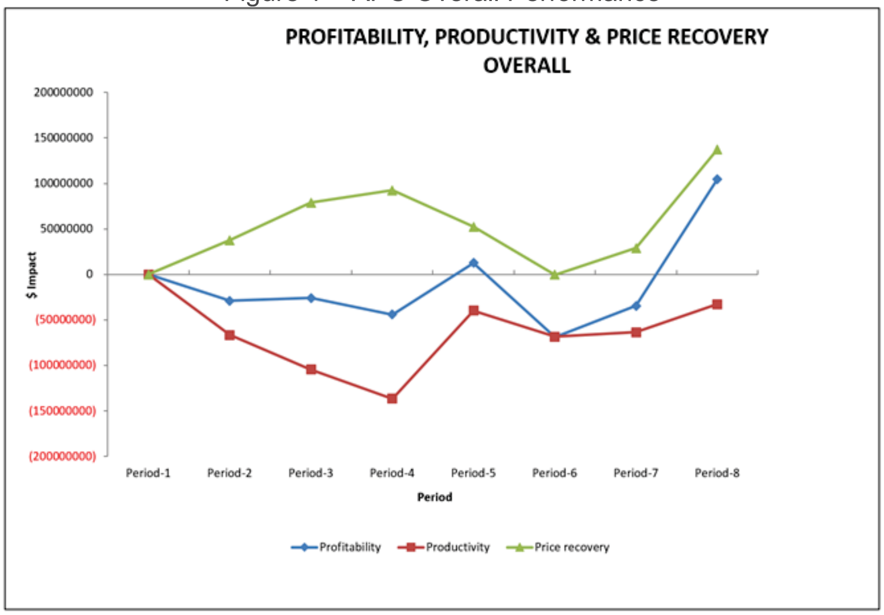
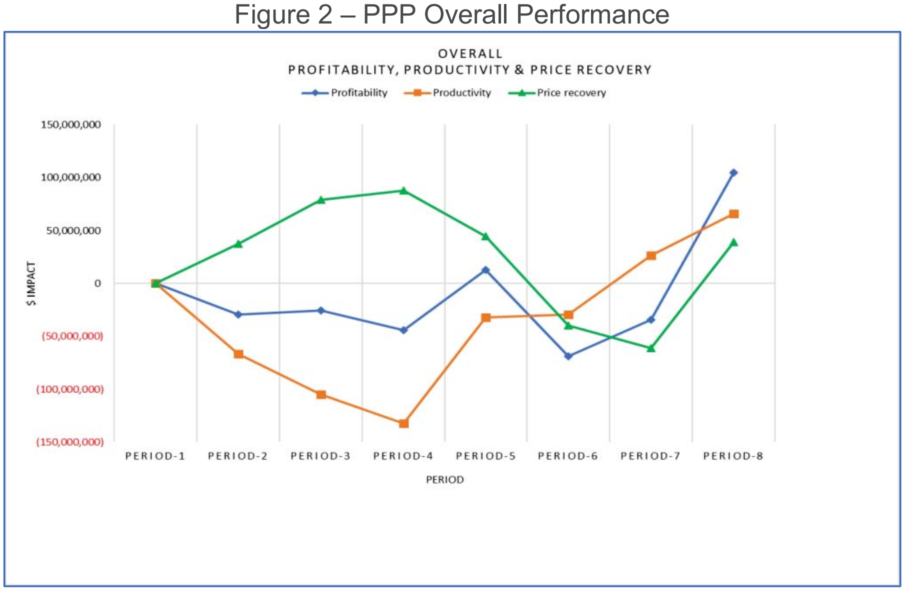


Figure 2 – PPP Overall Performance



In summary, there are differences and similarities between the models. We are not making a statement that one model is superior to the other, but merely indicating that different models will generate different results, even if the models have the same data. In this instance, the deflating approach utilized by the PPP model, offers different results when more periods are part of the calculation.

A few more differences to consider include the fact APC uses mainly ratios, while PPP uses mainly dollar amounts and profit margins. Moreover, the APC method is generally used for 2-period data. On the other hand, the PPP method is mostly used for 3 or more periods. The results of our study clearly demonstrate the influence of differences listed above.

SUMMARY AND CONCLUSIONS

Overall, most of the results differ from the APC to the PPP models. It is critical to recognize the difference in deflating techniques. As we mentioned earlier, it does not mean one model is better and the other model is worse. This study merely points out the differences between the models. In addition, both models can be easily created from available accounting data. As our study demonstrates with Finolex’s data, basic accounting data itemized by expense category provides the necessary basis to use either of the models.

Along with benefits of charts and graphs, each model provides the benefits of being able to drill down by category. For instance, we recognized the significant influence materials has on Finolex’s profitability. However, in order to truly comprehend business expense drivers, we were able to utilize both models to drill further by analyzing the effects of polyethylene and other traded items.

In summary our research compares two similar yet different measurement models that extend prior research by Miller and Rao (1989). The context of this study is water conservation. With the expected global population growth trends and the potential impact of climate change, micro irrigation could play a key role in water conservation and irrigation strategies.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE
The Complementarity between Human Resource Information Systems and Strategic Human Resource Practices: A Preliminary Empirical Study

Joonghee Lee
Auburn University
jzl0165@auburn.edu

ABSTRACT

The past decade witnessed two tends in HR field: emphasis of strategic role of HR and increase of use information technology (IT). This study examines the complementarity between human resource information systems (HRIS) and HR practices. Our findings suggest that HRIS positively moderate the relationship between HR practices and productivity.

KEYWORDS: Human resource information systems, Strategic human resource management,

INTRODUCTION

All specialized department in a corporation desires to demonstrate the value of what they do for the rest of organization. Human resource (HR) department is no exception. Although the human resource (HR) function is ubiquitous and human capital is considered a critical for organizational performance (Becker & Gerhart, 1996; Liu et al., 2017; Sun et al., 2007), human resource (HR) practitioners and researchers have been attempting to prove that HR practices make a contribution to the organization (Wright et al., 2005). In response to these pressures, a critical trend in HR management (HRM) has been seen over the past years. That is the emphasis on the strategic roles of HRM (SHRM) rather than its administrative role. In other words, HR practitioners are under pressure to be strategic partners and to contribute to a firm performance (Collins & Clark, 2003; Paauwe, 2004; Shin & Konrad, 2017). In addition, the past decade has witnessed a surge in the adaptation of e-HRM technology to achieve administrative and strategic benefits (Bondarouk et al., 2017; Marler & Parry, 2016). These are obvious HRM trends in the past. So previous research studied the impacts of HR practices and human resource information systems (HRIS) on economic efficiencies, but relatively fewer studies have focused on the complementarity between HRIS and SHRM. This study propose that these two phenomena are related and that the performance benefits of HR practices depend on HRIS.

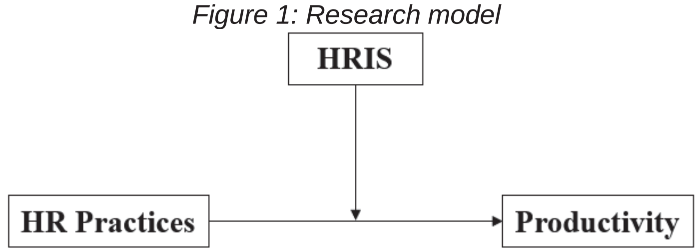
This study contributes to the literature in three ways. First, it deepens our understanding of the economic value of HR practices in an organization by proposing a holistic mode. In addition, this study contributes to our understanding of HRIS strategic values by testing the suggested model. There is a substantial body of research on HRIS with different foci, such as implementation (Lippert & Michael Swiercz, 2005), administrative efficiency (Bondarouk et al., 2009), employee perceptions towards HRIS (Wickramasinghe, 2010), and job satisfaction and turnover intention (Maier et al., 2013). However, little literature exists regarding HRIS's economic value. Third, it deepens our understanding the business value of IT by investigating a specialized information system within a traditional functional area of an organization. Demonstrating the value of information technology (IT) is a central issue in the information systems (IS) discipline, and a significant body of research has examined the organizational performance impacts of IT (e.g. Chae et al., 2014; Mithas & Rust, 2016; Mithas et al., 2017). IT in literature is frequently used as aggregate variables, such as IT investments (Mithas et al., 2012) and IT portfolio (Oh & Pinsonneault, 2007). However, scholars in the IS field have argued that we need to examine the

IT-performance link at a disaggregated level, such as at the level of specific information systems (Aral & Weill, 2007; DeSanctis, 1986; Mithas et al., 2006). By examining the link between specialized information system within a traditional functional area of an organization and the business value, this study extends our understanding regarding the IT-related value creation at an IT system level.

RESEARCH MODEL AND HYPOTHESIS

This study investigates the effects of HRIS on the relationship between strategic HR practices and firm performance. Specifically, the proposed model is shown in Figure 1.

Figure 1: Research model



HR practices and Productivity

The resource-based view (RBV) of the firm suggests that organizations can increase firm performance by using valuable and inimitable internal resources. Human capital is a potential source of competitive advantage (Barney, 1991). In this perspective, firms can enhance organizational performance by investing in people that constitute the organization-specific resource results in higher organizational performance. In addition, the behavioral perspective suggests that HR practices increase employee engagement in productive behavior (Shin & Konrad, 2019).

These theoretical explanations of the effect of HR practices on performance are empirically supported by previous research. For example, MacDuffie (1995) found that “bundles” of HR practices were related to productivity. Delery and Doty (1996) found significant relationships between HR practices and accounting profits. Youndt et al. (1996) found that combinations of HR practices were related to operational performance indicators. Also, many previous studies proved the positive relationship specific HR practice and firm performance, such as employee training (Delaney & Huselid, 1996; Kalleberg & Moody, 1994) and performance-based payment (Lazear, 2000; Way, 2002). More recent studies by Shin and Konrad (2019), and Wright et al. (2005) also showed the positive relationships between HR practice and firm performance.

Productivity is a commonly used measure of firm performance in the strategic human resource management field (Combs et al., 2006). HR practices increases productivity by enhancing employee knowledge, skills, and ability (Lawler, 1986). Also, the RBV and previous studies provide the theoretical and empirical base for predicting a positive effect of HR practices on firm productivity.

Hypothesis 1: HR practices is positively associated with productivity.

The Moderating Role of HRIS

According to the resource-based view (RVB), resources are not solely responsible for organizational performance (Wade & Hulland, 2004). In other words, the effects of organizational resources are viewed as both contingent and complementary. Complementarity refers the effect that one resource in organizations may have on another. Resources may influence performance when combined with another resource (Teece, 1986). In particular, IT resources can be more effective for improving organizational performance in conjunction with other firm resources, such as structure, culture, and skills (Chae et al., 2014). In particular, previous literature showed the complementary between a specific technology and a set of organizational practices (Aral et al., 2012; Black & Lynch, 2001; Bresnahan et al., 2002). For instance, Aral et al. (2012) found that the human capital management software adoption, HR analytics practices, and performance pay together act as a complementary system.

HRIS can help HR departments play that strategic role. Literature has broadly suggested two main benefits of HRIS: the improvement of efficiency and reduction of costs associated with HR activities involving day-to-day transactions (e.g. payroll and selection), and facilitation of a more strategic role for the HR function (Parry, 2011). A large proportion of administrative roles of HRM are now delivered through information technology (IT), which enables HR people to focus more on strategic roles. Literature in HRM and IS suggests that HRIS reduce administrative burdens of HR staff and allow them to focus on more strategic support for their organizations (Bondarouk & Ruël, 2013; Nagi & Wat, 2006). Specifically, HR departments can automate routine HR tasks and replace 'filing cabinets' by adopting HRIS (Parry et al., 2007), expend more effort on the improvement of HR strategies and policies for the strategic management of employees (Martin et al., 2008; Ruël et al., 2004; Shrivastava & Shaw, 2004) and support strategic decision-making (Bondarouk & Ruël, 2009; Hussain et al., 2007; Parry & Tyson, 2011). Thus, HRIS help transform HR professionals from administrative workers to strategic partners (Bell et al., 2006; Burbach & Dundon, 2005; Haines & Lafleur, 2008; Voermans & Van Veldhoven, 2007). The adoption of HRIS may allow the HR functions to increase its value in strategic levels.

Thus, I hypothesize that the effects of HR practices are more leveraged when they are implemented with HRIS.

Hypothesis 2: The HRIS positively moderate the relationship between HR practices and productivity.

METHOD

Sample and Procedure

The data come from the Human Capital Corporate Panel (HCCP) Survey from the Korean Research Institute for Vocational Education and Training (KRIVET, 2015). The survey data of the sample companies are combined with the corporate financial data held by the Korea Investors Service (KIS) using a corresponding firm code (KIS, 2015).

KRIVET has conducted the HCCP survey every two years since 2005; the survey includes a battery of questions regarding HR practices. We focused on the survey data and financial data in 2011 – 2015 for these reasons. First, the HCCP survey sample was changed in 2009. The

9th Korean standard industrial classification was placed into effect on February 1, 2008. Second, to eliminate the effects of the Great Recession, I focused on the data after that time.

For the survey process, KRIVET contacted HR managers at targeted firms, and they administered the HCCP survey using on-site interviews. The referent for the survey items was the firm, and HR managers responded to the items. The population for HCCP sample extraction consisted of companies listed in the company overview information of the KIS database with at least 100 employees across six industries: (1) Manufacturing, (2) Finance, banking and insurance, (3) Publishing/video/broadcasting and telecommunications, and information service, (4) Specialist/scientific and technical service (5), education service, and (6) Arts/sports and leisure service. In 2009 – 2015, companies for the survey sample numbered between 467 and 500.

Variable Operationalization and Descriptive Statistics

Table 2 provides the operationalization and the descriptive statistics for the variables in our study.

Strategic HR practices (SHRP). Theoretically, strategic HR practices are those that are related to overall organization performance. Prior research (Delery & Doty, 1996; Shin & Konrad, 2014) suggested practices that are considered strategic HR practices. These are internal career opportunities, training, compensation, employment security, employee involvement, and work design. Using HCCP data, SHRP is measured with four sets of practices in the area of internal career opportunities, training, compensation, and employment security (see Table 1). I use additive indices of HR practices following previous studies (Shin & Konrad, 2014; Wright et al., 2005). Each practice is coded as a binary variable based the response to a question for HR managers in the firm. This indicates whether they implement each HR practice (1 = yes, 0 = no). Adoption of practices is obtained by calculating the mean across the specific practices.

The first practices, Internal career opportunities, indicate the extent to which organizations can use internal labor markets to hire. The training measure refers to the formal training programs or policies given to employees for helping them develop their knowledge and skills. The HR practices in the compensation category focus on results-oriented compensation system that ties pay to individual and organizational performance. The fourth practice is the degree to which employees are given employment security.

HRIS adoption (HRIS). This is a binary variable that is based on the response to a question for HR managers in the firm that indicates whether they use HRIS. Response choices were 0 indicating the use of HRIS, and 1 meaning no use of HRIS.

Productivity. This variable is commonly used as a measure of financial performance in the SHRM field (Comb et al., 2006; Huselid, 1995). Consistent with prior research (Konrad & Mangel, 2000), I use a standard measure of productivity calculated as the logarithm of the gross operating revenue per employee.

Table 1: Strategic HR Practices

HR practice	Items
Internal career opportunities	- Does your company operate the internal job posting system? - Does your company develop core talents internally? - Does your company implement career development planning? - Does your company implement succession planning? - Does your company operate the job circulation?
Training	- Does your company provide orientation for new employees? - Does your company provide collective internal training? - Does your company provide a domestic training program? - Does your company provide an abroad training program? - Does your company operate the training leave system (provision of vacations and financial support for personal competence development)? - Does your company provide mentoring (or coaching)? - Does your company provide on the job training? - Does your company provide financial support for studying a domestic college program? - Does your company provide financial support for studying a domestic graduate school program? - Does your company provide financial support for studying an overseas graduate school program?
Incentive compensation	- Does your compensation system include the individual incentive? - Does your compensation system include the team incentive? - Does your compensation system include the division incentive? - Does your compensation system include the corporate incentive? - Does your compensation system include the profit sharing? - Does your compensation system include the stock option? - Does your compensation system include employee stock ownership plan?
Employment security	- Over the last 2 years, did your company ever adjust employment by lay-off, resignation advice, voluntary retirement, transfer to affiliate/partner, corporate division, and others?

Controls. The research model in this study controls for firm age (AGE), firm size measured as the logarithm of the number of employees (SIZE), management systems [fully managed by the owner (OWNER), managed by a management specialist with significant owner intervention (MAS_OWNER1), managed by a management specialist with slight owner intervention (MAS_OWNER2), fully managed by a management specialist (SPECIAL)], the top priority in the company management [improvement the quality of product/service(QUALITY), cost reduction (COST), new product development (DEV)], having an organization exclusively responsible for HR (HRDPT), number of HR employees (HREMP), number of HR managers (HRMG), the number of HR senior managers (HRSMG).

Table 2. Variable Description

Quantitative variable	Operationalization	Mean	Std. dev
SHRP	The average ratings on the items in Table 1 in a firm *i* in year *t*	8.079	1.038
Productivity	The logarithm of the gross operating revenue per employee in a firm *i* in year *t*	9.965	1.242
AGE	A firm *i*'s age in year *t*	33.914	17.620
SIZE	the logarithm of the number of employees in a firm *i* in year *t*	5.792	1.092
HREMP	The number of HR employees in a firm *i* in year *t*	2.841	4.941
HRMG	The number of HR managers in a firm *i* in year *t*	2.566	4.979
HRSMG	The number of HR senior managers in a firm *i* in year *t*	.421	.668
Categorical variable	Operationalization	Proportion	Std. dev
HRIS	1 if a firm *i* adopts and operates HRIS in year *t*	.700	.459
OWNER	1 if a firm *i* is fully managed by the owner, otherwise, coded as 0 in year *t*	.469	.499
MAS_OWNER1	1 if a firm *i* is managed by a management specialist with significant owner intervention, otherwise, coded as 0 in year *t*	.165	.371
MAS_OWNER2	1 if a firm *i* is managed by a management specialist with slight owner intervention, otherwise, coded as 0 in year *t*	.166	.372
QUALITY	1 if a firm *i*'s top priority in the company management is improvement the quality of product/service, otherwise, coded as 0 in year *t*	.523	.500
COST	1 if a firm *i*'s top priority in the company management is cost reduction, otherwise, coded as 0 in year *t*	.296	.457
HRDPT	1 if a firm *i* has an organization exclusively responsible for HR in year *t*	.685	.465

Preliminary analysis

Because of the panel nature of our data set, we specify the following equation for the panel models:

$$Y_{it} = X_{it}\beta + u_i + \varepsilon_{it} \quad (1)$$

where *Y* represents the endogenous variable (i.e. Productivity); *X* represents a vector of firm characteristics, such as HRIS, SIZE and other control variables; *β*s are the parameters to be estimated; subscript *i* indicates firms and subscript *t* indicates time; u_i represents unobserved time invariant fixed factors associated with a firm *i*; and *ε* is the error term associated with each observation. Our main independent variable (HR practices) is not a time-constant variable, and Hausman test rejected the null ($X^2(14) = 39.04$, $p < 0.01$) which is taken to mean that the key random effect assumption, the unobserved effect is uncorrelated with each explanatory variable, is false so that a fixed effect model is used (Wooldridge 2013).

RESULTS

A summary of the results of the fixed effects regression models are shown in Table 2. Model 1 includes the variable composite HR practices only. The coefficient of SHRP is not significant. The interaction term SHRP*HRIS in model 3 is positive and significant, supporting hypothesis 2.

Table 3: Fixed Effects Results

Independent Variables	Dependent variable: Productivity		
	Model1	Model2	Model3
SHRP	-.044 (.041)	-.045 (.041)	-.156* (.076)
HRIS		.088 (.094)	-.943 (.560)
SHRP*HRIS			.124^{+*} (.072)
SIZE	-.502** (.152)	-.494 (.151)	-.494** (.151)
AGE	-.008 (.018)	-.016 (.020)	-.012 (.020)
OWNER	.010 (.164)	.111 (.163)	.115 (.163)
MAS_OWNER1	-.190 (.168)	-.183 (.168)	-.176 (.168)
MAS_OWNER2	-.019 (.136)	-.020 (.135)	-.013 (.135)
QUALITY	.069 (.109)	.081 (.108)	.080 (.108)
COST	.067 (.129)	.070 (.127)	.073 (.126)
HRDPT	-.268* (.111)	-.273 (.110)	-.282* (.109)
HREMP	-.007 (.010)	-.008 (.010)	-.008 (.010)
HRMG	.007 (.013)	.008 (.013)	.007 (.013)
HRSMG	.066 (.050)	.069 (.051)	.079 (.051)
*Note: Standard errors in parentheses; *** p<0.001, ** p<0.01, *p<0.5, $^{+}$p<0.5 (All analyses employ two-tailed tests of statistical significance)*			

CONCLUSION AND FUTURE RESEARCH

The objective of this study was to analyze and investigate the complementary relationship between HRIS and strategic HR practices. Based on our preliminary analyses, we find that HRIS positively moderates the impact of HR practices on productivity. It is a preliminary analysis. In the future, we will do alternative analysis by using different econometric models, such as 2SLS, to handle any endogeneity that might exist in the data and present more detailed findings and implications for future research.

REFERENCES

Aral, S., Brynjolfsson, E., & Wu, L. (2012). Three-way complementarities: Performance pay, human resource analytics, and information technology. *Management Science*, *58*(5), 913-931.

Aral, S., & Weill, P. (2007). IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation. *Organization Science*, *18*(5), 763-780.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of management journal*, *39*(4), 779-801.

Bell, B. S., Lee, S. W., & Yeung, S. K. (2006). The impact of e-HRM on professional competence in HRM: Implications for the development of HR professionals. *Human Resource Management*, 45(3), 295–308.

Black, S. E., & Lynch, L. M. (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics, 83*(3), 434–445.

Bondarouk, T., Parry, E., & Furtmueller, E. (2017). Electronic HRM: four decades of research on adoption and consequences. *The International Journal of Human Resource Management*, *28*(1), 98-131.

Bondarouk, T., & Rue¨l, H. (2009), Electronic human resource management: Challenges in the digital Era. *The International Journal of Human Resource Management*, 20(3), 505–514.

Bondarouk, T., & Ruël, H. (2013). The strategic value of e-HRM: Results from an exploratory study in a governmental organization. *The International Journal of Human Resource Management*, *24*(2), 391-414.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. 2002. Information technology, workplace organization and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics, 117*(1), 339–376.

Burbach, R., & Dundon, T. (2005). The strategic potential of human resource information systems: Evidence from the Republic of Ireland. *International Employment Relations Review*, *11*(1/2), 97-117.

Chae, H. C., Koh, C. E., & Prybutok, V. R. (2014). Information technology capability and firm performance: Contradictory findings and their possible causes. *MIS Quarterly*, *38*(1), 305-326.

Collins, C. J., & Clark, K. D. (2003). Strategic human resource practices, top management team social networks, and firm performance: The role of human resource practices in creating organizational competitive advantage. *Academy of management Journal*, *46*(6), 740-751.

Combs, J. G., Liu, Y., Hall, A. T., & Ketchen, D. J. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, *59*(3), 501-528.

Delaney, J. T., & Huselid, M. A. (1996). The impact of human resource management practices on perceptions of organizational performance. *Academy of Management Journal*, 39(4), 949-969.

Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of management Journal*, *39*(4), 802-835.

DeSanctis, G. (1986). Human resource information systems: a current assessment. *MIS Quarterly*, 10(1), 15-27.

Haines, V. Y., & Lafleur, G. (2008). Information technology usage and human resource roles and effectiveness. *Human Resource Management, 47*(3), 525-540.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of management journal*, *38*(3), 635-672.

Hussain, Z., Wallace, J., & Cornelius, N. E. (2007). The use and impact of human resource information systems on human resource management professionals. *Information & Management*, *44*(1), 74-89.

Kalleberg, A. L., & Moody, J. W. (1994). Human resource management and organizational performance. *American Behavioral Scientist*, 37(4), 948-962.

KIS 2015. KISData. Available at: http://www.niceinfo.co.kr/business/KISData.nice

Konrad, A. M., & Mangel, R. (2000). The impact of work-life programs on firm productivity. *Strategic Management Journal*, 21(12), 1225-1237.

Korea Research Institute for Vocational Education & Training (KRIVET) (2015). About HCCP. Available at: http://www.krivet.re.kr/eng/eu/eh/euDDALs.jsp

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, *90*(5), 1346-1361.

Lawler, E. E. (1986). *High-involvement management*. San Francisco: Jossey-Bass.

Lippert, S. K., & Michael Swiercz, P. (2005). Human resource information systems (HRIS) and technology trust. *Journal of Information Science*, *31*(5), 340-353.

Liu, D., Gong, Y., Zhou, J., & Huang, J. C. (2017). Human resource systems, employee creativity, and firm innovation: The moderating role of firm ownership. *Academy of Management Journal*, *60*(3), 1164-1188.

MacDuffie JP. (1995). Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. *Industrial & Labor Relations Review, 48*(2), 197–221.

Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2013). Analyzing the impact of HRIS implementations on HR personnel's job satisfaction and turnover intention. *The Journal of Strategic Information Systems*, *22*(3), 193-207.

Marler, J. H., & Parry, E. (2016). Human resource management, strategic involvement and e-HRM technology. *The International Journal of Human Resource Management*, *27*(19), 2233-2253.

Martin, G., Reddington, M., & Alexander, H. (2008). *Technology, Outsourcing and Transforming HR*, Amsterdam: Elsevier.

Mithas, S., Almirall, D., & Krishnan, M. S. (2006). Do CRM systems cause one-to-one marketing effectiveness? *Statistical Science,* 21(2), 223-233.

Mithas, S., & Rust, R. T. (2016). How information technology strategy and investments influence firm performance: Conjecture and empirical evidence. *MIS Quarterly*, *40*(1), 223-245.

Mithas, S., Tafti, A., Bardhan, I., & Goh, J. M. (2012). Information technology and firm profitability: mechanisms and empirical evidence. *MIS Quarterly*, 36(1), 205-224.

Mithas, S., Whitaker, J., & Tafti, A. (2017). Information technology, revenues, and profits: Exploring the role of foreign and domestic operations. *Information Systems Research*, *28*(2), 430-444.

Ngai, E. W. T., & Wat, F. K. T. (2006). Human resource information systems: A review and empirical analysis. *Personnel Review*, *35*(3), 297-314.

Oh, W., & Pinsonneault, A. (2007). On the assessment of the strategic value of information technologies: conceptual and analytical approaches. *MIS Quarterly*, 31(2), 239-265

Paauwe, J. (2004). *HRM and Performance: Achieving Long Term Viability*, Oxford: Open University Press.

Parry, E. (2011). An examination of e-HRM as a means to increase the value of the HR function. *The International Journal of Human Resource Management*, *22*(5), 1146-1162.

Parry, E., & Tyson, S. (2011). Desired goals and actual outcomes of e-HRM. *Human Resource Management Journal*, *21*(3), 335-354.

Parry, E., Tyson, S., Selbie, D., & Leighton, R. (2007), *HR and Technology: Impact and Advantages*, London: Charted Institute of Personnel and Development.

Rue¨l, H.J.M., Bondarouk, T.V., & Looise, J.C. (2004), *E-HRM: Innovation or Irritation. An Exploration of Web-Based Human Resource Management in Large Companies*, Utrecht: Lemma Publishers.

Shin, D., & Konrad, A. M. (2017). Causality between high-performance work systems and organizational performance. *Journal of Management*, *43*(4), 973-997.

Shrivastava, S., & Shaw, J. B. (2003). Liberating HR through technology. *Human Resource Management*, *42*(3), 201-222.

Sun, L. Y., Aryee, S., & Law, K. S. (2007). High-performance human resource practices, citizenship behavior, and organizational performance: A relational perspective. *Academy of management Journal*, *50*(3), 558-577.

Teece, D. J. (1986). Firm Boundaries, Technological Innovation and Strategic Planning,. in *The Economics of Strategic Planning*, G. L. Thomas (ed.), D. C. Heath, Lexington, MA, 187-199.

Voermans, M., & Van Veldhoven, M. (2007), Attitude Towards e-HRM: An Empirical Study at Philips. *Personnel Review*, 36(6), 887–902.

Wade, M., & Hulland, J. (2004). The resource-based view and information systems research: Review, extension, and suggestions for future research. *MIS Quarterly*, *28*(1), 107-142.

Way, S. A. (2002). High performance work systems and intermediate indicators of firm performance within the US small business sector. *Journal of management*, *28*(6), 765-785.

Wickramasinghe, V. (2010). Employee perceptions towards web-based human resource management systems in Sri Lanka. *The International Journal of Human Resource Management*, *21*(10), 1617-1630.

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach 5th*, Mason, OH: South-Western Cengage Learning

Wright, P. M., Gardner, T. M., Moynihan, L. M., & Allen, M. R. (2005). The relationship between HR practices and firm performance: Examining causal order. *Personnel Psychology*, 58(2), 409-446.

Youndt MA, Snell SA, Lepak DP. (1996). Human resource management, manufacturing strategy, and firm performance. *Academy of Management Journal, 39*(4), 836–866.

DECISION SCIENCES INSTITUTE

Varied Impacts and Destination Choices of Different Interhospital Transfer Types

Raymond Lei Fan
C.T. Bauer College of Business
University of Houston
4750 Callhoun Road, Houston, Texas 77204
Email: raymondfan@bauer.uh.edu

Ming Zhao
University of Houston
Email: mzhao@bauer.uh.edu

David Xiaosong Peng
University of Houston
Email: xpeng@bauer.uh.edu

ABSTRACT

We first differentiate between clinical transfer and non-clinical transfer based on their varied coordination properties, and develop a practical method to separate the two transfer types using patient-level discharge data. We evaluate the impacts of the two transfer types on care outcomes (measured by readmission, mortality, and LOS). Our analysis yields two major findings. First, clinical transfers are more routinized and are associated with better care outcomes than non-clinical transfers. Second, transfer destination choices matter differently for the two transfer types. Specifically, for non-clinical transfers, destinations within the same multihospital system are associated with better care outcomes than those out of the system, whereas for clinical transfers, destinations out of the system are associated with better care outcomes than those in the system. The robustness of the models are checked using propensity score matching techniques. The results shed light on the nature of the two different transfer types, and provide practical destination choice strategies to practitioners.

KEYWORDS: hospital operations management; interhospital transfer types; destination choice strategy

INTRODUCTION

Interhospital transfer (IHT) of patients occurs commonly between hospitals (Iwashyna, 2012) when patients' needs exceed the resources available at the current hospital. The increasing specialization of a single hospital, especially in urban areas, combined with the demand for sophisticated medical treatments, generates a great need for IHT. Although IHT represents an effort of hospitals to provide better care through inter-provider collaboration, it often increases risks to patients and leads to the development of clinical complications (Barry & Ralston, 1994). In order to develop strategies to better manage IHT, it is important to identify different IHT types and understand the individual operations nature of each type.

Because the reasons underlying IHT are largely unknown, most literature focuses on disease and site-specific studies so that a typical transfer type can be identified. In this paper, we specify an IHT from a sending hospital to a receiving hospital as *clinical transfer* if the sending hospital (also referred to as the "sender") does not offer the specialty or higher levels of care needed for

the patient while the receiving hospital (also referred to as the “receiver”) does, otherwise as *non-clinical transfer*. Our definition of clinical transfer depends on a hospital’s clinical service scope, which is a long term strategic plan and forms the basis of the lower-level operational and tactical decisions (Ivanov, 2010). Services that are outside of the scope but are required for patients will need to be “outsourced”, i.e., those patients need to be transferred to another hospital for further care. In contrast, a non-clinical transfer occurs due to the ad hoc status of operational factors. One major cause of non-clinical transfers is the unavailability of medical staff or beds due to ED crowding (Morley, Unwin, Peterson, Stankovich, & Kinsman, 2018).
The two IHT types have fundamentally different decisions models. As a result of a hospital’s service scope strategy, clinical transfer decisions are dictated by the strategic-level service scope decision, and the decisions tend to be repeated from case to case. Non-clinical transfer decisions, on the contrary, need ad hoc tactical evaluation of the operational factors, causing decision inconsistency because of stochastic operational situations. In the literature, there is no known analysis of the different intrinsic natures between the two IHT types, or methodology to operationalize the differentiation of them. Our study fills the gap.
We develop our hypotheses using Organizational Information Processing Theory (OIPT) (Galbraith, 1974) and empirically test it with the 2014 Florida community hospitals’ emergency department and inpatient discharge data, which are made available by the Healthcare Cost and Utilization Project (HCUP), and the 2014 American Hospital Association (AHA) annual survey data. We find that non-clinical transfers are associated with worse care outcomes as measured by 30-day hospital readmission, in-hospital mortality, and LOS, than clinical transfers. Distinct from previous disease- and site-specific studies about IHT, to the best of our knowledge, our paper is the first to estimate the varying effects of the two IHT types on care outcomes.
Hospitals in the US have a strong tendency to transfer patients within the same hospital system (Lu & Lu, 2018; Bosk, Veinot, & Iwashyna, 2011). Hospitals may have more effective communication for transfer arrangement with hospitals within the same system than with those outside the system. Hospitals in the same system are less likely to reject the transfer requests from within the system and more likely to have implemented transfer protocols among them. Therefore, choosing the IHT destinations inside the same system may improve care efficiency. However, without knowing the specific IHT types, it is hard to tell if this practice is the optimal choice, as the routinization of clinical transfers may offset the communication convenience of in-system coordination. Is it better for a patient to be transferred out of the system or within the system? Will the answer be the same for the two IHT types? We investigate the effect of system-based destination choices under the scope of coordination mechanisms, and find that the optimal choices diverge. For non-clinical transfers, in-system destinations offer higher care efficiency (measured by LOS) with the same care quality (measured by readmission and mortality), while for clinical transfers, out-of-system destinations offer the same care efficiency with higher care quality. This finding extends the findings of Lu and Lu (2018), and offers more general destination choice strategies for hospitals to use for different IHT types, considering both care quality and care efficiency.

LITERATURE REVIEW

IHT is an understudied area where little is known about the process of information transfer and its association with patient outcomes (Herrigel, Carroll, Fanning, Steinberg, Parikh, & Usher, 2016). Our research contributes to the empirical healthcare operations literature on categorizing IHT types and comparing their effects on the care performance.

Impacts of IHT

Inter-organizational communication of IHT is challenging. Once a patient has been identified to need transfer, the medical staff of the sending hospital needs to request and negotiate with the candidate receiving hospitals (Bosk et al., 2011). The negotiation can be an intense process (Iwashyna, 2012), as the sending physician needs to explain why the patient needs to be transferred and provide transfer timing and method to the referred physician (Dunn, Gwinnutt, & Gray, 2007). The negotiation is a time consuming process (Bosk et al., 2011). Identifying a hospital that accepts the patient may take multiple phone calls and multiple rounds of negotiation (Rosenberg, Hofer, Strachan, Watts, & Hayward, 2003).
Because of the above challenges, IHT can impact patient care performance, which are commonly measured by readmission, mortality and LOS. Table 1 summarizes findings of several recent papers on the relationship between IHT and care performance. Most studies generally find that transfers has unfavorable impacts on care performance. However, several studies find neutral or even favorable impacts. This difference is likely due to the scope of these studies. Most the findings are based on a single hospital setting or a single disease. Furthermore, most studies do not differentiate between transfer types. Without understanding the nature of each IHT type, it is hard to discern their varied effects on care performance or act on specific cases to minimize the potential negative consequences. Our study uses a state-wide discharge data that covers all hospitals in the region and all conditions, and we differentiate between the two IHT types.

Table 1 Recent Studies of the Clinical Impact of IHT on Patient Care

Paper	**Subject**		**Effects on Care**		
	Disease	**Setting**	**Readmission**	**Mortality**	**LOS**
Mun˜oz et al., 1988	All	Single hospital		+	+
Kerr and Byrd, 1989	All	One hospital		+	+
Bernard et al., 1996	All	One hospital			+
Gordon and Rosenthal, 1996	All	One hospital		+	+
Selevan et al., 1999	All	Multiple hospitals	0	0	0
Duke and Green, 2001	All	One hospital		0	+
Rosenberg et al., 2003	All	One hospital		+	+
Nathens et al., 2003	Trauma	One hospital		0	0
Combes et al., 2005	All	One hospital		+	+
Stolte et al., 2006	All	Multiple hospitals			+
Golestanian et al., 2007	All	One hospital		+	+
Flabouiris et al., 2008	All	Regional hospitals		+	+

Odetola et al., 2009a	All	Regional hospitals		+	+
Odetola et al., 2009b	All	One hospital		+	+
Wiggers et al., 2011	Hip fracture	One hospital	+	+	+
Hill et al., 2011	Trauma	Multiple hospitals		0	0
Barratt et al., 2012	All	Regional hospitals			+
Russell et al., 2015	All	Two hospitals	-	-	-
Sokol-Hessner et al., 2016	All	Multiple hospitals		+	+

"+" indicates significant and positive; "-" indicates significant and negative; "0" indicates not significant.

IHT Types

Several studies have taken various approaches to categorize IHT types. Duke and Green (2001) categorize transfers for specialty care or higher level of care as Category A transfer and those due to a hospital's temporary unavailability of beds as Category B transfer. Gray et al. (2003) split transfers into transfers for specialty care and transfers for non-clinical reasons. Stolte et al. (2006) differentiate transfers due to capacity from those for specialty services. While these categorization concepts make sense, they do not identify the fundamental difference in the nature of the decisions in various IHT types.

We differentiate between two IHT types. The first type is transfers for clinical reasons, e.g., for specialty care or higher level of care. The transfer decision in this type originates from a hospital's services scope, which is a strategic level decision (Butler, Leong, & Everett, 1996; Li, Benton, & Leong, 2002). The second type is transfers for non-clinical reasons, e.g., bed unavailability, patient preference, etc. The transfer decision in this type is made after tactically evaluating several operational factors such as a hospital's capacity resource and the patient's hospital preference. We categorize the first type as clinical transfer, and the second type as non-clinical transfer.

In many cases, it is technically difficult to separate the data based on the transfer type definitions. Our study develops a methodology to identify different transfer types based on the definition above.

IHT Destination Choices

Our study also contributes to the optimal destination choices in IHT. According to the Future of Emergency Care report released by the Institute of Medicine, "regionalization is an active process by which patients are appropriately matched to appropriate resources" (Carr & Martinez, 2010). For care regionalization to succeed, it is critical to choose the best destinations for IHT.

Literature on IHT destination hospital choice suggests that hospitals generally choose destinations based on same-system relationship (Lu & Lu, 2018). Transfers within the same system can have several advantages, e.g., it is easier to implement transfer protocols within the same multihospital system. Hospitals within the same system typically have a better relationship since they have shared goals, shared knowledge and better communication platform (Gittell, Fairfield, Bierbaum, Head, Jackson, Kelly, Laskin, Lipson, Siliski, & Thornhill, 2000). Although preferred by hospitals, transferring within the same system also can have drawbacks. Iwashyna

(2012) suggests that transfers based on relationship tend to overlook service quality and neglect patients' clinical interests. Lu and Lu (2018) find that relationship-based (within the same system) patient transfers led to worse clinical outcomes than distance-based and quality-based transfers.
We view system-based IHT destination choice as a tool that a hospital may use to coordinate the complex transfer procedures. Inter-organizational tasks such as IHT require multi-functional teams and departments to provide a wide variety of services. These teams and departments may have different task specializations and work priorities (Tucker, Nembhard, & Edmondson, 2007), which may clash. The good relationship between hospitals within a system may minimize the conflicts.

HYPOTHESIS DEVELOPMENT

Clinical transfers and non-clinical transfers originate from different levels of decisions and thus possess fundamentally different operations characteristics. Clinical transfers originate from a hospital's limited service scope. In this case, hospital decision makers can accurately predict that patients out of service scope will be transferred, and preplan for that situation. The relative stability of the service scope and the preplanning nature of clinical transfers lead to the repetition of transfer practice: whenever a patient requires some services out of scope, they must be transferred, unless the patient refuses further treatment. They also promote the sending hospital to establish long-term relationships with the receiving hospitals, e.g., tertiary hospitals and medical centers that offer specialty care or higher levels of care. In practice, sending hospitals tend to establish IHT protocols and agreements with the capable receiving hospitals to govern the long-term relationships. Practice repetition and interhospital protocols nurture the decision and process routinization in clinical transfers. Non-clinical transfers, in contrast, do not nurture routinization due to the ad hoc nature of the decisions. First, whether a patient is transferred depends on the operational situations at a specific time for a specific patient. For example, the decision to transfer a patient due to bed unavailability may be a slow process as bed occupation level is time dependent and stochastic. The patient may just need to wait for a short while before beds become available. Second, patients' choices vary. For example, between two patients who need to be admitted when no bed is available, one patient may prefer to wait while the other may prefer to be transferred. Third, hospitals view patients as sources of revenue (Veinot et al., 2012), and generally are reluctant to transfer patients and their revenues out (Kahn, Asch, Iwashyna, Rubenfeld, Angus, & Asch. 2008). In non-clinical transfers, hospitals typically have the option of retaining the patient instead of transferring. Hospitals need to weigh the benefits and costs of retaining versus transferring. Therefore, nonclinical transfers are less routinized than clinical transfers, and we advance our following hypothesis.

Hypothesis 1. Clinical transfers are more routinized than non-clinical transfers.

As a process becomes more complex, the amount of information to be processed increases, and if not managed well, will affect a firm's performance. Organizational information processing theory (Galbraith, 1974) prescribes several coordination mechanisms to organizations to increase their information processing capabilities. Formal mechanisms include departmentalization of organizational units, centralization of decision making through the formal hierarchy, procedure standardization, strategic planning, and output control such as financial performance goal. Informal mechanisms include lateral (cross-unit or cross-organizational) relations, informal communication through personal relationships, and socialization through culture-building (Martinez & Jarillo, 1989).

Clinical transfers are coordinated with both formal and informal coordination mechanisms. Formal mechanisms for clinical transfers include strategic planning and procedure standardization. A hospital usually preplans for scenarios where it is incapable of caring, because it knows what patients can be cared and what cannot. The plan is typically strategic as it does not change in the short term. Furthermore, clinical transfers are usually routinized, therefore, they usually have standard procedures. Informal mechanisms such as cross-organizational relations and personal relationships are also prevalent in clinical transfers. Furthermore, personal relationships, if exist between the sending physician and the receiving physician, may dominate the referral decision (Forrest, Nutting, Starfield, & Von Schrader, 2002). In non-clinical transfers, formal mechanisms are weak and informal coordination mechanism is largely missing. Due to decision inconsistency, non-clinical transfers are less likely to have sender-receiver transfer protocols. Lack of routinization prevent the transfer process to follow standard procedures. The ad-hoc nature of the transfer decisions makes the sender-receiver relationship temporary, suppressing long-term personal relationships between the referring and referred physicians.

Having both formal and informal coordination mechanisms, clinical transfers may be better coordinated than non-clinical transfers, which essentially have no informal coordination mechanism. Better coordination reduces communication breakdowns, care delays, and medical errors, improving patient care quality and care efficiency. Therefore, we hypothesize the following.

Hypothesis 2A. Clinical transfers have better care quality (lower readmission and lower mortality) than non-clinical transfers.

Hypothesis 2B. Clinical transfers have better care efficiency (shorter LOS) than non-clinical transfers.

In cases where a sender and a receiver belong to the same hospital system, the sender has the option to choose destinations based on its system membership. If the sender choose to transfer within the system, formal coordination mechanisms such as centralization of decision making and output control by financial performance goal may affect the IHT coordination.

In practice, a physician' primary concern about IHT is the fractured and contentious process of negotiating with the candidate destination hospital (Iwashyna & Courey 2011; Iwashyna, 2012). The concern is especially valid for non-clinical transfers because in such cases the sender is also capable of treating the patient, and the receiver needs to be convinced that the patient indeed needs to be transferred (Bosk et al., 2011). The negotiation would be much simpler if the receiver belongs to the same hospital system, since hospitals within the same system are less likely to reject a transfer request. Therefore, for cases of same-system transfers, same-system convenience implicitly carried out the mechanism of centralization of decision making. Further, a hospital system may explicitly centralize transfer navigation system, transfer case management and administrative procedures such as transmission of IHT paperwork.

Moreover, hospitals within the same system typically share some common financial interests (Veinot et al., 2012). Upon choosing a same-system hospital as the destination, the sender retains the revenue from the care within the system (Bosk et al., 2011).

The two system-induced coordination mechanisms discussed above may have varied effects on the two IHT types, as clinical transfers already have other strong formal and informal mechanisms, and non-clinical transfers only have other weak mechanisms. For clinical transfers, the coordination mechanisms of strategic planning and procedure standardization are likely to dominate the system-induced coordination mechanism. This is especially true for clinical transfers due to regionalization of care, where the regionalization protocols instead of system memberships direct where to transfer. For example, a Level 1 trauma center has the

responsibility to accept patients from lower level providers and the transfers are governed by established protocols, regardless of the hospitals' system.
For non-clinical transfers, the system-induced coordination mechanisms weighs more since the existing coordination mechanisms are weak. Therefore, we argue that non-clinical transfers benefit more from same-system relationship than clinical transfers.

Hypothesis 3A. Non-clinical transfers within the system have better care quality (lower readmission and lower mortality) than those outside of the system.

Hypothesis 3B. Non-clinical transfers within the system have better care efficiency (shorter LOS) than those outside of the system.

Hypothesis 3C. Clinical transfers within the system have better care quality (lower readmission and lower mortality) than those outside of the system, but the difference is smaller than for non-clinical transfers.

Hypothesis 3D. Clinical transfers within the system have better care efficiency (shorter LOS) than those outside of the system, but the difference is smaller than for non-clinical transfers.

METHOD

To test our hypotheses, we obtain patient level visit data that has patients' demographics, visit information, diagnosis and treatment service details, as well as hospital level data that has hospital characteristics. Next, we develop a method to differentiate the IHT types based on the definition. Lastly, we compare the differentiated IHT and test the hypotheses.

Data

We use the Healthcare Cost and Utilization Project (HCUP) 2014 Florida State Emergency Department Databases (FL-SEDD) and the State Inpatient Databases (FL-SID) obtained from the Agency for Healthcare Research and Quality (AHRQ) for the patient level variables, service details and visit information, and use the 2014 American Hospital Association (AHA) annual survey data, as well as the readmission data from the Centers for Medicare & Medicaid Services (CMS) website (https://www.cms.gov) for the hospital level variables. There are 2,381,173 inpatient discharge observations from FL-SID, and 6,887,904 ED discharge observations from FL-SEDD. The two datasets include 299 common hospitals, 217 of which also appear in the AHA and readmission data.

Transfer Definition and Scope

Similar to Lu and Lu (2018) in a study of IHT of heart attack patients, we restrict our analysis to cases where patients are transferred from one hospital's ED to another hospital's inpatient department with emergent admissions (an example of a non-emergent admission is an elective admission). The restriction is due to two considerations. First, compared with transfers out of inpatient, transfers out of the ED are more time sensitive, and more critical to the patients and need quicker decisions. Second, emergent transfers happen within a short period of time, during which little unobserved endogenous factors may affect the LOS. For example, in emergent transfers, patients and families have less influence on the transfer decisions. Our study focuses on emergent transfers that need some procedures (one or more procedures) at the receiving hospitals since these cases involves higher risks to the patients and need more clinical resources than cases where no procedures are performed.

Out of the 2,381,173 inpatient records, 755,867 observations are emergent admissions needing procedures, and 12,016 of them are identified as transfers. At the hospital-level, 208 out of 237 hospitals have sending records and 236 out of 237 hospitals have receiving records.

Differentiating IHT Types

The HCUP data includes complete service codes that a hospital has performed on each patient. DX codes represent diagnosed conditions and PR codes represent performed procedures. Based on the codes, we deduct whether a transfer is a clinical transfer or not. If the receiver has performed any DXs or PRs that the sender never performed during the full year, we consider that the sender transferred the patient due to clinical reasons. Otherwise, the transfer is made due to non-clinical reasons. The data generates 9,716 non-clinical transfers and 9,854 clinical transfers. Following the above differentiation process, we process the data and obtain the differentiation result summarized in Table 2. We use the dichotomous variable CLN_TRANS to indicate IHT types.

Table 2 Transfer Type Classification Result

CLN_TRANS	Frequency	Percent
0 (Non-clinical transfer)	4,434	36.90
1 (Clinical transfer)	7,582	63.10

Table 3 (see Appendix) lists the most frequent principle diagnosis codes in IHT. The 23 diagnosis codes in the list account for 60% of all transfers. For each of the frequent diagnosis codes, the table shows the frequency of transferring inside the system (in-system) versus outside the system (out-of-system), as well as the frequency of clinical transfers versus non-clinical transfers. Note that the percentages of in-system and out-of-system transfers do not sum to 1 ‘because there are cases where the sender does not belong to any hospital system. Table 4 (see Appendix) shows frequencies of age group, patient insurance type, hospital ownership type, teaching status, and bed size, for each group of destination choice, as well as for each IHT type.

Variables

Transfer routinization is a construct that reflects the repetition and stability of the process. We operationalize routinization in IHT by measuring a sender’s repetition of destination relationships across quarters of the year (DEST_OVERLAP). For example, a high routinization case is when the sender keeps sending a certain types of patients (as measured by MDC, Major Diagnostic Category) to the same receiving hospital from one quarter to another. On the contrary, a lower routinization case is when the sender does not use the same receiving hospitals across the two consecutive quarters.

We measure care quality by 30-day hospital-wise readmission and in-hospital mortality. A readmission occurrence is identified if a patient is admitted to any hospital within 30 days of discharge from the previous hospital. We measure care efficiency by LOS. Considering that patient LOS at EDs is normally within one day, we consider hourly length of stay (HLOS) as the dependent variable. An initial check of the data shows that HLOS follows an approximate exponential distribution. The log transformation of HLOS (LNHLOS) approximates a normal

distribution. Therefore, we use LNHLOS as the dependent variable in our model. Table 5 and 6 (see Appendix) summarize the statistics of the dependent variables.
We control for a variety of patient-level and hospital-level variables. Table 7 summarizes the control variables. Total charge is controlled for as it indicates the amount and complexity of the care delivered to the patient. Geographic distance is controlled because transferring a patient over a long distance increases risks for the patient, especially for critically ill patients (Duke & Green, 2001; Bosk et al., 2011). Furthermore, in our analysis of destination choice (H3a - H3d), we also require that the destinations be within 75 miles of geographic distance, in order to minimize the sample selection bias.

Table 7 List of Control Variables

Variable Type	Variable Name
Patient demographics	Female (1 is female, 0 is male)
	Age
	Race
Conditions and care	Principle diagnosed conditions
	Number of diagnoses,
	Number of chronic conditions
	Number of comorbidities
	Number of procedures
	Number of current procedural terminology
Visit information	Log of total charge
	Primary payer/insurance
Hospital variables	number of beds
	ownership type
	teaching status
	hospital referral regions code
	hospital-wide readmission rate
Other variables	geographic distance

Econometric Models

We use the following linear model to model the relationship between transfer type and routinization.

$$DEST\_OVERLAP = \beta_0 + \beta_1 \times TransferType + \beta_2 \times HospitalVariables \quad (1)$$

We use generalized linear models for the hypothesized relationships between transfer related variables and care performance. The following shows a general structure of the models.

$$Care\_Outcomes = \beta_0 + \beta_1 \times TransferRelatedVariables + \beta_2 \times PatientDemographics + \beta_3 \times ConditionsAndCare + \beta_4 \times VisitInformation + \beta_5 \times HospitalVariables + \beta_6 \times OtherVariables \quad (2)$$

Readmission and mortality performance is only evaluated for the receiver's section of a transfer case since the sender has transferred out the patient to the receiver, thus does not have any readmission and mortality directly tied to it. LOS performance is evaluated for both the sender and the receiver.

RESULTS

The result for transfer routinization is shown in Table 8. The odds ratio of 1.416 for the IHT type factor means that clinical transfers have 41.6% higher chance of choosing the same receiving hospitals for transfer destinations from quarter to quarter than non-clinical transfers. The difference is statistically significant. Therefore, H1 is supported. For clinical transfers, the senders tend to have a stable set of receiving hospitals. Non-clinical transfers, however, have a changing set of destination hospitals from time to time. To the receiver, a non-clinical transfer relationship with the sender is temporary because the sender may not make the same transfer decision for the next patient.

Table 8 Destination Quarterly Overlap

Factors	Odds Ratio	95% Wald Interval	
		Lower	Upper
IHT Type: Clinical vs Non-clinical	1.416	1.264	1.586
N	7,086		

Control variables: Hospital variables.

Table 9 shows the result of the effect of transfer types on readmission and mortality. From the table, clinical transfers 16.4% smaller chance of incurring readmissions than non-clinical transfers. However, the difference's significant level is at the borderline. The table also shows that clinical transfers have 21.3% lower chance of mortality than non-clinical transfers, and the difference is statistically significant. The result suggests that overall, clinical transfers have better care quality than non-clinical transfers, supporting H2a.

Table 9 IHT Type on Readmission and Mortality

Factors	Readmission			Mortality		
	Odds Ratio	95% Wald Interval		Odds Ratio	95% Wald Interval	
		Lower	Upper		Lower	Upper
Clinical vs Non-clinical	0.836	0.675	1.037	0.787	0.632	0.980
N		11,000			11,619	

Control variables: Patient demographics, Conditions and care, Visit information, Hospital variables.

Column 1 and 3 of Table 10 correspond to the results of the linear regression for the sender's LOS and the receiver's LOS. As the table shows, clinical transfers are negatively associated with LOS for both the sending and the receiving hospitals, supporting Hypothesis 2b. On average, transferred patients for clinical reasons stay 5.2% (0.35 hour, or 21 minutes) shorter in EDs than patients transferred for non-clinical reasons. Even though the saved time is not much, considering that the average ED stay of nonclinical transferred patients is only about 5 hours and process efficiency is critical to emergent patients, the saved 21 minutes LOS has practical significance. At the receiving hospital, clinical transfers incurs 3.5% (about 3.35 hours) less LOS than non-clinical transfers, whose average LOS is 136 hours. That is an approximate saving of 3.4 hours.

Table 10 IHT Type and Destination Choice on LOS

	Ln (Sender's HLOS)		Ln (Receiver's HLOS)	
	1	2	3	4
Clinical transfer	−0.052 (0.013)	−0.057 (0.022)***	−0.035 (0.012)***	−0.055 (0.018)***
Clinical * in-system transfer		−0.039 (0.022)*		−0.001 (0.018)
Non-clinical * in-system transfer		−0.057 (0.023)**		−0.016 (0.020)
N	9,320	7,352	8,866	6,967
$R2$	0.236	0.234	0.725	0.735

Control variables: Patient demographics, Conditions and care, Visit information, Hospital variables.

*p <0.1, **p <0.05, ***p <0.01, ****p <0.001 (Standard errors in parentheses)

The results for H3a and H3c are shown in Table 11 - 12. For either clinical or non-clinical transfers, choosing in-system does not affect patients' readmission probability. Table 12 shows the results for mortality. While choosing in-system has no effect on mortality probability for non-clinical transfers, it does have a positive effect for clinical-transfers. Overall, the results imply that non-clinical transfers' care quality does not benefit from nor is it hurt by choosing in-system destinations, while clinical transfers' care quality (mortality) suffer by choosing in-system destinations. Therefore, neither H3a nor H3c is supported.
The results for H3a and H3c may be explained by an alternative view. While we argue that transferring within the same system may help improving care quality, it may also reduce the probability of a patient getting the best care quality. In Florida, on average, a multihospital system has 4 to 5 hospitals, while outside of the system there are over 200 hospitals. Even considering a certain driving distance (e.g., 75 miles), there are still many times more hospitals outside of the system than inside of it. Therefore, the chance that the best quality hospital exists inside the system is much lower versus outside of the system. Furthermore, clinical transfers may suffers more from choosing in-system destinations since the receiver's clinical competence is more of a concern for a clinical transfer, than for a non-clinical transfer. This view is consistent with that of Lu and Lu (2018), who suggest that transfers out of system have higher chance of landing better clinical quality than those in the same system.

Table 11 Destination Choice on Readmission for Each IHT Type

Factors	Readmission (Non-clinical)			Readmission (Clinical)		
	Odds Ratio	95% Wald Interval		Odds Ratio	95% Wald Interval	
		Lower	Upper		Lower	Upper
In-system vs out-of-system	0.876	0.533	1.440	0.925	0.581	1.473
N		3,038			3,929	

Control variables: Patient demographics, Conditions and care, Visit information, Hospital variables.

Table 12 Destination Choice on Mortality for Each IHT Type

Factors	Mortality (Non-clinical)			Mortality (Clinical)		
	Odds Ratio	95% Wald Interval		Odds Ratio	95% Wald Interval	
		Lower	Upper		Lower	Upper
In-system vs out-of-system	1.263	0.723	2.209	1.627	1.022	2.591
N		3,196			4,131	

Control variables: Patient demographics, Conditions and care, Visit information, Hospital variables.

Column 2 and 4 of Table 10 shows the results for H3a and H3d. As Column 2 shows, for the sender: choosing in-system destinations saves about 5.7% (about 18 minutes) for non-clinical transfers (significant at $p < 0.05$), while it saves only 3.9% for clinical transfers (significant at $p < 0.10$). For the receiver, choosing in-system destinations does not affect LOS for either non-clinical or clinical transfers. Considering both the sender and the receiver, the result supports H3b and H3d.
The results for H3a to H3d unveils an interesting situation for a hospital's transfer destination choice strategy. Considering both care quality and care efficiency, in-system is a better destination for non-clinical transfers since it is associated with better care efficiency while maintains the same care quality. Out-of-system is a better destination for clinical transfers, since it is associated with better mortality and similar care efficiency. The results imply that non-clinical transfers may benefit from the coordination mechanisms of same-system relationships without missing the possibly better quality of out-of-system hospitals. On the contrary, clinical transfers may miss the higher chance of better quality possibly existing in out-of-system hospitals while not benefiting much from the ease of internal coordination, because preplanning and routinization have already helped cultivate the formal and informal coordination mechanisms that are independent of system coordination mechanisms, as discussed in the Hypothesis Development section.

Robustness Check

To check the robustness of the results, we focus on addressing endogeneity issue, which is a common concern for studies using observational data. In our study, the issue is addressed both in concept and with propensity score matching techniques. Concept wise, endogeneity is less of a concern in our study for three reasons. First, the HCUP and AHA are regarded by many researchers as among the best quality and most reliable archival data sources for healthcare operations research, and measurement errors in the data are not much concerned. Second, simultaneous causality between IHT type and is of little concern as well. Clinical transfers are decided by a sending hospital's service scope. The receiver's care performance is unlikely to reversely affect clinical transfer decisions. The receiver's care performance may reversely affect nonclinical transfer decisions. However, such effects can only make our conclusion stronger because positive receiver's performance attracts more non-clinical transfers and therefore, the observed readmission, mortality and LOS measures for non-clinical transfers are overestimated. Correcting the overestimation only strengthens our conclusion that clinical transfers have better care quality and care efficiency than non-clinical transfers. Third, our model controls for both patient and hospital level variables. The chance of missing non-ignorable variables is small.

Next, we discuss the process and results of using propensity score matching techniques to address two possible endogeneity issues.

The Varied Effect of IHT Types

Patients who are directly admitted could have been transferred for non-clinical reasons (Note that the case is not valid for clinical transfers). We use one to one matching method to identify the direct admissions that are equally probable to be non-clinically transferred. Based on a 1% propensity score radius, we obtain 3,377 pairs, and compare the matched direct admissions with both clinical and non-clinical transfers. The result shows that non-clinical transfers have 22.1% more LOS than direct admissions ($p < 0.001$), while clinical transfers do not have significant different LOS than direct admissions. The result for readmission and mortality is not significant, possibly due to limited pairs matched. Overall, the result supports our previous findings that clinical transfers have better care performance than non-clinical transfers. It further shows that non-clinical transfer's care efficiency is worse than direct admissions while clinical transfer's efficiency is similar to direct admissions. This finding would not have been possible without differentiating the two IHT types. Existing literature's contradictory findings about the effect of IHT on LOS (e.g., Nathens et al. (2003) and Hill et al. (2011)) could have been due to mixing the two IHT types.

The Effect of System-based Destination Choices

Choosing in-system versus out-of-system a transfer destination may be an endogenous decision: the care performance associated with an in-system destination choice may be affected by the transfer case itself instead of the destination choice. To solve the issue, we use a one-to-one propensity score matching method to match the transfers within the same system with those outside of the system. Based on 1% propensity score radius, we obtain 446 pairs for the clinical transfer group, and 314 pairs for the non-clinical transfer group. The result for readmission and mortality is not significant, probably due to the limited pairs matched. The result for LOS, though, is consistent with our original findings. As Table 13 shows, for non-clinical transfers, in-system destinations are better as it saves 16.8% of the sender's LOS, and 22.7% of the receiver's LOS. For clinical transfers, the benefit of in-system destinations is not statistically significant.

Table 13 IHT Type's Destination Choice on LOS

	LNHLOS (Non-clinical)		LNHLOS (Clinical)	
	Sender	**Receiver**	**Sender**	**Receiver**
In-system transfer	−0.168 (0.071)**	−0.227 (0.104)**	−0.038 (0.062)	−0.041 (0.088)
Propensity Matched Pairs	314	314	446	446

*p <0.1, **p <0.05, ***p <0.01, ****p <0.001 (Standard errors in parentheses)

DISCUSSION AND CONCLUSIONS

EDs are getting more crowded due to the increasing level of diagnostic and treatment intensity (e.g., more clinical tests) for many maladies and the use of advanced medical technology (e.g., advanced imaging, particularly CT scanning) (Pitts, Pines, Handrigan, & Kellerman, 2012). Between 1997 and 2007, the increase in total annual ED visits almost doubled compared to what would be expected from population growth (Tang, Stein, Hsia, Maselli, & Gonzales, 2010). Transferring patients to another hospital may not only provide timely care for patients with time-sensitive conditions, but also helps alleviate hospital congestion (Iwashyna, Christie, Kahn, & Asch, 2009a; Russell et al., 2015). Moreover, hospital nursing resources in the US are in short demand (Rosseter, 2014). Transferring patients to hospitals where they can be cared for more efficiently may save overall nursing resources.

Contributions

IHT management is associated with challenging communication and information sharing that frequently breaks down, causing adverse clinical outcomes and wasting care resources. Understanding the characteristics and the impact of IHT on care efficiency may help hospitals make better transfer decisions. Our study investigates this subject using data from all of the hospitals within the state of Florida and considering all patient conditions. This is in stark contrast to existing studies that use a single hospital setting or consider a single disease. Our study makes two main contributions.

First, our study delineates the fundamentally different characteristics between clinical and nonclinical transfers, use OIPT to analyze the different coordination mechanisms adopted in each IHT type, and technically differentiates the two IHT types using empirical data. We show that clinical transfers are more routinized than non-clinical transfers. Very few studies have conceptually differentiated between the two IHT types, and none have technically separated them. Our empirical approach rigorously differentiates between clinical transfers and non-clinical transfer and the method can be used by future studies.

Second, we demonstrate the varied impacts of the two IHT types on care quality and care efficiency. Specifically, non-clinical transfers are associated with worse care quality and care efficiency than clinical transfers. Further, compared with direct admissions, non-clinical transfers have clearly worse care efficiency, while clinical transfers have not much different care efficiency. The varied impacts imply that coordination mechanisms (which are abundant in clinical transfers but lack in non-clinical transfers) are critical to meeting information processing demand in IHT.

Transferring within the same healthcare system has been shown to overlook care quality for certain diseases when IHT types are not differentiated (Lu & Lu, 2018). Our findings complement this view. We find Pareto sets of destination choices for the two IHT types. When both care quality and care efficiency are considered, in-system destinations are better for non-clinical transfers and our-of-system destinations are better for clinical transfers.

Managerial Implications

Hospital staffs generally view IHT as a burden because of the process complexity and the additional coordination work involved (Bosk et al., 2011). In many situations, the feeling of burdensomeness originates from the lack of understanding of the characteristics and impacts of IHT, as well as evidence-based destination choice guidelines. Our research provides new specific suggestions to IHT operations. Knowing that non-clinical transfers are less routinized, lack coordination, and incur worse care quality and care efficiency than clinical transfers, hospitals management may implement more coordination mechanisms, e.g., standardized processes, strategic planning, and long-term stable partnership with receivers, to improve the care performance. Hospitals can even consider reducing the volume of non-clinical transfers though better planning and resource allocation.

Hospitals prefer to transfer within healthcare systems for the convenience of communication and coordination as well as for the system's financial interest. This practice may be improved.

Knowing the transfer-type-specific impacts of system-based destination choices, hospitals may form different transfer destination guidelines for the two IHT types. For non-clinical transfers, hospitals may keep their current same-system preference for the communication and coordination advantage. For clinical transfers, hospitals may choose out-of-system destinations for better care quality.

Limitations and Future Research Directions

We take an evidence-based approach to differentiate IHT types, i.e., the classification completely base on the historical diagnosis and procedure records in discharge data. While this approach is objective, in some scenarios, it may not completely reflect the actual transfer situations. Moreover, our study is in the setting of Florida and the results may not hold in states that have drastically different healthcare policies than that of Florida. Future research may verify our results in settings of other states.

Our IHT type classification method may be used by future IHT-related researches. For example, studies predicting patient transfers may consider that in some cases (clinical transfers) hospitals don't have the option not to transfer while in other cases (non-clinical transfers) have that option. Studies of care regionalization may need to filter non-clinical transfers when examining the effect of regionalization on care quality or care efficiency.

Most of the extant studies of the IHT destination choice problem focus on the sending hospitals' decision, however, a transfer arrangement is a result of negotiation between the sending and the receiving hospitals. The receiving hospital faces the problem of how to prioritize bed assignment between different IHT types as well as between IHT and direct admissions. Furthermore, with the implementation of the readmission penalty by CMS, hospitals probably don't want to expose themselves to penalty risks and need to make decisions of considering the different odds of readmission between clinical and non-clinical transfers.

Information sharing and task coordination is the key to IHT management. Hospitals' adoption of electronic health records (EHR) systems and their interoperability affects IHT management (Bosk et al., 2011; Usher, Sahni, Herrigel, Simon, Melton, Joseph, & Olson, 2018). Future research may explore the effect of hospitals' use of EHR on the coordination mechanisms existing in different transfer types.

APPENDIX - TABLES

Table 3 Most Frequent Principle Diagnosis Codes

Code	Description	Count	In-system	Out-of-system	Clinical	Non-clinical
109	Acute cerebrovascular disease	994	43.46%	38.03%	58.35%	41.65%
100	Acute myocardial infarction	931	54.99%	25.24%	84.43%	15.57%
2	Septicemia (except in labor)	596	34.56%	42.28%	64.09%	35.91%
233	Intracranial injury	414	40.82%	44.93%	70.05%	29.95%
237	Complication of device; implant or graft	363	25.34%	58.95%	66.39%	33.61%
142	Appendicitis and other appendiceal conditions	317	38.49%	48.58%	16.72%	83.28%

Differentiating Interhospital Transfer Types

226	Fracture of neck of femur (hip)	316	27.53%	43.99%	47.15%	52.85%
238	Complications of surgical procedures or medical care	301	25.91%	57.48%	47.84%	52.16%
229	Fracture of upper limb	268	32.46%	53.36%	62.69%	37.31%
230	Fracture of lower limb	266	27.82%	47.37%	50.38%	49.62%
197	Skin and subcutaneous tissue infections	253	40.32%	43.08%	41.11%	58.89%
231	Other fractures	250	23.20%	63.20%	66.80%	33.20%
131	Respiratory failure; insufficiency; arrest (adult)	238	21.01%	47.48%	65.97%	34.03%
153	Gastrointestinal hemorrhage	225	39.11%	45.33%	49.78%	50.22%
149	Biliary tract disease	219	28.77%	46.58%	55.71%	44.29%
106	Cardiac dysrhythmias	208	29.81%	37.50%	67.79%	32.21%
228	Skull and face fractures	189	38.62%	46.03%	89.42%	10.58%
108	Congestive heart failure; nonhypertensive	177	27.68%	42.37%	61.58%	38.42%
240	Burns	175	16.57%	67.43%	72.57%	27.43%
101	Coronary atherosclerosis and other heart disease	166	36.75%	34.34%	77.71%	22.29%
160	Calculus of urinary tract	155	52.26%	27.74%	29.68%	70.32%
115	Aortic; peripheral; and visceral artery aneurysms	152	46.71%	38.16%	78.95%	21.05%
42	Secondary malignancies	142	35.92%	42.96%	68.31%	31.69%

Table 4 Statistics for Each Group of Destination Choice / IHT Type

Category	In-system	Out-of-system	Clinical	Non-clinical
Age Group				
0 to 10	2.77%	6.19%	3.27%	6.18%
11 to 20	4.34%	7.30%	4.05%	8.21%
21 to 30	7.06%	7.55%	8.03%	5.98%
31 to 40	7.36%	7.76%	7.74%	7.06%
41 to 50	11.37%	10.45%	11.43%	9.72%
51 to 60	18.15%	16.39%	18.54%	15.52%
61 to 70	18.96%	16.73%	19.12%	16.26%
> 70	30.00%	27.64%	27.80%	31.08%
Insurance Type				
Medicaid	16.40%	20.99%	17.94%	20.00%

Differentiating Interhospital Transfer Types

Medicare	48.50%	45.85%	48.01%	48.17%
No charge	3.16%	1.15%	2.02%	1.69%
Other	3.44%	4.36%	3.94%	3.50%
Private insurance	19.35%	18.07%	18.32%	17.88%
Self-pay	9.16%	9.59%	9.77%	8.75%
Hospital Ownership				
Governmental	9.57%	7.05%	10.90%	8.32%
Not for profit	28.99%	69.71%	55.96%	48.78%
Private	61.44%	23.25%	33.14%	42.91%
Hospital Teaching Status				
Non-teaching	28.46%	10.90%	19.89%	19.95%
Teaching	71.54%	89.10%	80.11%	80.05%
Hospital Bed Size				
Large (over 80 percentile)	40.84%	71.65%	62.73%	56.26%
Medium (50 to 80 percentile)	49.12%	21.94%	28.90%	34.72%
Small (under 50 percentile)	10.03%	6.41%	8.37%	9.02%

Table 5 Statistics of 30-Day Readmissions and Mortalities

Total Admissions	**Readmissions**	**Readmission rate**	**Mortalities**	**Mortality rate**
755,867	55,045	*7.28%*	26,531	*3.51%*

Table 6 Statistics of LOS

LOS Categories	**Percent**
ED	
0 - 1 hour	22.42%
1 - 2 hours	23.74%
2 - 3 hours	18.61%
3 - 4 hours	12.23%
4 - 5 hours	7.12%
5 - 10 hours	8.43%
> 10 hours	7.45%
Inpatient	
0 - 1 day	7.86%

1 - 2 days	18.16%
2 - 3 days	19.00%
3 - 4 days	14.61%
4 - 5 days	10.17%
5 - 10 days	21.06%
> 10 days	9.14%

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Flexible Staffing of Retail Sales and E-Fulfillment Associates to Maximize Profits

Fred Easton (†)
Formerly of Syracuse University
†Deceased June 29, 2018

Gerard Campbell
Western Washington University
Email: Gerard.Campbell@wwu.edu

ABSTRACT

Some brick-and-mortar stores utilize existing sales associates to fill online orders during slow periods, while others rely on dedicated staff for each operation. We formulate a profit-maximizing mathematical program to evaluate alternative staffing strategies for retailers with in-store fulfillment operations. In experiments simulating different customer traffic patterns, traffic variability, and forecast accuracy, we measure expected profits, the flows of flexible staff between sales and fulfillment, and the sensitivity of those metrics to different market and employee characteristics. We find that fulfillment operations with as few as one cross-trained employee are more profitable than those with dedicated sales and fulfillment staff.

KEYWORDS: Retailing, Workforce staffing, Shift scheduling, Mathematical programming

INTRODUCTION

By Q2, 2017, online purchases represented about 8.9% of all U.S. retail sales and were increasing at an annual rate of 16.2%; much faster than overall retail sales growth rate of 4.1% (US Dept Commerce, 2017). Recognizing opportunities for growth, nearly half of the traditional retailers responding to a recent survey were planning or had already implemented an omni-channel distribution strategy (Forrester Research, 2014). An event study by Xia & Zhang (2010) found that publicly-traded retailers with an omni-channel presence experienced improved sales, expenses, and ROI. Besides the top-line benefits from increased geographic reach, assortment, and 24-7 availability, omni-channel strategies reduce costs and improve ROA through increased buying power with suppliers, lower transaction costs, and smaller inventories. To achieve these synergies, however, the marketing and operational strategies for in-store and online channels must be carefully integrated (Bendoly, Blocher et al., 2007; Zhang, Farris et al., 2010).

Order fulfillment remains one of the most costly processes for online retail enterprises (de Koster, 2002). The key elements of an online order fulfillment strategy include inventory location (centralized, decentralized, or drop-shipped directly from the manufacturer or wholesaler), warehouse organization (integrated or segregated picking locations for store replenishment and online orders), physical distribution (allocation of on-line orders to fulfillment sites, picking and packing, delivery destinations, transportation modes, and pick-up or delivery scheduling), and merchandise returns. Because the relevant information is easily digitized, some or all parts of the online order fulfillment process can be readily outsourced to third-party vendors (Perdikaki, Peng & Heim, 2015). However, many prominent U.S. omni-channel merchants (Best Buy, Dick's Sporting Goods, Lowe's, Nordstrom's, and Wal-Mart, among others) fulfill online orders from their own stores – for two important reasons (Rueter, 2014; Sheldon et al., 2014). First, a retail store is often closer to the buyer than any distribution center, so a ship from store strategy may reduce shipping time and shipping costs. Second, the real-time inventory visibility needed to support a ship from store strategy enables virtual pooling of retail demand across all store locations and the online channel (Mahar, Bretthauer & Venkataramanan, 2009), reducing chain-wide markdowns and stockouts due to forecast errors.

Countering these benefits, retailers pursuing in-store order fulfillment strategies must invest in staff training and provide enough space, equipment, packaging and staff time to support fast error-free picking and packing (Cunnane, 2015). Some ship from store retailers rely on their sales associates to fill online orders during slow periods (Fowler & Dodes, 2010; Reuter, 2014). However, diverting associates from the selling floor can imperil in-store sales (Fisher, Krishnan & Netessine, 2009; Ryan, 2014) and may adversely affect productivity through changeover losses whenever workers transition between sales and fulfillment tasks (Nembhard, 2014). To avoid such losses, other stores employ dedicated associates to pick and pack online orders (Partridge, 2013). However, the work rules governing employee schedules and the inherent variability of online order volume make it difficult to achieve a perfect match between the quantity of scheduled labor and the amount of labor needed to fill online orders, resulting in costly under- or over-staffing (Mani, Kesavan & Swaminathan, 2015).

While Quan (2004) and Lang (2010) summarized many of the issues related to retail labor staffing and scheduling, few analytical models have been developed for such decisions (Van den Bergh et al., 2013). Traditionally, retail staffing and scheduling decisions begin with demand planning (Netessine, Fisher & Krishnan, 2010), or forecasts of hourly (or more frequent) store sales. Those forecasts are then translated to the labor requirements for each planning interval, for up to a few weeks into the future, that drive workforce scheduling decisions. However, existing retail staffing and scheduling models may be ill-suited for retail operations with in-store fulfillment because they do not allow for the transfer of cross-trained workers between sales and e-fulfillment areas and fail to recognize heterogeneous task priorities.

To more accurately evaluate staffing alternatives for in store fulfillment operations, we propose a contribution-oriented workforce staffing and scheduling model for retail operations that allows for two demand streams – in-store customers and online orders. The sales floor and e-fulfillment areas each have their own skill requirements, and in-store customers have a higher priority because sales may be lost if an insufficient number of sales associates are available when the customers are present. Our model allows modest levels of cross-training and more fully exploits task-switching flexibility and the priority differences between store traffic and online orders than extant retail staffing and scheduling models. To demonstrate the benefits that flexible staffing can provide to retailers with instore fulfillment operations, we applied it in a full-factorial numerical experiment designed to compare the performance of alternate staffing strategies under uncertain in-store traffic.

The remainder of our paper is organized as follows. In the next section, we review the relevant literature. Then, we present our mathematical programming models and solution methods. Starting with a two-stage stochastic program for a single-period, we propose and evaluate a heuristic solution strategy based on expected traffic. We then extend that deterministic approach to a multi-period version of the problem. After that, we describe a numerical experiment based on that model and its results. In the final section, we summarize our findings, point out the limitations of our study, and suggest directions for future research.

LITERATURE REVIEW

Labor-related expenses such as hiring, training, and wage costs account for 10 – 20% of a retailer's costs, typically the largest expense after cost of goods sold (Nettesine et al., 2010).

Thus, a retailer's profitability is significantly affected by the quality of its staffing decisions. For this literature review, we first classify and review retail staffing strategies and the timetabling decisions needed to ensure sufficient capacity for in-store fulfillment of online orders. Then, we take a closer look at the sales response functions incorporated into current contribution-oriented workforce staffing and scheduling models for retail workers.

Retail Workforce Staffing and Scheduling Decisions

Fulfillment operations tend to be labor intensive, and handling costs tend to be the largest single expense for most fulfillment operations. Largely due to space and equipment limitations, in-store fulfillment operations may have higher labor requirements and labor costs per order than centralized fulfillment (Bendoly et al., 2007; Bretthauer et al., 2010). It is also likely that those costs will vary with the types of employees charged with order fulfillment. For example, some retailers like Saks staff their in-store fulfillment operation with dedicated specialists who focus on picking and packing online orders (Partridge, 2013), while other omni-channel retailers ask their sales associates to fill online orders during slow periods on the floor (Fowler & Dodes, 2010; Rueter, 2014). The latter approach combines online and in-store workload, offering the potential for productivity gains through reduced idle times (Schultz, et al., 2003). However, those gains may be offset by the cycle of learning, forgetting, and relearning following an assignment change. For example, Shultz et al. (2003) found that even short-term interruptions due to task switching adversely affects the productivity of cross-trained operators. Nembhard (2014) provides an excellent summary of the research pertaining to the productivity of cross-trained workers.

Staffing decisions for instore fulfillment operations can impact both labor expenses and sales. Although store labor budgets and staffing levels for traditional retailers are often based on monthly sales forecasts (Fisher et al., 2009), empirical studies (Lam, Vandenbosch & Pearce, 1998; Fisher et al., 2006; and Perdikaki, Kesavan & Swaminathan, 2012) find that sales vary with store traffic and staffing levels. Improving the match between store capacity and customer traffic can boost store profitability by 3% or more (Netessine et al., 2010; Chapados et al., 2011). However, retail stores usually operate beyond the normal 8 AM – 5 PM work day and often experience significant hour-to-hour and day-to-day variability in store traffic. In addition, certain times of the year are much busier than others. Due to restrictions on work hours, meal breaks and rest periods, state and federal statutes, and local custom, it can be difficult to match scheduled labor with forecasted demand. Furthermore, work schedules for full- and part-time retail employees are often developed well in advance of the realization of actual demand, so aligning capacity with actual demand may require real-time adjustments in employee work hours and/or task assignments.

Although the literature related to workforce staffing and scheduling is extensive (Van den Bergh, et al. 2013; De Breucker et al., 2015), few existing models are applicable to the retail setting assumed for this research (Mou et al., 2018). Among its distinguishing features, this setting includes multiple demand streams with different skill requirements and priorities, serviced by a mix of specialized and cross-trained employees, where workforce management decisions influence both income and expenses. The key workforce management decisions in such an environment include staffing decisions (S), training decisions (X), scheduling decisions (shift (SS), day-off (DO), or tour (TS) scheduling), and flexible worker assignment (A) decisions (Abernathy et al., 1973). In Table 1 we summarize the key decisions, the primary components of the objective function, and some of the key constraints for a selection of workforce staffing and scheduling models applicable to the assumed operating environment. All the models in Table 1 integrate two or more of these key decisions and can accommodate multiple demand streams, each with its own skill requirements. In the discussion below, we focus on the key similarities and differences between the models' objectives and constraints.

Most of the models in Table 1 track wage costs, but only the recent model by Cuevas et al. (2016), which views labor costs as sunk in the short run, addresses the diminishing impact of additional staff on revenue. Easton (2011) and Henao et al. (2015) incorporate quasi-fixed costs such as benefit and training costs that vary with headcount rather than hours worked. Cuevas et al. (2016) also model task switching costs as linear functions of the duration of the assigned task, incurred whenever a flexible employee is transferred from one task assignment to another. Other models (e.g., Brusco & Johns, 1998; Campbell, 2011; Easton, 2011**)** approximate this loss with reduced productivity when flexible workers are assigned outside their home area. However, Berman & Larson (2004) suggest switching costs in retail services arise from a brief period of non-productive "re-orientation time" by the transferred worker and vary with the number of transfers.

While all of the above models assume multiple demand streams with different skill requirements, where cross-trained employees can service more than one type of demand, none of them appear to recognize priority differences among the different demand streams Billionnet (1999), Campbell (2011), Henao et al. (2015) and Cuevas et al. (2016) confine cross-trained workers to the same task assignment for their entire shift, potentially limiting the value of cross-training. Loucks & Jacobs (1991), Brusco & Johns1998), Billionnet (1999), Cai & Li (2000), Bard (2004), and Easton (2011) require solutions that meet or exceed the staffing goals for each task in each period, and do not allow shortages. In contrast, Campbell (2011) and Henao et al. (2015) both allow under-staffed solutions, while Restrepo et al. (2017) consider target staffing levels as goals and penalize solutions (biased in favor of overstaffing) that deviate in either direction from staffing targets for each task and period. Campbell (2011) imposes a quadratic penalty on any shortage, while Henao et al. (2015) and Restrepo et al. (2017) rely on linear penalties. Loucks & Jacobs (1991) penalize surplus staff, while Cai & Li (2000) propose a secondary objective that encourages its even distribution. Cuevas et al. (2016) do not specify minimum staffing levels but impose upper bounds on the number of workers assigned to a task in each period.

Cross-training flexibility has been shown to mitigate supply – demand mismatches arising from forecast errors, unplanned absences, and employee turnover (Easton, 2011). When cross-trained sales associates also pick and pack online orders, the in-store and online demand streams are effectively pooled, reducing relative demand variability. This should allow retailers to achieve better store performance than a system staffed by the same number of specialists. However, any of the multiskill models in Table 1 assume deterministic demand, a shortcoming previously noted in the extensive review by De Breucker et al. (2015). Campbell (2011) and Restrepo et al (2017) address uncertainty through two-stage stochastic programs, establishing time-tabling decisions for employees in stage 1, postponing recourse decisions such as task and rest-break assignments in stage 2 after actual demand is revealed.

Finally, well-known service quality instruments like SERVQUAL (Parasuraman et al., 1988) have long assigned a prominent role to "responsiveness" in the assessment of customer satisfaction. By prioritizing in-store customers over online order fulfillment, increased service levels may result in increased sales and revenue. While events such as the arrival of a customer at the store may require a near-immediate response, order fulfillment tasks are often interruptible as well as postponable (Berman & Larson, 2004). An online order could remain queued without ill effect provided it is processed some time before the latest outbound parcel pickup of the day. However, none of the models in Table 1 acknowledge priority differences among the multiple demand streams. The distinct temporal priorities for each demand stream, as well as the flexibility to divert online orders to other fulfillment centers when necessary, could also temper the effects of positively correlated demand streams on cross-training performance.

Thus, these limitations may cause existing retail staffing and scheduling models to undervalue the benefits of cross-training flexibility.

Table 1: Key decision variables, objective function components, and constraints for selected integrated staffing, scheduling and allocation models for multi-skill workforce

Multi-skill Workforce Mgt Model	Decis-ions	Objective Function Terms					Constraints				
		sales/ utility/ surplus values	Short-age costs	Lab-or cost	Trai-ning costs	Task swit-ching costs	Task switch-ing fre-quen-cy	Exogenous staffing goals with: Sur-plus vars.	Exogenous staffing goals with: Short-age vars.	Tas k prio r-ities	Un-cer-tain-ties
Loucks & Jacobs, 1991	TS, A	surplus penalty	No	No	No	No	≤ Hr	Yes	No	No	Det
Brusco & Johns,1998	SS,A	No	No	Yes	No	Yes	≤ Hr	No	No	No	Det
Billionnet, 1999	DO,A	No	No	Yes	No	No	Shift	No	No	No	Det
Cai & Li, 2000	S,TS, A	Even distrib.	No	Yes	No	No	≤ Hr	Yes	No	No	Det
Bard, 2004	TS, A	No	No	Yes	No	No	≤ Hr	No	No	No	Det
Campbell, 2011	DO,A	No	Qua-dratic	Yes	No	Yes	Shift	No	Yes	No	labor reqs.
Easton, 2011	S,X, TS,A	No	No	Yes	Yes	Yes	≤ Hr	No	No	No	labor reqs.
Henao et al., 2015	S,X, TS,A	No	Linear	Yes	Yes	No	Shift	No	Yes	No	Det
Cuevas et al., 2016	TS,A	trun-cated	No	sunk	No	Yes	Shift	No	No	No	Det

Sales response functions

Queueing theory suggests that decreasing the number of sales associates on the floor will increase waiting time and the likelihood that impatient customers in the store will leave without buying. Grewal, Baker et al. (2003) confirmed the importance of perceived waiting, finding that once in the store, customer purchase intentions are influenced by their expectations about waiting. More store employees visible on the sales floor tends to produce favorable wait expectations, while increased customer density on the sales floor results in more negative expectations. Fisher & Raman (2010) explained that retail staffing levels influence store sales in two ways: directly through the level of sales assistance available to shoppers, and indirectly through execution of store operational activities such as stocking shelves, tagging merchandise, and maintaining the overall store ambience. Netessine et al. (2010) found the practice of matching sales associates to store traffic rather than sales forecasts resulted in larger mean purchases. Perdikaki et al. (2012) confirmed that retail store sales tend to increase with store traffic, but at a decreasing rate that is moderated by sales associate staffing levels.

Lam et al. (1998) were among the first to propose a two-input sales response function conditioned on store traffic (λ) and the sales associate staffing level (W_h) during time h, where:

$$S_{Wh} = \alpha\, [\lambda_h^{\beta}]\, [e^{-\Upsilon/Wh}]. \tag{1}$$

In their model, S_{Wh} represents sales in period h conditioned on store traffic and staffing levels. Parameter α represents the "sales potential," or mean purchase amount for customers who make a purchase. The bracketed expression following α is the number of shoppers in the store who are "buyers." Reasoning that sales should increase with traffic at a decreasing rate, Lam et al. (1998) confined traffic elasticity β to the interval $0 < \beta < 1$. The parameter Υ in the bracketed expression on the right is usually defined as the responsive-ness of sales to labor. With $\Upsilon > 0$ and staffing level $W_h \geq 1$, the expression $\exp(-\Upsilon/W_h)$ can be interpreted as the fraction of "patient" store traffic willing to remain in the store and complete their purchases, which decreases in Υ at a rate moderated by W_h.

In their empirical study, Mani et al. (2015) found that the fitted values for traffic elasticity β for an apparel retailer tended to be higher during weekdays (when the ratio of traffic to staffing is lower) and lower on the weekends (when the ratio of traffic to staffing is larger). Asserting that traffic elasticity β should depend on both traffic and staffing levels, Chuang et al. (2016) proposed alternate ways to model and estimate sales response functions. However, both measuring and predicting store traffic remain challenging (Chapados et al., 2014; Song, 2015), particularly at the resolutions required for operational decisions such as staffing and scheduling. Complicating factors include nested seasonality and the impact of traffic drivers such as pricing, promotion, product mix, competitive actions, and political, economic, meteorological and cultural events.

MODELS, SOLUTIONS, AND BOUNDS

Our literature review suggests that omni-channel retailers planning instore fulfillment operations may find existing retail staffing and scheduling models inadequate, for the following reasons:

- Most contribution-oriented retail staffing and scheduling models tend to ignore differences in the skills needed to respond to different requirements, and they ignore priority differences between in-store traffic and postponable, interruptible online orders.
- Despite a well-established relationship between store traffic, staffing levels and store sales, most cost-oriented retail workforce management models ignore the diminishing marginal revenue produced by additional staff and, in some cases, assess penalties if scheduled staffing exceeds exogenous targets.

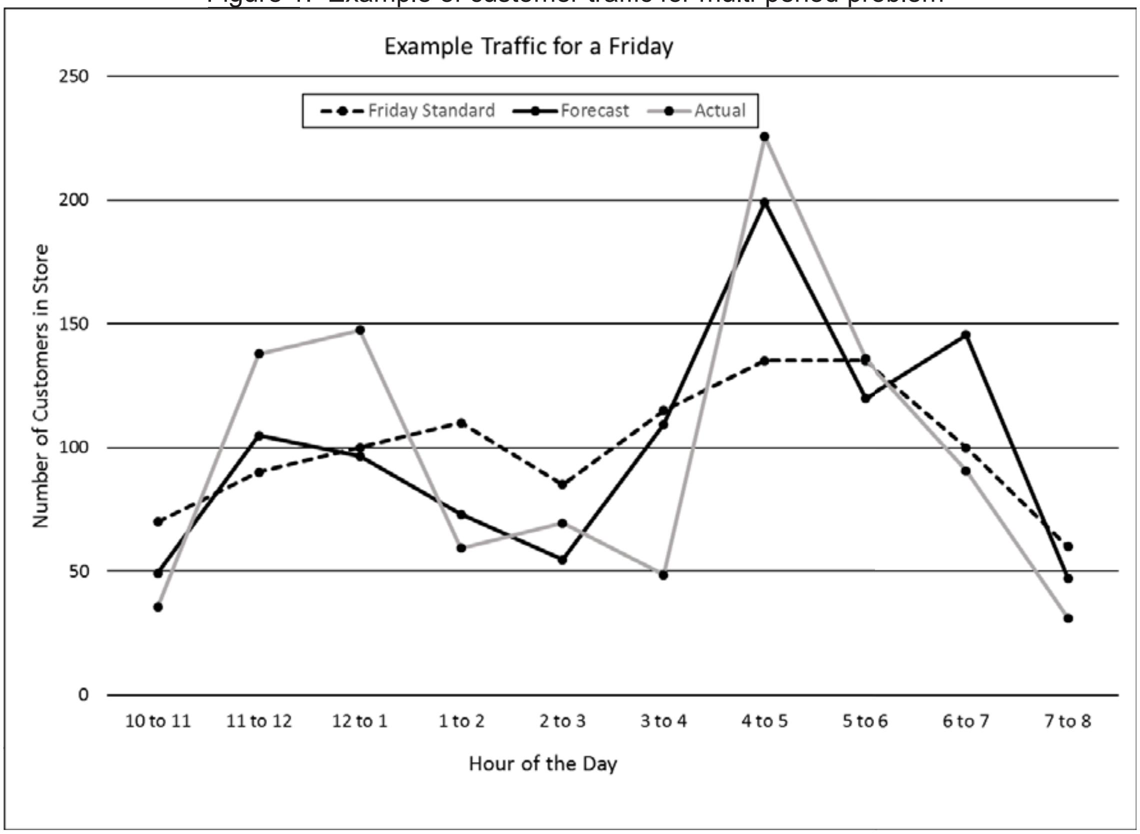
This study begins to address these limitations by developing models for use in investigating alternative staffing strategies (dedicated versus multi-skilled workers) for retailers with in-store customers and online orders.

Overview of models and solution methods

We assume the retailer services two distinct demand streams: assisting in-store customers and fulfilling online orders. An example of in-store customer traffic for the multi-period decision problem appears in Figure 1. Hourly store traffic varies by time of day, day of the week, and season. The three hourly traffic curves shown in Figure 1 represent mean, forecasted, and actual traffic for a Friday. When actual traffic is less than forecasted, we assume that any on-duty sales associate can be temporarily assigned to help fulfill online orders and contribute to fulfillment capacity, albeit less proficiently than a fulfillment specialist. We also assume there are a limited number of in-store fulfillment associates trained to assist on the sales floor. At any hour, store managers can review current actual traffic and decide whether sales associates can profitably be assigned to fulfillment tasks, or whether any cross-trained fulfillment associates can be profitably diverted to the sales floor.

The complexities of the stochastic, multi-period environment shown in Figure 1 make it difficult to obtain optimal solutions to problems of realistic size, even for versions that do not allow for cross-utilization. For purposes of evaluating the value of cross-utilization, which is the main focus of our study, we develop heuristic solution procedures based on deterministic versions of the multi-period problems. To demonstrate the effectiveness of the heuristics, we compare their performance against optimal solutions based on single-period versions of the stochastic problems, obtained using Monte Carlo methods.

Figure 1: Example of customer traffic for multi-period problem



We begin by presenting two single-period stochastic staffing models. The first is similar to the single-period contribution-oriented staffing model for sales associates described in Lam et al. (1998) and Henao et al. (2016). The second model has two stages: stage 1 staffing, which must be completed long before actual traffic is known; and stage 2 recourse, which allows flexible sales and fulfillment staff to be reassigned at the beginning of a period based on realized customer traffic. We use the two-stage single-period stochastic model to define upper bounds and demonstrate the performance of a heuristic solution approach based on a deterministic version of the problem, which is then extended to a multi-period environment.

Contribution-oriented stochastic program for single-period staffing.

In seeking to maximize profit contributions for a single period, equation (1) can be used within an objective function. W, the number of workers, is an integer decision variable, for which we set lower and upper limits of L and M, respectively. At the time when staffing is performed, customer traffic is a random variable, $\boldsymbol{\lambda}$, with mean $\hat{\lambda}$. A single-period model based on Lam et al. (1998) can be stated as follows:

$$\underset{W}{\text{Maximize}} \{g \ \alpha\boldsymbol{\lambda}^{\beta}e^{-\Upsilon/W} - \ cW\} \tag{2}$$

$$\text{Subject to:} \quad L \leq W \leq M \text{ and integer,} \tag{3}$$

where g is the gross profit margin (e.g., 51% as in Lam et al., 1998) and c is the periodic wage rate for the sales associates. Definitions for the sales response function parameters α, λ, β, and γ were introduced previously following equation (1).

Lam et al. (1998) present optimality conditions derived from mean forecasted store traffic and forecast error variance. For our study, we adopt a heuristic approach by solving a deterministic version of the problem that substitutes $\hat{\lambda}$ for $\boldsymbol{\lambda}$:

$$\underset{W}{\text{Maximize}} \{g \ \alpha\hat{\lambda}^{\beta}e^{-\Upsilon/W} - \ cW\} \tag{2'}$$

$$\text{Subject to:} \quad L \leq W \leq M \text{ and integer} \tag{3}$$

For the single-period deterministic version of the problem, an optimal solution W* is easily obtained through enumeration. As a heuristic solution strategy, substituting the unbiased estimator $\hat{\lambda}$ for the random variable $\boldsymbol{\lambda}$ is both computationally efficient and practical, especially when forecast error variance cannot be reliably estimated. Furthermore, as Chapados et al. (2014) noted, the contribution function for profit-oriented retail scheduling decisions tends to be fairly flat near its maxima. Using Monte Carlo simulations with 100 replications of actual customer traffic for each of 60 different hourly customer traffic distribution parameter/sales model parameter combinations (described in section 4), the optimal solutions to the model given by (2') - (3) resulted in staffing levels with the highest profits for 52 of the 60 cases. For the 8 cases where heuristic solutions resulted in lower profits based on the Monte Carlo sampling, the average difference in profits was 0.12%, and the maximum difference was 0.29%. This supports the use of the deterministic approach based on mean forecasted customer traffic $\hat{\lambda}$ for generating solutions to single period problems that do not allow for cross-utilization. The next section builds on this result to model situations where reallocation of flexible workers is permitted at the beginning of the period based on actual customer traffic.

Stochastic program with recourse for staffing and scheduling for a single period.

Extending the single-period staffing model for sales associates in (2) – (3), we next consider a two-stage, single-period model with multi-skilled sales associates and a limited number of cross-trained fulfillment staff. Using forecasted traffic, the Stage 1 model determines the ideal staffing level for sales associates, each of whom is also trained to perform fulfillment tasks when necessary. The Stage 2 recourse model allows appropriately-trained workers to shift between the sales floor and the warehouse in response to differences between predicted and actual customer traffic. A summary of model parameters and decision variables is shown in Table 2.

Table 2: Model parameters and decision variables

Model Parameters:	
L, M	lower, upper bounds for number of sales associates on duty. Sales associates are indexed k = 1, …, M.
$\boldsymbol{\lambda}$, $\hat{\lambda}$, λ	random, mean forecasted, or actual customer traffic during the hour.
c	sales associate hourly wage rate.
b	fulfillment associate hourly wage rate, where $b \leq c$.
g	gross profit margin (e.g., 51% as in Lam et al., 1998).
P_f	relative proficiency of workers transferred *from* sales to fulfillment operation.
P_t	relative proficiency of workers transferred *to* sales from fulfillment operation.
U	maximum number of workers that can be moved to the sales floor each hour.
Stage 1 decision variable:	
W	number of sales associates scheduled in advanced for the hour.
Stage 2 recourse variables:	
T	number of fulfillment associates moving to sales floor at the beginning of the hour.
F	number of sales associates moving from sales floor at the beginning of the hour.

Single-Period Stage 1 Model: Contribution-oriented sales associate staffing

$$\underset{W}{\text{Maximize}} \ \{E_{\boldsymbol{\lambda}}\,[Q(W,\boldsymbol{\lambda})] - cW\} \tag{4}$$

$$\text{Subject to:} \quad L \leq W \leq M \text{ and integer} \tag{3}$$

Note that $E_{\boldsymbol{\lambda}}[\,]$ denotes the mathematical expectation with respect to the random variable $\boldsymbol{\lambda}$.

Single-Period Stage 2 Model: Cross-utilization recourse after actual customer traffic is known

$$Q(W,\lambda) = \underset{F,T}{\text{Maximize}} \ \{g\,\alpha\,\lambda^{\beta}\,e^{-\Upsilon/(W+P_tT-F)} + P_f\,F\,b - T\,b\} \tag{5}$$

$$\text{Subject to:} \quad T \leq U \tag{6}$$

$$F \leq (W - L) \tag{7}$$

$$F, T \in \{\text{non-negative integers}\} \tag{8}$$

The problem shown above is a two-stage stochastic program with integer decision variables (see Birge & Louveaux, 2011). The solution to the stage 1 model establishes the staffing levels for retail sales associates (W) for the period, which constrain the subsequent stage 2 assignment decisions (T and F, the recourse variables) that are made at the beginning of the period. At that time, sensors reveal actual store traffic, λ, and management decides how to best deploy their on-duty multi-skilled staff.

When T fulfillment associates with sales proficiency P_t are reassigned to the sales floor to assist the W sales associates already on duty, we expect a proportionate increase in sales. However, we assume that the capacity diverted from the in-store fulfillment operation must be replaced, possibly through assistance from sales associates at other times of the day, or by either increasing local fulfillment staffing levels at cost Tb or diverting some local fulfillment load to another fulfillment center in the chain. Similarly, when F sales associates are temporarily reassigned to fulfillment duties, store sales will decrease. However, the shift also reduces labor expenses in the warehouse by P_fFb, where P_f is the average relative proficiency of sales associates assigned to fulfillment duties.

Constraint (6) limits the number of cross-trained fulfillment associates that can be deployed to the sales floor (T) to the number trained for such work (U). Constraint (7) ensures that at least L sales associates will remain on the sales floor at all times. Due to the restricted range of values for integer variables T and F, stage 2 optimal solutions can be obtained through simple enumeration. The objective function ensures that at most one of T and F will be non-zero, because with both P_t and P_f both < 1, it would never be profitable to transfer workers both to and from the sales floor in the same period.

To optimally solve stage 1 of the stochastic programming problem, we require additional information about the distribution of λ. Barring that, however, we can still obtain an upper bound for its solution. For upper bounds, we define the following:

Z_{PI} -- the optimal stage 1 profit with perfect store traffic information.

Z_{ub} -- an upper bound corresponding to the optimal profit when stage 2 is solved with:

1) no constraint on U; 2) the wage premium (c-b) charged only for worker-hours used for transfers to the sales floor; and 3) $P_f = P_t = 1.0$.

For the single-period problem, $Z_{ub} = Z_{PI}$. However, because work rules such as such as the minimum or maximum shift length tie the periods together, this may not be the case for multi-period problems.

To establish a heuristic solution, the following stage 1 problem based on mean forecasted customer traffic ($\hat{\lambda}$) can be solved:

$$\underset{W}{\text{Maximize}} \{g\ \alpha\hat{\lambda}^{\beta}e^{-\Upsilon/W} - cW\} \tag{2'}$$

$$\text{Subject to:} \qquad L \leq W \leq M \text{ and integer} \tag{3}$$

This formulation is the same deterministic model presented in section 3.2. Let W* be the optimal solution for (2') – (3), easily obtained through enumeration. Once actual traffic λ, is known, the following represents the profit that would be realized if no cross-utilization were used:

$$\text{Profit without cross-utilization} = Z_{w/o} = \{g\ \alpha\lambda^{\beta}e^{-\Upsilon/W*} - cW^*\} \tag{9}$$

With W* as a starting point, the stage 2 problem can be solved to obtain the following:

$$\text{Profit with cross-utilization} = Z_{xu} = Z_2{*} - cW^* - U(c-b), \tag{10}$$

where $Z_2{*} = Q(W^*, \lambda)$ and sales associate wages cW^* and wage premiums (c-b), if any, paid to the U cross trained fulfillment workers on duty (which are paid regardless of where they are assigned to work).

<u>Performance measures:</u>

Let **\$xu** be the mean value for (Z_{xu} - $Z_{w/o}$), representing the average value of cross-training flexibility under uncertain demand, and **%xu** be the percentage (**\$xu**/$Z_{w/o}$). Let **\$pi (%pi)** represent the mean (percentage) difference between Z_{PI} and $Z_{w/o}$, a measure that captures the impact of demand uncertainty without cross-training flexibility. Finally, define **\$ub** (**%ub**) as the mean (percentage) difference between Z_{ub} and $Z_{w/o}$, a measure of the potential improvement in profits attainable with completely accurate forecasts, complete scheduling flexibility, unconstrained redeployment of cross-trained staff, and flexible workers with task proficiencies comparable to dedicated workers. In our computational study, we also report the mean flows from and to the sales floor (consequence variables **Fxu**, **Txu** from the stage 2 cross-utilization solution) and compare them to **Fub and Tub**. Although these performance measures have been defined based on the single-period models and solutions, they also apply for multi-period cases, which are described next.

Deterministic staffing and scheduling models for multiple periods.

The multi-period stage 1 problem builds on (2') by including constraints for allowable schedule patterns such as the minimum and maximum allowable shifts, allowable start times, etc. We add the index h, representing one of the H periods in the planning horizon, to model parameters and variables $\hat{\lambda}$, λ, W, T, and F (defined in Table 2). Table 3 defines additional notation used with the multi-period model.

Table 3: Parameters and decision variables for multi-period formulation

Additional Model Parameters:	
H	number of hours in planning horizon. Hours are indexed h = 1,…H.
N	number of different shift patterns. Shift patterns are indexed j = 1,…,N.
a_{hj}	1 if hour h is a work period for shift pattern j, 0 otherwise; ∀ h, j.
$\hat{s}_{hk}$, s_{hk}	mean forecasted marginal sales, actual marginal sales (based on eq. 1) produced during hour h by k^{th} sales associate based on forecasted, actual traffic; ∀ h, k.
C_j	wage cost per sales associate assigned to shift pattern j.
Additional Stage 1 Decision Variables:	
X_j	number (integer) of sales associates who work shift pattern j, j=1,…,N.
Y_{hk}	binary variable for k^{th} on-duty sales associate in hour h, equal to 1 if $\sum_{j=1}^{N} a_{hj}X_j \geq k$, 0 otherwise.

<u>Multi-Period Stage 1 Model: Deterministic contribution-oriented sales associate staffing</u>

$$\underset{\forall X,Y}{\text{Maximize}} \quad Z_1 = g\sum_{h=1}^{H}\sum_{k=1}^{M} Y_{hk}\hat{s}_{hk} - \sum_{j=1}^{N} C_jX_j \tag{11}$$

Subject to:

Relate sales associates on duty during period h (W_h) to number assigned to shift pattern j (X_j):

$$\sum_{j=1}^{N} a_{hj}X_j - W_h = 0; \; h=1, \ldots, H \tag{12}$$

Fix minimum and maximum staffing level for sales floor during operating hours:

$$L \leq W_h \leq M; \quad h=1, \ldots, H \tag{13}$$

Count on-duty sales associates to compute revenue for each hour:

$$W_h - \sum_{k=1}^{M} Y_{hk} \geq 0; \; h=1, \ldots, H \tag{14}$$

$$Y_{hk} - Y_{h,k-1} \leq 0; \; h=1, \ldots, H \text{ and } k=1, \ldots, M. \tag{15}$$

Other constraints: $X_j \geq 0$ and integer $\forall$ j; $Y_{hk} \in \{0,1)$ $\forall$ h,k (16)

Period h Stage 2 Model: Cross-utilization recourse after actual customer traffic is known in hr. h:

$$\underset{F_h, T_h}{\text{Maximize}} \; Z_{2h} = \left[g\left[\sum_{k=1}^{W_h-F_h} s_{hk} + P_t \sum_{k=W_h+1}^{W_h+T_h} s_{hk}\right] + P_f F_h b - T_h b\right] \quad (17)$$

Subject to:

$$T_h \leq U \quad (18)$$

$$F_h \leq (W_h - L) \quad (19)$$

$$F_h, T_h \in \{\text{non-negative integers}\} \quad (20)$$

For the multi-period stage 1 problems in this paper's computational study (see section 4), shift patterns (a_{hj}) are defined based on four-hour shifts, which is consistent with Lam et al. (1998), Mani et al. (2015), and Quan (2004). With multi-hour shifts in the formulation, the stage 1 multi-period problem is more complex than the deterministic single-period problem given by (2') - (3) and requires integer programming solution methods. For moderate problem instances, such as the stage 1 problems modeled in our computational study (7 general integer and 400 binary integer variables), optimal solutions for the stage 1 model given by (11)- (16) can be readily obtained with commercial optimization software such as the IBM ILOG CPLEX (2017).

The hourly stage 2 recourse problems are always feasible, and optimal solutions (Z_{2h}*) can be obtained through simple enumeration over the restricted range of values for integer recourse variables T_h and F_h (the objective function ensures that at most one will be non-zero). The hourly problems are independent of each other because the number of sales associates scheduled for period h (W_h) is fixed when stage 1 is solved – in practice, this corresponds to shift schedules for sales associates that are established in advance based on forecasted customer traffic.

For multi-period problems, the following apply:

$$\text{Profit with cross-utilization} = Z_{xu} = \sum_{h=1}^{H}[Z_{2h}^* - cW_h - U(c-b)] \quad (21)$$

$$\text{Profit without cross-utilization} = Z_{w/o} = g\sum_{h=1}^{H}(\alpha \lambda_h^{\beta} e^{-\Upsilon/W_h}) - \sum_{j=1}^{N} C_j X_j \quad (22)$$

Also, with multi-hour shifts, the following relationship between bounds and heuristic solutions apply:

$$Z_{ub} \geq Z_{PI} \geq Z_{w/o} \quad (23)$$

Where:

Z_{PI} -- the optimal stage 1 profit with perfect store traffic information; and

Z_{ub} -- an upper bound corresponding to the optimal profit when stage 2 is solved with: 1) no constraint on U; 2) the wage premium (c-b) charged only for worker-hours used for transfers to the sales floor; and 3) $P_f = P_t = 1.0$.

Performance measures for multi-period problems are the same as those defined earlier for the single-period problem. These performance measures are summarized in Table 5, presented in the next section.

NUMERICAL EXPERIMENTS

To evaluate the benefits of flexible workers, we simulated the operation of a hypothetical omni-channel retailer that operates for seven days/week, 10 hours per day. Assuming part-time sales associates who work 4-hour shifts/day, we established bounds on profitability and measured changes in profitability and the flows of flexible workers between the sales floor and the fulfillment operation when a few fulfillment associates are trained to assist on the sales floor.

In the US, mean hourly wages for retail employees and for warehousing & storage workers are nearly the same (BLS, 2015), so we fixed the wages for both sales and fulfillment associates at $12/hr. After Schultz et al. (2003), we assumed that cross-trained associates are less proficient than dedicated employees when working outside their primary area. In our main numerical experiment, we fixed P_f and P_t at 80%, and limited the number of cross-trained fulfillment associates to U = 1. However, the value of flexibility depends on both the number and the relative proficiency of cross-trained associates, so we performed additional sensitivity analysis with values of P_f and P_t ranging from 0.70 to 1.0 and for $1 \leq U \leq 3$.

To facilitate replication, we adopted the sales response model S (eq. 1) and gross profit margin (51%) from Lam et al. (1998). Due to the heterogeneity of the different retail enterprises they examined, Lam et al. (1998), Kabak et al. (2008), and Mani et al. (2015) found quite different values for the sales response model coefficients used in their studies. Viewing the staffing level W_h as endogenous, we used their fitted values to establish ranges for the experimental sales response model coefficients α (average purchase amount), Υ (related to customer patience), and β (traffic elasticity) used in our study. For our main numerical experiment, we fixed the sales response model coefficients at their medium values. To isolate the influence of these parameters on results, however, we also performed additional sensitivity analysis with these parameters at their low, medium, and high values. Those values are shown in Table 4.

Table 4: Experimental parameters for store traffic and sales response function

<table>
<tr><td rowspan="4">Mean Traffic $\bar{\lambda}_h$ by hour and day</td><td>hour (h)</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr>
<tr><td>Mon-Thr</td><td>45</td><td>55</td><td>60</td><td>65</td><td>50</td><td>70</td><td>80</td><td>80</td><td>60</td><td>35</td></tr>
<tr><td>Fr</td><td>70</td><td>90</td><td>100</td><td>110</td><td>85</td><td>115</td><td>135</td><td>135</td><td>100</td><td>60</td></tr>
<tr><td>Sa-Su</td><td>80</td><td>100</td><td>120</td><td>130</td><td>120</td><td>130</td><td>110</td><td>90</td><td>70</td><td>50</td></tr>
<tr><td colspan="5"></td><td colspan="7">Day of week</td></tr>
<tr><td colspan="2" rowspan="5">Hypothetical Sales Response Parameters</td><td colspan="3" rowspan="2">Level</td><td colspan="3">Mon – Thr</td><td rowspan="5"></td><td colspan="3">Fr – Su</td></tr>
<tr><td>α</td><td>β</td><td>Υ</td><td>α</td><td>β</td><td>Υ</td></tr>
<tr><td colspan="3">Low</td><td>18</td><td>0.60</td><td>2.09</td><td>30</td><td>0.60</td><td>4.00</td></tr>
<tr><td colspan="3">Med</td><td>27</td><td>0.75</td><td>2.75</td><td>45</td><td>0.75</td><td>5.25</td></tr>
<tr><td colspan="3">High</td><td>36</td><td>0.9</td><td>3.40</td><td>60</td><td>0.9</td><td>6.50</td></tr>
</table>

Finally, we combined store traffic data from Mani et al. (2015), Lam et al. (1999), and other publicly available retail traffic studies to create the average store traffic patterns $\bar{\lambda}_h$ shown in Table 4. We assumed that weekday (Mon – Thr) traffic averaged 600 customers per day, while on weekends (Fri-Sun), mean traffic per day increased to 1,000. Because the benefits of flexibility tend to increase with increasing variability (Iravani et al., 2011, among others) we simulated traffic forecasts $\hat{\lambda}_h$ for each period by sampling from the normal distribution $N(\bar{\lambda}_h, CV\bar{\lambda}_h)$. Consistent with Thompson (1992) and Mani et al. (2015), we used three different CV levels (CV = 0.2, 0.4, or 0.6) to generate the forecasts. The sample values were then normalized to ensure that the average daily demand for the test problems corresponded to the assumed values (e.g., weekdays = 600/day; weekends = 1000/day), with negative hourly traffic forecasts truncated to zero. Actual hourly traffic values λ_h were simulated by adding a random forecast error term ε_h to the forecast (e.g., actual hourly traffic $\lambda_h = \hat{\lambda}_h + \varepsilon_h$). The error terms were generated by sampling from a normal distribution with mean 0 and a standard deviation corresponding to one of the three levels of forecast accuracy: % mean absolute deviation (%MAD) = 12%, 24%, or 36%. The 12% and 24% values are similar those reported in Mani et al. (2015), and the 36% level represents higher forecast errors.

Main Experiment Results

Our main numerical experiment examined the impact of store traffic on performance: specifically, the mean hourly traffic for each day, the coefficient of variation of forecasted traffic, and forecast error. Fixing the sales response parameters (α, Υ, and β) at their medium values (see Table 4), with $P_f = P_t = 0.8$ and U = 1, we used a full factorial design with three levels each for store traffic (Mon-Thr, Friday, and Sa-Su), CV, and %MAD factors, with five replications for each of the 27 experimental cells. For each test problem, we solved with and without cross-utilization to obtain the four objective measures described in the previous section ($Z_{w/o}$, Z_{xu}, Z_{ub}, and Z_{PI}). From them, we computed ten absolute and relative measures to help establish the value of cross-training flexibility. Table 5 shows the overall mean values and standard errors for those measures.

Compared with dedicated sales and fulfillment associates, cross-utilization helped increase expected daily profits by an average of **$xu** = $ 39.51 (**%xu** = 1.81%). In contrast, perfect information about customer traffic lifted daily profits for dedicated staffing strategies by only **$pi** = $18.11 (**%pi** = 0.96%) -- roughly half that achieved with flexible workers -- supporting previous experimental findings that cross- training flexibility may be more valuable than perfect information (Campbell, 2011). The upper bound results of $95.88 (4.16%), when compared against $**xu** and %**xu**, provide an indication of the losses associated with the 80% worker efficiencies and limitation on U assumed for the cross-utilization solutions. These losses are evaluated more closely in the sensitivity analysis presented in a later section.

Table 5: Results summary, main numerical experiment

Performance Measure	**Grand Mean**	**Standard Error**
$ profit increase with cross-utilization (**$xu)** = ($Z_{xu}$ - $Z_{w/o}$)	$39.51	$4.66
$ profit increase with perfect information (**$pi)** = ($Z_{PI}$ - $Z_{w/o}$)	$18.11	$2.16
Upper bound for $ profit increase (**$ub) = (**$Z_{ub}$ - $Z_{w/o}$)	$95.88	$7.92
% profit increase w/cross-utilization (**%xu) =** [(Z_{xu} - $Z_{w/o}$)/ $Z_{w/o}$]*100	1.81%	0.24%
% profit increase w/perfect information (**%pi)** = [(Z_{PI} - $Z_{w/o}$)/ $Z_{w/o}$]*100	0.96%	0.17%
Upper bound on % profit increase (**%ub)** = [(Z_{ub} - $Z_{w/o}$)/ $Z_{w/o}$]*100	4.16%	0.35%
Mean workers/hour from sales floor F_h | P_f = 0.8 (**F_{xu}**)	0.86	0.07
Mean workers/hour to sales floor T_h | P_t = 0.8 (**T_{xu}**)	0.19	0.01
Mean u. bound, workers/hr from sales floor F_h | P_f=1.0, U = ∞ (**F_{ub}**)	1.43	0.08
Mean u. bound, workers/hr to sales floor T_h | P_t=1.0, U = ∞ (**T_{ub}**)	1.09	0.05

Table 5 also summarizes the mean number of flexible workers moving from and to the sales floor. For solutions with cross-utilization and 80% proficiency, $\mathbf{F_{xu}}$ and $\mathbf{T_{xu}}$ averaged 0.86 and 0.19, respectively. For Z_{ub}, the upper bound solutions, cross-trained worker proficiencies equal their dedicated colleagues (i.e., $P_f = P_t = 1$) and U is relaxed to permit unrestricted movement to the sales floor. The upper bound solution values $\mathbf{F_{ub}}$ = 1.43 and $\mathbf{T_{ub}}$ = 1.09 suggest that with increased cross-training and proficiency, movement to the sales floor is likely to increase at a faster rate than movement from the sales floor.

Table 6: ANOVA results for the main numerical experiment

Perf. Measure	ANOVA Output	Main Effects			Interaction Effects			
		Day	CV	MAD	Day * CV	Day * MAD	CV * MAD	Day * CV * MAD
$xu	*F-stat. (sig.)*	**6.3 (.00)**	**12.0 (.00)**	**6.6 (.00)**	0.9 (.48)	0.6 (.70)	1.1 (.34)	0.7 (.66)
	effect size (η^2)	**0.073**	**0.138**	**0.076**	0.020	0.013	0.026	0.034
$pi	*F-stat. (sig.)*	**6.0 (.00)**	0.1 (.92)	**30.7 (.00)**	0.2 (.91)	**3.6 (.01)**	0.5 (.70)	0.2 (.99)
	effect size (η^2)	**0.060**	0.001	**0.306**	0.005	**0.072**	0.011	0.008
$ub	*F-stat. (sig.)*	**15.2 (.00)**	**16.9 (.00)**	**10.0 (.00)**	1.5 (.22)	1.3 (.28)	1.3 (.29)	0.5 (.85)
	effect size (η^2)	**0.143**	**0.159**	**0.094**	0.027	0.024	0.024	0.019
%xu	*F-stat. (sig.)*	1.1 (.34)	**7.1 (.00)**	**6.5 (.00)**	0.2 (.96)	0.8 (.55)	0.8 (.54)	0.6 (.81)
	effect size (η^2)	0.015	**0.096**	**0.087**	0.004	0.021	0.021	0.030
%pi	*F-stat. (sig.)*	2.7 (.07)	0.3 (.77)	**11.4 (.00)**	0.4 (.80)	1.5 (.20)	0.3 (.88)	0.4 (.93)
	effect size (η^2)	0.036	0.004	**0.153**	0.011	0.041	0.008	0.020
%ub	*F-stat. (sig.)*	1.1 (.35)	**10.8 (.00)**	**8.9 (.00)**	0.2 (.96)	0.9 (.46)	0.9 (.46)	0.5 (.83)
	effect size (η^2)	0.013	**0.134**	**0.110**	0.004	0.023	0.023	0.026
Fxu	*F-stat. (sig.)*	**16.8 (.00)**	**14.4 (.00)**	**7.9 (.00)**	0.9 (.45)	0.6 (.69)	1.5 (.22)	0.9 (.52)
	effect size (η^2)	**0.164**	**0.140**	**0.077**	0.018	0.011	0.028	0.035
Txu	*F-stat. (sig.)*	**3.2 (.04)**	**6.4 (.00)**	**4.3 (.02)**	0.5 (.73)	0.7 (.58)	0.7 (.62)	1.5 (.15)
	effect size (η^2)	**0.042**	**0.083**	**0.055**	0.013	0.018	0.017	0.079
Fub	*F-stat. (sig.)*	**41.2 (.00)**	**12.1 (.00)**	**7.4 (.00)**	0.6 (.65)	0.8 (.50)	1.6 (.18)	0.7 (.70)
	effect size (η^2)	**0.334**	**0.098**	**0.060**	0.010	0.014	0.026	0.022
Tub	*F-stat. (sig.)*	**34.5 (.00)**	**6.3 (.00)**	**5.8 (.00)**	1.4 (.23)	1.8 (.13)	0.5 (.73)	1.0 (.47)
	effect size (η^2)	**0.308**	**0.057**	**0.052**	0.025	0.032	0.009	0.034

(Shading indicates F-statistics with better than 0.05 significance. Effect size (η^2) indicates the proportion of variance in the dependent variable explained by the independent variable.)

While the overall results presented in Table 5 provide a general indication of performance for the main numerical experiment, Table 6 summarizes the Analysis of Variance (ANOVA) results. The main experimental factors (traffic differences by day of week, forecasted traffic CV and MAD forecast error) exhibit a statistically significant relationship with almost all of the financial and flexible worker flow measures. The effect sizes (η^2), which reflect the number of cell replicates involved in their calculations, are consistent with the main and interaction effects. With one exception, however, interaction effects appear insignificant. Based on the ANOVA results, our discussion of the results will focus on the main effects of the experimental factors.

Effects of Daily Traffic Pattern (Day)

Table 7 presents the cell means for the 10 performance measures for each level of each of the three traffic factors we examined. The upper left portion of Table 7 shows the mean and percentage increase in profits for 80% cross-utilization (**$xu, %xu**), perfect information (**$pi, %pi**) and upper bounds (**$ub, %ub**) under each of the three daily traffic patterns. The 80% cross-utilization solutions and upper-bound solutions both exhibit larger absolute profit increases relative to solutions without cross-utilization on weekends (and especially Fridays). The solutions with perfect information, which do not benefit from cross-utilization, have much smaller improvements with all three traffic patterns. While mean percentage increases on weekends appear smaller, the differences are not statistically significant (see Table 6).

The lower left portion of Table 7 summarizes the mean number of workers moving from and to the sales floor each hour for the 80% cross-utilization ($\mathbf{F_{xu}}$, $\mathbf{T_{xu}}$) and the upper-bound ($\mathbf{F_{ub}}$, $\mathbf{T_{ub}}$) solutions. Average flow from the sales floor exceeds the mean flow to the sales floor, with the largest flows occurring on Fridays and the weekend. With increased proficiency and no restrictions on flows to the sales floor, the mean number of fulfillment associates assisting with sales ($\mathbf{T_{ub}}$) soared above $\mathbf{T_{xu}}$ by more than six-fold on Fridays and weekends. To justify such an increase, actual customer traffic would have to significantly exceed the forecasted level in some periods. Considered together, the daily traffic pattern results suggest that the value of cross-training flexibility tends to increase with store traffic.

Table 7: Performance measures by treatment levels of experimental factors

	Daily Traffic Pattern			Coeff. of Dem. Var.			Forecast Error Mag		
Perf. Measure	M-Th	Fri.	Sat.-Sun	0.2	0.4	0.6	12%	24%	36%
$xu	$19.90	$54.75	$43.89	$19.12	$32.67	$66.75	$20.46	$41.33	$56.75
$pi	$12.62	$26.63	$15.06	$17.07	$18.60	$18.64	$3.48	$14.16	$36.68
$ub	$47.24	$125.67	$114.72	$60.68	$80.65	$146.30	$58.36	$102.98	$126.29
%xu	2.10%	1.98%	1.35%	0.86%	1.65%	2.91%	0.87%	1.71%	2.84%
%pi	1.43%	0.96%	0.49%	0.80%	0.97%	1.10%	0.15%	0.69%	2.03%
%ub	4.64%	4.31%	3.52%	2.59%	3.71%	6.17%	2.47%	4.22%	5.78%
F_{xu}	0.41	1.10	1.07	0.56	0.76	1.26	0.56	0.95	1.07
T_{xu}	0.16	0.20	0.22	0.16	0.17	0.24	0.16	0.20	0.23
F_{ub}	0.70	1.81	1.77	1.14	1.34	1.81	1.13	1.53	1.63
T_{ub}	0.60	1.31	1.37	1.01	0.97	1.30	0.90	1.16	1.22

Effects of Traffic Variability (CV)

The middle three columns of Table 7 display mean values for the 10 performance measures with respect to the variability of the mean hourly arrival rates (CV). For 80% cross-utilization and upper bound solutions, the absolute and percentage profit difference measures increase at an increasing rate in CV. The effect is most pronounced for the upper-bound results. In contrast, the solutions with perfect information, which do not benefit from cross-utilization, appear relatively insensitive to CV. This observation, which is supported by the ANOVA results reported in Table 6, suggests that improvements may be due to cross-utilization flexibility. The lower half of the middle three columns of Table 7 shows average hourly flows of flexible workers to and from the sales floor for the 80% cross-utilization and upper bound solutions. In general, the flows from the sales floor increase with CV for both solutions. On the other hand, CV appears to have much less of an influence on flows to the sales floor.

Effects of Forecast Error Magnitude (%MAD)

The final three columns of Table 7 show the effects for forecast error magnitude %MAD on the absolute and percentage increases. The effects on the solutions with cross-utilization are similar to those induced by traffic variability, with larger forecast errors creating more profitable opportunities for flexible workers. The bottom half of the last three columns in Table 7 show the effects of forecast errors on the flow of flexible workers to and from the sales floor. The strength of the effect is as expected, with increased flows at higher levels of forecast error.

Sensitivity analysis

In our main numerical experiment, we focused on the influence of traffic and cross-training on profitability. Other potentially important environmental characteristics were held fixed, such as the assumed parameters for the sales response function (α, β, and Υ), the proficiency of cross-trained workers (P_f and P_t), and the number of cross-trained fulfillment associates allowed to assist on the sales floor (U). We now examine the sensitivity of the main experimental results to changes in those previously fixed environmental parameters.

Sensitivity to Sales Model Parameters

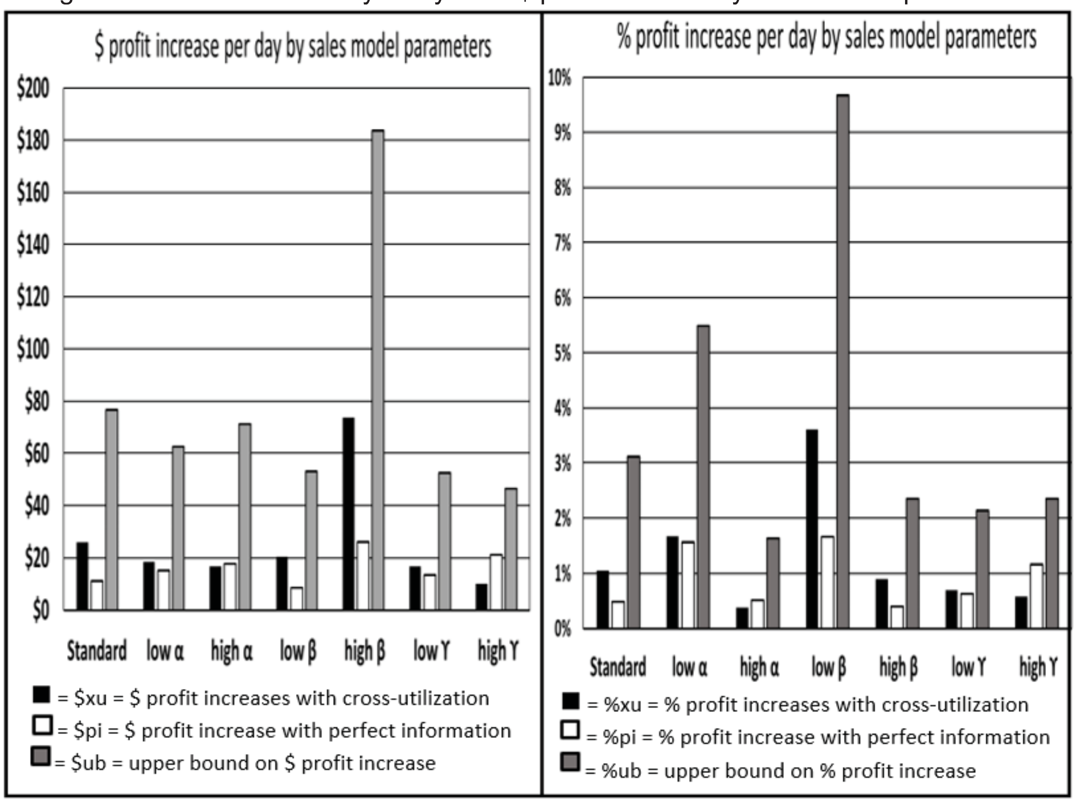
We first consider sensitivity to the assumed parameter values for the sales response function (eq. 1). Guided by the parameter values reported in Lam et al. (1998) and Mani et al. (2015), we established low, medium and high values for sales response function parameters α, β, and Υ. To assess the sensitivity of our results to changes in those parameters, we developed a compact experimental design to estimate the $ and % profit changes as the sales response parameters are varied from their low to medium to high settings while the other parameters remained fixed at their medium values (see Table 4).

With traffic variability CV = 0.4 and uncertainty %MAD = 24% (same as main numerical experiment) we first fixed one of the three sales response parameters at its "low" level and the other two parameters at their medium settings (as used with the main numerical experiment). Reusing the 5 replicates for each of the 3 daily traffic patterns generated for the main numerical experiment, we solved the problems with the new sales response parameters and obtained $Z_{w/o}$, Z_{xu}, Z_{PI} and Z_{ub} for all 15 problem instances. We then reset the test parameter to its high value and solved the problem sets again. Repeating this process for the second and third sales response model parameters, we solved a total of 90 additional problems to assess profit sensitivity to the three sales response parameters.

Figures 2A and 2B summarize differences between profit metrics **$xu/%xu**, **$pi/%pi**, and $**ub/%ub** as each of the sales response model parameters are varied. For comparative purposes, we also display the mean values from the main numerical experiment, hereafter referred to as the "standard" case. In both absolute and percentage terms, parameter β (traffic elasticity) appears to have a much greater influence on profit than either α or Υ. In Fig. 2A, increasing β from low to high more than triples all three absolute profit differences. As β approaches 1.0, customers become less sensitive to congestion and the proportion of buyers among the arrivals increases. With more willing buyers among the arriving customers, it is easier to justify increased staffing levels. However, **$pi** benefits the least from increasing β.

Even with accurate forecasts, scheduling constraints make it difficult to efficiently match staffing levels to demand with dedicated sales associates, especially at higher CVs. For the staffing options that allow cross-utilization, flexible workers can be shifted to and from the sales floor for short periods of time to achieve a better match between capacity and demands during busy periods. Those same workers can be deployed to assist with fulfillment during low traffic periods, reducing labor costs for the fulfillment operation. With the ability to perfectly match highly proficient staff to demand, **$ub** shows the largest absolute gains from increasing β. In percentage terms, however, profit improvement decreases with increasing β, as seen in Fig. 2B. Total profits are very sensitive to β, averaging \$629 at low β, \$2463 at standard β, and \$7216 at high β. These differences in total profit are much wider than the percentage difference in dollar profit increases seen in Fig. 2A, resulting in the reversal of the impact direction for β when going from a dollar basis to a percentage.

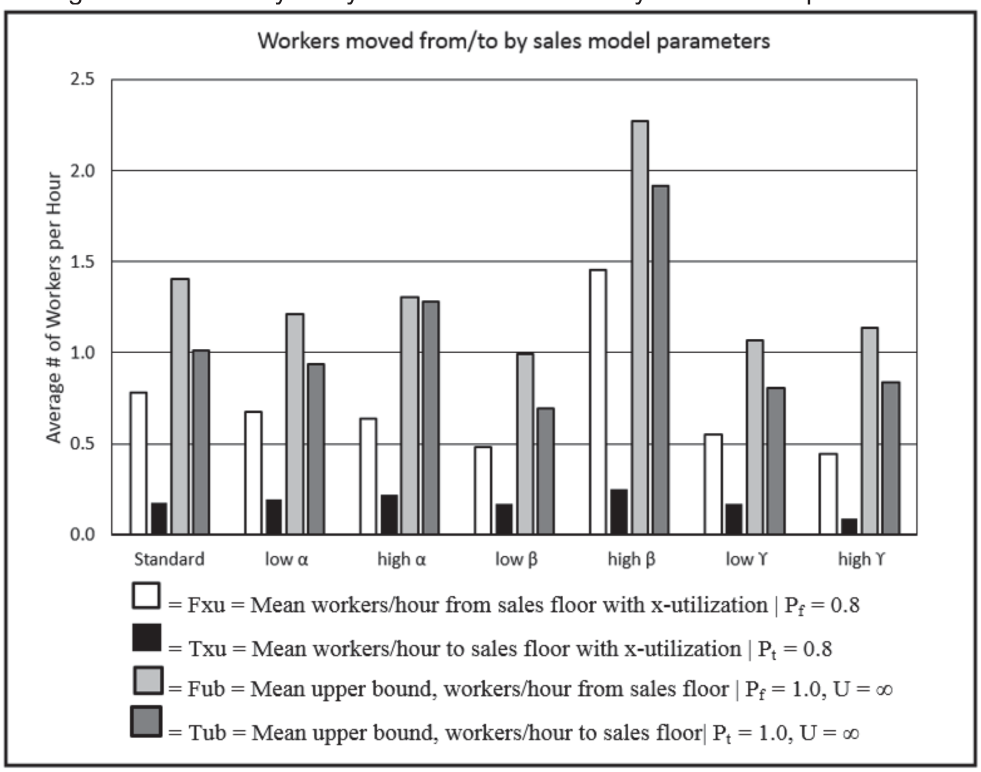
Figures 2A & 2B: Sensitivity analysis of $ profit increase by sales model parameters



With respect to sales potential α, the differences **$xu**, **$pi** and **$ub** in Figure 2A are surprisingly small as α is varied across its range. As α increases, the revenue per transaction increases while traffic remains the same. The increased margin supports a larger sales staff, which in turn captures more of the potential revenue. However, as more of the arriving customers receive prompt assistance, there are fewer opportunities to generate additional sales and it becomes more difficult to profitably add sales staff. The profit differences **$xu**, **$pi** and **$ub** are relatively unchanged as α increases, even though profits for the dedicated staffing strategy $Z_{w/o}$ steadily increase from \$1,256 when α is at its low setting to \$4136 at high α. This suggests that most of the benefits of cross-training, forecast accuracy, and scheduling flexibility are already present with α at its low setting. Thus, the percentage increases from those strategies, as shown in **%xu**, **%pi** and **%ub,** diminish as $Z_{w/o}$ increases with α. Finally, profit differences are less sensitive to sales response parameter Υ, which governs the sales responsiveness to store labor. For a given staffing level, the fraction of arrivals willing to wait for service decreases in Υ. As Υ increases, more staff are needed to meet customer wait time expectations, increasing labor costs and driving profits downward.

Figure 3 summarizes how the flow of cross-trained workers between sales and fulfillment operations varies with the sales model parameters. Here again, β appears to have the largest impact and Υ the smallest. Compared with the standard case, increasing β drives larger flows of flexible workers in both directions. When flows to the sales floor are limited to U=1, mean flows from the sales floor (**F_{xu}**) greatly exceed flows to the sales floor (**T_{xu}**). For the unconstrained upper bound solutions, with $P_t = P_f = 1.0$ and U unlimited, cross-utilization in both directions (**F_{ub}** and **T_{ub}**) increases with increasing β and notably, **T_{ub}** >> **T_{xu}**. The large differences**,** particularly at high values of β, hint at opportunities to improve profitability through investments in training and process improvements.

Figure 3: Sensitivity analysis of workers moved by sales model parameters



<u>Sensitivity to Cross-Utilization Efficiencies and Constraints</u>

To develop additional insights about the environmental conditions driving cross-utilization T_h and F_h, we examine how performance changes as the values of U, P_f, and P_t are varied while all other parameters remain at their medium levels. The magnitude of the changes in performance is based on the difference between the experimental treatment solution and the corresponding mean values for the standard problem set (i.e., CV = 0.4, %MAD = 24%, with the values of α, β and ϒ at their medium settings, for each of the three daily traffic patterns and 5 replications per cell.). To isolate the effects of P_f from P_t, we varied one while maintaining the other at 80%. Parameter U influences the flow of workers to the sales floor, and is therefore varied in conjunction with P_t. Figure 4A shows changes in **$xu** and **%xu**, relative to the standard solutions, as the proficiency of workers moved from sales (P_f) is varied from 70% to 100% in 5% increments. The steady increase in relative profitability in P_f is due solely to the increasingly efficient fulfillment assistance from sales associates. However, at P_f = 100%, the cross-utilization solutions must conform to U = 1 and P_t = 0.8. Therefore, the profit increase for cross-utilization solutions remains below the level of the upper bound solutions, where we assume P_f = P_t = 100% and U is unconstrained.

Figure 4B depicts the average number of sales associates/hour who assist with fulfillment operations over the range of P_f and largely explains the profit increases shown in Figure 4A. The flows from the sales floor are proportional to the profit increases, rising steadily from P_f = 75% to 100%. At P_f = 100%, the average number of workers moving from the sales floor for the cross-utilization solutions matches the flow of workers from the sales floor for the upper bound solutions.

Figures 4A & 4B: Sensitivity analysis of profit increases and number of workers moved from sales floor by proficiency of workers moved from sales (P_f)

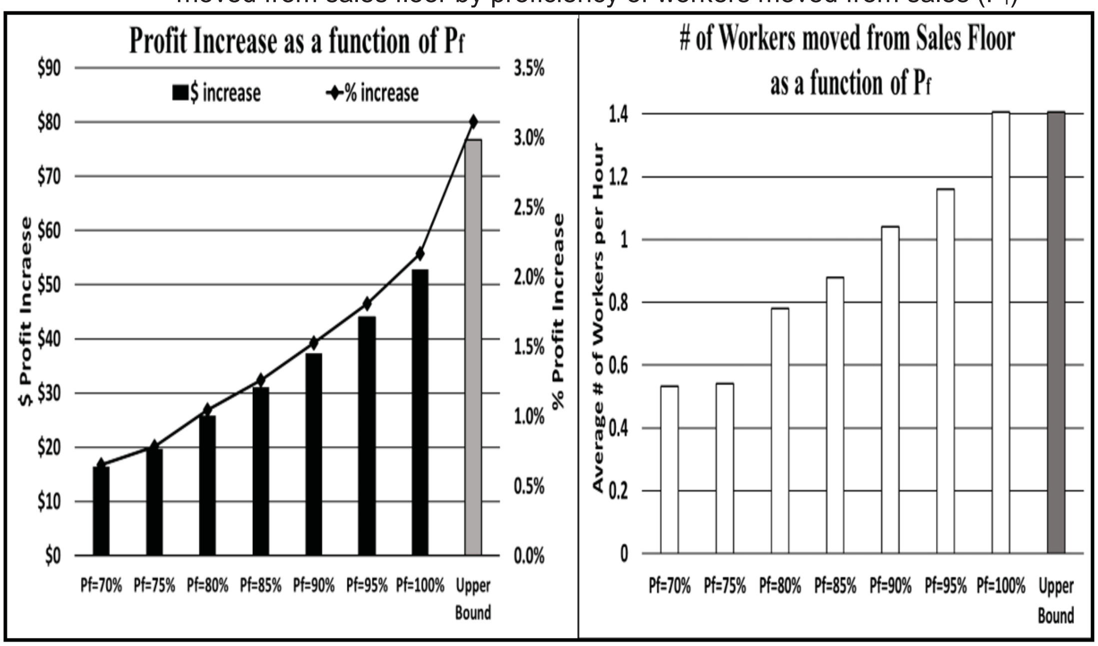
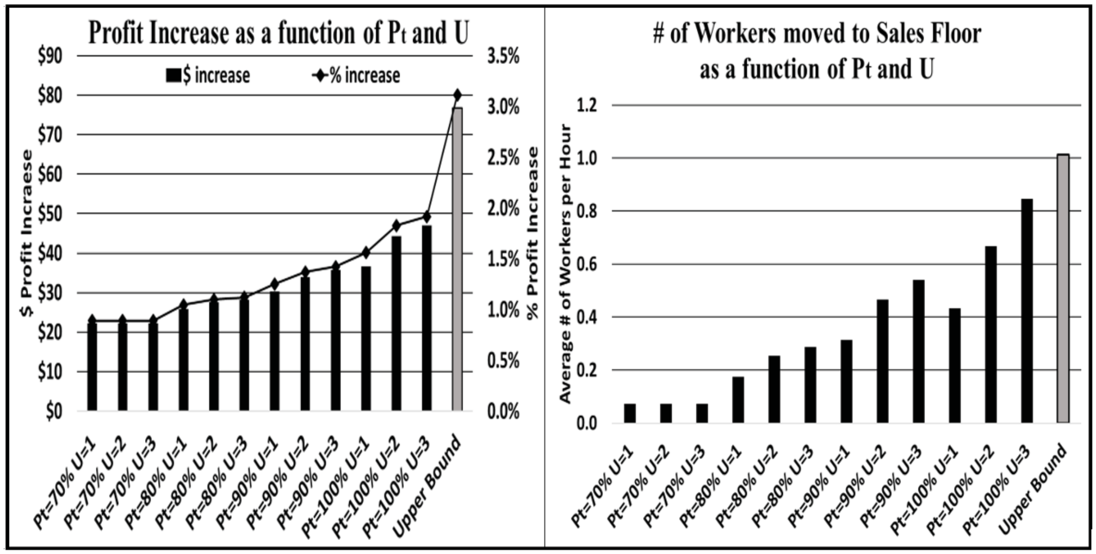


Figure 5A shows the effects on **$xu** and **%xu** as the proficiency of workers moved to sales (P_t) varies between 70% and 100% at 10% intervals, and the upper bound on number of workers moved to sales (U) varies from 1 to 3 for each value of P_t (P_f remains fixed at 80%). In general, profits increase steadily in P_t, and for a given level of P_t, exhibit diminishing returns to scale in U. Although decreasing returns to scale for U have been reported in several previous studies, including Henao et al. (2015), our results also show a strong interaction between P_t and U, with the effects of U more pronounced at higher levels of P_t.

Figure 5B displays the mean number of workers moving to the sales floor per hour over the same range for P_t and U. As with Figure 5A, we see a strong interaction between P_t and U, with

U exerting a larger influence at higher levels of P_t. The number of workers moved to the sales floor when U is constrained never quite equals the number of workers moving to the sales floor for the upper bound solutions. Flows to the sales floor increase in U, but even with P_t = 100%, U = 3 remains a binding constraint on the flow of cross-trained fulfillment workers to the sales floor.

Figure 5A & 5B: Sensitivity of profit increase and # of workers moved to sales floor by proficiency of workers moved to sales (P_t) and upper bound on # of workers moved (U)



SUMMARY AND CONCLUSIONS

In this study, we explored the impact of cross-trained flexible staff on store-level profitability for the large and rapidly growing cadre of omni-channel retailers that fulfill online orders from their stores. Our framework contributes to the workforce staffing and scheduling literature by proposing a model with the following features: a profit-oriented objective, employees with multiple skills, two types of tasks with different priorities and uncertain arrivals, and task-switching flexibility. We apply the model in a computational study to ascertain and compare the benefits of specialized and partially cross-trained fulfillment workers for daily customer traffic patterns, hour-by-hour traffic variability, and forecast error magnitude. We measure the benefits of cross-training in terms of actual and potential profit increases and the flows of cross-trained workers between fulfillment and the sales floor, as well as the sensitivity of our results to the assumed values for key model parameters. Compared with dedicated staffing options, the results of our study suggest that:

- The additional profit attributable to cross-trained staff increase with traffic variability (CV) and forecast error magnitude (MAD); approaching 3.0% when both experimental factors were at their highest settings.
- Cross-training flexibility helps mitigate forecast errors, exceeding the profitability of systems staffed by specialists even when the specialist system benefits from 100% accurate demand forecasts.
- Consistent with earlier cross-training studies (e.g. Nembhard, 2014), we find decreasing marginal improvements with increased cross-training. However, even a single cross-trained fulfillment worker boosts average daily profits significantly (in our case, by 1.81%),

suggesting that both low- and high-volume in-store fulfillment operations are likely to benefit from flexible fulfillment staff.

- Increasing the proficiency of cross-trained sales associates who assist with fulfillment (P_f) has a strong positive influence on profits. Increasing the proficiency of fulfillment workers who assist on the sales floor (P_t) has a similar effect, with greater impact when also accompanied by an increase in the number of cross-trained fulfillment workers U.

Because sales model and store traffic parameters were adapted from recent empirical studies, this study should provide useful insights for many retailers. However, the retail landscape is very diverse, and our ability to generalize our results to that larger domain is not assured. For example, parameters used in this study that might not apply in some retail settings include those related to employee skills, their training costs and proficiencies, their mobility, their wage rates, scheduling constraints, the availability of real-time traffic information, or some other implicit or explicit exogenous factors.

Besides considering a wider range of problem parameters, this research could also be extended in other ways. A straightforward and practical extension is to develop a timetabling model, based on this framework, for both full- and part-time employees with a planning horizon of a week or more. Along with this, other practical scheduling issues such as meal and rest breaks or unplanned absences could possibly be considered. Another useful extension would be to address retail environments with more than two types of cross-trained workers. Other possibilities for extending this research relate to modeling and solution methods. While deterministic heuristics for the inherently stochastic problem were shown to work well for the single period problem without cross-utilization, the two-stage problem may benefit from other stochastic programming solution methods, such as the L-shaped algorithm used for a two-stage integrated staffing and scheduling model by Kim and Mehrotra (2015).

Considering the dramatic impact that e-commerce continues to have on the retail landscape, flexibility is likely to remain an important element for future research and practice. Furthermore, while our study focused on omni-channel retailers with in-store fulfillment, other services involving both direct customer contact and interruptible back-office tasks (e.g., retail banks, auto repair, public safety, health care, government services) may benefit from a similar analysis.

REFERENCES

Abernathy, W. J., Baloff, N., Hershey, J. C., & Wandel, S. (1973). A Three-Stage Manpower Planning and Scheduling Model-A Service-Sector Example. *Operations Research, 21*(3), 693-711.

Bard, J. F. (2004). Staff scheduling in high volume service facilities with downgrading. *IIE Transactions*, 36(10), 985-997.

Bendoly, E., Blocher, D., Bretthauer, K. M., & Venkataramanan, M. A. (2007). Service and cost benefits through clicks-and-mortar integration: Implications for the centralization /decentralization debate. *European Journal of Operational Research*, 180(1), 426-442.

Berman, O., & Larson, R. C. (2004). A queueing control model for retail services having back room operations and cross-trained workers. *Computers & Operations Research*, 31(2), 201-222.

Billionnet, A. (1999). Integer programming to schedule a hierarchical workforce with variable demands. *European Journal of Operational Research*, 114(1), 105-114.

Birge, J.R. and F. Louveaux (2011) *Introduction to Stochastic Programming*, 2nd Edition, Springer Series in Operations Research and Financial Engineering, Springer Science & Business Media, LLC.

BLS (2015). Table B-3a: Average hourly and weekly earnings of all employees on private nonfarm payrolls by industry sector, seasonally adjusted. Current Employment Statistics. Washington, DC: US Dept of Labor.

Bretthauer, K. M., Mahar, S., & Venkataramanan, M. A. (2010). Inventory and distribution strategies for retail/e-tail organizations. *Computers & Industrial Engineering*, 58(1), 119-132.

Brusco, M. J., & Johns, T. R. (1998). Staffing a Multiskilled Workforce with Varying Levels of Productivity: An Analysis of Cross-training Policies. *Decision Sciences*, 29(2), 499-515.

Cai, X., & Li, K. N. (2000). A genetic algorithm for scheduling staff of mixed skills under multi-criteria. *European Journal of Operational Research*, 125(2), 359-369.

Campbell, G. M. (2011). A two-stage stochastic program for scheduling and allocating cross-trained workers. *The Journal of the Operational Research Society*, 62(6), 1038-1047.

Chapados, N., Joliveau, M., L' Ecuyer, P., & Rousseau, L.-M. (2014). Retail store scheduling for profit. *European Journal of Operational Research*, 239(3), 609-624.

Chuang, H. H.-C., Oliva, R., & Perdikaki, O. (2016). Traffic-Based Labor Planning in Retail Stores. *Production and Operations Management*, 25(1), 96-113.

Cuevas, R., Ferrer, J.-C., Klapp, M., & Muñoz, J.-C. (2016). A mixed integer programming approach to multi-skilled workforce scheduling. *Journal of Scheduling*, 19(1), 91-106.

Cunnane, C. (2015). Ship From Store: An Omni-Channel Journey. Logistics Viewpoints: A Blog for Logistics, Supply Chain, and 3PL Executives: ARC Advisory Group. http://logisticsviewpoints.com/2015/04/15/ship-from-store-an-omni-channel-journey/ (accessed 26-Nov- 2017)

De Bruecker, P., Van den Bergh, J., Beliën, J., & Demeulemeester, E. (2015). Workforce planning incorporating skills: State of the art. *European Journal of Operational Research*, 243(1), 1-16.

de Koster, R. (2002). The Logistics behind the Enter Click. In Klose A., Speranza M.G., van, & W. L.N. (Eds.), *Quantitative Approaches to Distribution Logistics & Supply Chain Management* (pp. 131-148). Berlin: Springer.

Easton, F. F. (2011). Cross-training performance in flexible labor scheduling environments. *IIE Transactions*, 43(8), 589-603.

Fisher, M., Krishnan, J., & Netessine, S. (2006). *Retail Store Execution: an Empirical Study*. Working paper, Operations and Information Management Dept, Univ. Pennsylvania. Philadelphia.

Fisher, M. L., Krishnan, J., & Netessine, S. (2009). Are Your Staffing Levels Correct? *International Commerce Review*, 8(2-4), 110-115.

Fisher, M. L., & Raman, A. (2010). *The New Science of Retailing.* Harvard Business Press.

Forrester Research (2014). *Customer Desires Vs. Retailer Capabilities: Minding The Omni-Channel Commerce Gap*. https://www.accenture.com/ca-en/~/media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Technology_7/Accenture-Customer-Desires-VS-Retailer-Capabilities.pdf (accessed 26-Nov-2017).

Fowler, G., & Dodes, R. M.(2010--Online). *Retailers Tap Stores to Speed Online Orders. Wall Street Journal.* http://www.wsj.com/articles/SB10001424052748703565804575238542925807102 (accessed 26-Nov-2017)

Grewal, D., Baker, J., Levy, M., & Voss, G. (2003). The Effects of Wait Expectations and Store Atmosphere Evaluations on Patronage Intentions in Service-Intensive Retail Stores. *Journal of Retailing*, 79(4), 259-268.

Henao, C. A., Ferrer, J. C., Muñoz, J. C., & Vera, J. (2016). Multiskilling with closed chains in a service industry: A robust optimization approach. *International Journal of Production Economics*, 179, 166-179.

Henao, C. A., Munoz, J. C., & Ferrer, J. C. (2015). The impact of multi-skilling on personnel scheduling in the service sector: a retail industry case. *Journal of the Operational Research Society*, 66(12), 1949-1959.

IBM (2017) *IBM ILOG CPLEX Optimizer*. http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/ (accessed 26-Nov-2017.

Iravani, S. M., Kolfal, B., & Van Oyen, M. P. (2011). Capability flexibility: a decision support methodology for parallel service and manufacturing systems with flexible servers. *IIE Transactions*, 43(5), 363-382.

Jacobs, L. W. & Bechtold, S.E. (1993). Labor utilization effects of labor scheduling flexibility alternatives in a tour scheduling environment. *Decision Sciences*, 24(1), 148-166.

Kabak, O., Ulengin, F., Aktas, E., Onsel, S., & Topcu, Y. I. (2008). Efficient shift scheduling in the retail sector through two-stage optimization. *European Journal of Operational Research*, 184(1), 76-90.

Kim, K. & Mehrotra S. (2015). A Two-Stage Stochastic Integer Programming Approach to Integrated Staffing and Scheduling with Application to Nurse Management, *Operations Research*, 63(6), 1431-1451.

Lam, S., Vandenbosch, M., & Pearce, M. (1998). Retail sales force scheduling based on store traffic forecasting. *Journal of Retailing*, 74(1), 61-88.

Lang, G. (2010). *Fulfillment systems in Multi-Channel Retailing - Customer Expectations and Economic Performance.* Paper presented at the 8th International Research Conference in Logistics and Supply Chain Management (RIRL), Bordeaux, France.

Loucks, J. S., & Jacobs, F. R. (1991). Tour Scheduling and Task Assignment of a Heterogeneous Work Force: A Heuristic Approach. *Decision Sciences*, 22(4), 719-738.

Mahar, S., Bretthauer, K. M., & Venkataramanan, M. A. (2009). The value of virtual pooling in dual sales channel supply chains. *European Journal of Operational Research*, 192(2), 561-575.

Mani, V., Kesavan, S., & Swaminathan, J. M. (2015). Estimating the Impact of Understaffing on Sales and Profitability in Retail Stores. *Production and Operations Management*, 24(2), 201-218.

Mou, S., Robb, D.J, and DeHoratius, N. (2018). Retail store operations: Literature review and research directions. *European Journal of Operational Research*, 265(2): 399-422.

Nembhard, D. A. (2014). Cross training efficiency and flexibility with process change. *International Journal of Operations & Production Management*, 34(11), 1417.

Netessine, S., Fisher, M., & Krishnan, J. (2010). *Labor Planning, Execution, and Retail Store Performance: an Exploratory Investigation*. Working Paper. Operations and Information Management Department, The Wharton School, University of Pennsylvania. Philadelphia.

Parasuraman,A., Berry, L.L., and Zeithaml, V. (1988). SERVQUAL: A Multiple-Item Scale For Measuring Consumer Perceptions of Service Quality, *Journal of Retailing*, 64(1): 12-40.

Partridge, A. R. (2013). Smart Shipping Strategies: Using Stores as Fulfillment Centers. *Apparel* (November 4): Columbia, SC. https://apparelmag.com/smart-shipping-strategies-using-stores-fulfillment-centers (accessed 26-Nov-2017).

Perdikaki, O., Kesavan, S., and, & Swaminathan, J. M. (2012). Effect of Traffic on Sales and Conversion Rates of Retail Stores. *Manufacturing & Service Operations Management*, 14(1), 145-162.

Perdikaki, O., Peng, D. X., & Heim, G. R. (2015). Impact of Customer Traffic and Service Process Outsourcing Levels on e-Retailer Operational Performance. *Production and Operations Management*, 24(11), 1794-1811.

Quan, V. (2004). Retail labor scheduling. *OR/MS Today*, 31(6), 32-35.

Restrepo, M.; B. Gendron, and L-M Rousseau (2017). A two-stage stochastic programming approach for multi-activity tour scheduling. European Journal of Operational Research 262, 620-635.Rueter, T. (2014). Store fulfillment advances. *Internet Retailer*. https://www.internetretailer.com/2014/04/02/store-fulfillment-advances (accessed 26-Nov-2017).

Ryan, T. (2014). Is ship-from-store a proven omnichannel benefit? *Retail Wire*. December 23,2014. Retrieved from http://www.retailwire.com/discussion/17986/is-ship-from-store-a-proven-omnichannel-benefit (accessed 26-Nov-2017).

Schultz, K. L., McClain, J. O., & Thomas, L. J. (2003). Overcoming the dark side of worker flexibility. *Journal of Operations Management*, 21(1), 81-92.

Sheldon, P., Silverman, A., & Lawrie, G. (2014). *Why Every Online Retailer Should "Ship-From-Store"*. Retrieved from https://www.forrester.com/report/Why+Every+Online+Retailer+Should+ShipFromStore/-/E-RES115372 (accessed 5-August-2019).

Song, Q. (2015). Lessons Learned and Challenges Encountered in Retail Sales Forecast. *Industrial Engineering & Management Systems*, 14(2): 196-209.

Talarico, L., Duque, M., & Andrés, P. (2015). An optimization algorithm for the workforce management in a retail chain. *Computers & Industrial Engineering*, 82, 65-77.

Thompson, G. M. (1992). Improving the Utilization of Front-Line Service Delivery System Personnel. *Decision Sciences*, 23(5), 1072-1098.

US Dept Commerce (2017). Quarterly Retail E-Commerce Sales 2rd Quarter 2017. Washington, DC: US Census Bureau. (accessed 14-Nov-2017).

Van den Bergh, J., Belien, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, 226(3), 367-385.

Xia, Y. and Zhang, G. P. (2010). The Impact of the Online Channel on Retailers' Performances: An Empirical Evaluation. *Decision Sciences*, 41(3), 517-546.

Zhang, J., Farris, P. W., Irvin, J. W., Kushwaha, T., Steenburgh, T. J., & Weitz, B. A. (2010). Crafting integrated multichannel retailing strategies. *Journal of Interactive Marketing*, 24(2), 168-180.

Zolfaghari, S., El-Bouri, A., Namiranian, B., & Quan, V. (2007) Heuristics for Large Scale Labour Scheduling Problems in Retail Sector. I*NFOR* 45(3), 111-122.

DECISION SCIENCES INSTITUTE

What's in a Name – Identity, Community, Controversy, Responsibility, Reputation - Cases A & B

John Davies
Victoria Business School, Victoria University of Wellington, New Zealand
john.davies@vuw.ac.nz

ABSTRACT

A set of inter-related cases, set in the context of a murderous terror attack on a religious group provides a platform for examination of matters as diverse as the conduct of international and national politics, the governance of social media and the identity and accountability of sporting organisations.
The cases provide opportunity to examine the responsibilities and accountabilities associated with functional management and governance, especially as they relate to crisis and media management, the management of stakeholder relationships, identity / brand management, on the one hand, then gun law reform and social media regulation, on the other hand.

KEYWORDS: Stakeholder Management, Brand Identity & Equity, Crisis Management, Governance, Ethical behaviour, Sport Culture,

CASE OVERVIEW

Inter-related Cases A, B & C

A set of inter-related case narratives provides a platform for a multi-level and multi-perspective analysis of how a white-supremacist terror attack on a religious group has impacted a country, a city and its community organisations.
The cases are set in the context of a murderous terror attack, a hate crime, and a spectrum of stakeholder reactions at local, national and international levels. The context also includes a perceived need for various community organisations to support and care for victims and their families, to foster diversity and inclusivity in the community, and address the concerns of others who see their very identity, for example, the Crusaders professional franchise sports team, as being insensitive, divisive and disrespectful.
Each case narrative provides opportunity to conduct a formal stakeholder analysis to document and understand the consequential impact of the attack on a broader set of stakeholders than just those subject to attack. They also provides opportunity to understand how responses to the attack both reflect and reshape the dynamics of stakeholder relationships, which have sparked both global action and impacted community organisations in unforeseen ways.
In particular, each narrative also provides opportunity to identify how dormant stakeholders or prior non-stakeholders may emerge post-event with significant attributes of power/influence, legitimacy and urgency that, in turn, require other stakeholder organisations to reflect on their identity, purpose, responsibilities especially in terms of how others perceive them.
The inter-related Cases A and B highlight the nature and impact of the changing stakeholder landscape that may emanate from unforeseen external events – in this case, an act of terror. Case A looks at community responses and the management of political relationships in the international domain to curtail the spread hate speech, the live-streaming of murderous activity etc. As such, it may explore the issue of social-media regulation taking place at national government level or via global, multi-country forums. Case B requires attention be given to the nature of stakeholder relationships and dynamics, and how they may impinge on the functional

management of brand identity and brand names at the organisational level.
In addition, both case narratives A and B provide a context for examining changing societal attitudes to religion, hate speech, extremist groups, minorities etc, specifically but not exclusively Muslims, on the one hand, and the management of terror attack risk, crisis and media management, on the other hand.
A further Case C, not addressed here, examines how the terrorist's choice to live-stream, via Facebook, his shooting/killing of innocent people at a place of worship, has had many unintended consequences for a spectrum of current and erstwhile stakeholders. They have also led to the emergence of new stakeholders and groups coalescing around issues of social media and internet regulation, and of the power and responsibilities social media companies like Facebook and Facebook-owned social media companies.

Here, we use Case B to direct emphasis to whether the community's most prominent and successful sporting organisation, the Crusaders, should respond to pressure to decide to change its name and rebrand. Its name and its identity has particular sensitivity given the connotations with mediaeval religious wars between followers of different faith, one being the dominant faith in the community, and the other being the faith that was subject to attack. Case B may be separated into Case B1 and B2.
Case B1 can explore the need for the Crusaders to develop an awareness of, and then act appropriately on matters relating how to its formal name and its brand identity as the Crusaders may be perceived as disrespectful and insensitive to ethnic minority and religious minority groups.
Case B2 can be used more broadly to explore issues of ethical behaviour and moral responsibility relating to individuals, to matters of management and governance at the Crusaders, and of national body governance. It foreshadows discussion of how such behaviour may be perceived by a spectrum of stakeholders with different stakes and interests, and then how stakeholder legitimacy intersects with influence and an ability to command urgency, and to influence the nature and timeliness of action taken by critical actors.
Case B2, may also be used to explore specific issues of locus of responsibility for branding decisions that relate both to the franchisee organisation, the Crusaders, and to the franchisor, the national governing body of the sport, the New Zealand Rugby Union (NZRU).
As such, Cases B1 and B2 not only foreshadow and encompass increasing calls for accountability in relation to the managerial and governance practices of professional sports organisations and their governing bodies. They also surface societal disquiet with the culture of sports organisations.

Here, the case narratives are set out in sequence, but each is organised similarly with respect to the flow of time. As such, although each case can be explored separately, they can also be explored in alternation with case information released in chunks, and then discussed as events lapse over time. For the latter, sample students discussion prompts are interspersed throughout the narrative, in illustration of how issues may be explored. Both modes provide an effective basis for class exercise.

CASE A1 - POLITICAL AND COMMUNITY RESPONSES

Week 1, Day1 -The Terror Attack

On **Friday afternoon, 15 March 2019**, at 1.15pm, a lone gunman brandishing a variety of semi-automatic weapons, entered the Masjid Al Noor Mosque in Deans Avenue, Christchurch, and then the nearby Linwood mosque. The gunman, later identified as a white-supremacist

Australian, who had been living in the provincial city of Dunedin proceeded to fire his weapons in a deliberate manner, killing fifty-one and injuring scores of Muslims in prayer.
The attack which was live-streamed on Facebook by the gunman, took place over a seventeen-minute period before he was captured by police (MFAT, 2019). The gunman had also posted a 'manifesto' online, and emailed a copy to media in New Zealand.

1 Stakeholder Analysis
1a Provide a list of **generic** stakeholders linked to the Terror Attack and state how they may have been impacted by the actions of the key players listed above.
1b Then for those generic stakeholder, identify a set of specific stakeholders, and the impact on those stakeholders relating to their roles, responsibilities and identities
2 Crisis Management
2a Use Bundy, Pfarrer et al's framework to explore what actions would be expected of government giving consideration to practical, humanitarian, media-information, political and legislative matters.

Responses to the Terror Attack

Week 1, Day1 - The Immediate Organised Response

The police responded to emergency calls **immediately**, dispatching anti-terror squads to the area, and capturing the terrorist as he fled from the Linwood mosque. Para-medics raced to the mosques to treat the wounded victims, and prepare the wounded for delivery to local hospitals who were ready having been alerted by two Muslims who had witnessed the shooting. (RNZ, 2019x).
Christchurch had faced a similar kind of terror, albeit of a natural kind in the 2011 earthquake that resulted in 167 deaths. That unfortunate experience meant that the response of the police, the military, the para-medics, hospitals, social services, the schools and community groups etc, was immediate, practiced and effective.
The visible response from the police, with extra officers on the ground outside mosques and elsewhere in the city, and helicopters flying overhead, blockades etc, was reassuring in the event that a second attack was a possibility. The invisible response from the police in meeting with, and briefing victims' families and Muslim groups, was also meant to provide reassurance to a group that felt under-attack, isolated yet exposed (Stuff, 2019a), and also to make them aware of the legal procedures and time frames required to identify and release the bodies for religious burial purposes. Overall, the response from other community groups and organisations almost seemed practised.
Within hours, a crisis centre was set up at Hagley Community College to assess the needs of victims and their families and to ensure that co-ordinated support could be provided urgently to those in need. The support provided by the existing *All Right? Campaign* was also immediate and reflective of a prior Christchurch response to the effects of the deadly 2011 Christchurch earthquake (all Right, 2019).
The *Campaign* organisation had been launched in 2013, as a Healthy Christchurch initiative led by the Canterbury District Health Board and the Mental Health Foundation of New Zealand, 'to improve Cantabrians' mental health and wellbeing as they recover and rebuild from the earthquakes' (CreativeNZ, 2019).
The *Campaign's* success and effectiveness in dealing with post-traumatic stress is evident in continuing support from the Ministry of Health and relationships with the Ministry of Social Development and other organisations including the Red Cross, Special Kids, Involved Parents (SKIP), the Christchurch City Council and the Waimakariri District Council (All Right, 2019). Similarly, the Ministry of Education trauma team was immediately mobilised on the Friday afternoon, to work alongside the team of mental health and support workers still embedded

within schools as part of the Ministry of Education Mana Ake Programme following the earthquake (Ardern, 2019b).

Week 1, Day1 - Outpourings of Grief, Empathy and Support

The response from the community on learning of the terror attack was of course, initial shock, but a heartfelt and instantaneous outpouring of grief, empathy and support for the Muslim victims, their families, and the Muslin community in general.
The response manifested as a wall of floral tributes, messages, candle-lit vigils at mosques in Christchurch and throughout New Zealand, and fund-raising events and campaigns that raised more than $10m in less than a week through fund-raising sites set up by the NZ Council of Victim Support Groups. Sporting events and theatrical productions were also cancelled, and remembrance services were organised.
Other symbols of support included the wearing of headscarves, and meaningful slogans like "They are us!" resonated with local Muslims who had been living in fear, and who had feelings of increasing isolation. Bariz Shah (2019), President of the University of Canterbury's Muslim Student Association, offered a view that after many years of feeling that he did not belong, the actions and words of Cantabrians and New Zealanders in general, and the Prime Minister, Jacinda Ardern, in particular, had provided him with a different sense of belonging.

Week 1, Day1 - The Response from the Prime Minister and Government

The Immediate Response to Reassure the Muslim Community
In the hours following the attack, Prime Minister, Jacinda Arden condemned the gunman's hate-filled atrocity, and said

> 'We should not be perpetuating, sharing, giving any oxygen to this act of violence and the message that sat behind it' – a message aimed at both individuals and social media platforms (Ardern, 2019aa; Cooke, 2019).

She had already acted to contact the latter, and had received assurances that social media platforms were working to remove the footage. She indicated that she was calling for a ban on the spread of messages of hate.
A Prime Ministerial statement (2019a) on the **Saturday morning** was designed to reassure the Muslim community that they would be supported in their grief and protected from harm and from extremism. On the same day, she met with Christchurch Muslim leaders at the Canterbury Refugee Resettlement and Resources Centre in the company of Deputy Prime Minister Winston Peters, and Leader of the Opposition National party, Simon Bridges.
She indicated her awareness of possible roots of the terror attack, and said that the police and the intelligence community were focused on extremism of every kind, and that security agencies would be assessing whether there had been any activity on social media that should have prompted a response. She said that the offender had obtained a gun licence and weapons whilst living in NZ:

> "While work is being done as to the chain of events that led to both the holding of this gun licence, and the possession of these weapons, I can tell you one thing right now.
> Our gun laws will change.
> There have been attempts to change our laws in 2005, 2012 and after an inquiry in 2017.
> Now is the time for change**."**

She also said that:

> "Police are aware of distressing material relating to this event being online and are reminding people it is an offence to distribute objectionable material."

It later transpired that these comments were not empty rhetoric, or devoid of meaning! Neither was the image of the Prime Minister wearing a hijab devoid of meaning! 'It's a photo which has gone around the world. She looks forward, her brow raised with concern and hands clasped. She's not speaking and doesn't show any sign that she is about to launch into a political speech. Jacinda Ardern is simply listening' (McConnell, 2019).

2 The Impact of the Prime Minister's and Government's Statements and Actions Politics and Brfanding

Consider the statements made, and actions taken by the Prime Minister, Jacinder Arden, and the Labour/Greens/NZ First collation government, and whether their statements:

(i) Would have led to reputational gain as individuals or politicians.
(ii) created beneficial brand equity at a political level

3 Brand Management - Evaluating Actions by the Crusaders organisation

Students may be introduced to issues of Branding, Brand Equity, Brand Management at this stage – that relate to politicians and political parties.

(i) Students should engage in preliminary discussion of whether the Prime Minister has enhanced an identity around decisiveness, compassion etc –that constitutes a political brand identity, or constitute a brand in itself.

Week 1, Day4 - A Call to Immediate Arms Reform

On **Monday, 18 March 2019**, in its first meeting following the attack, Cabinet made 'in-principle decisions' around the reform of gun laws, hoping to finalise details of gun-law reforms at its next meeting (Arden, 2019c). Cabinet also decided that there would be an inquiry to look at the specific circumstances leading up to the Christchurch attack. Ardern referred to the terror attack in Christchurch on Friday as 'the worst act of terrorism on our shores', and that 'the time to act is now'

Week 1, Day 7 – A Promise for Gun Law Reform

On **Thursday, 21 March**, the government announced how it would proceed with gun law reform placing a ban on all military style semi-automatics (MSSA) and assault rifles in New Zealand (Ardern, 2019e; Mann, 2019), stating that:

> 'On 15 March our history changed forever. Now, our laws will too. We are announcing action today on behalf of all New Zealanders to strengthen our gun laws and make our country a safer place. All semi-automatic weapons used during the terrorist attack on Friday 15 March will be banned - the Government will take immediate action today to restrict the potential stock-piling of these guns.'

Week 2, Day 1 – Hate Speech

On **22 March 2019**, when addressing Organisation of Islamic Cooperation Emergency Meeting on the Christchurch Terror Attack, Deputy Prime Minister Winston Peters referred to the innocent people who were killed as they practiced their religion, and said:

> 'In a country that practices religious tolerance, an attack on one of us, observing their beliefs, is an attack on all of us. … Whenever and wherever a terrorist strikes, the aim is to provoke fear and panic. In New Zealand, it has failed. It failed because our thoughts are not the terrorist's thoughts, and his extremist ways are not our ways. And to be clear, in New Zealand hate speech is not tolerated.'

Week 4, Day 6 – Gun Law Reform - Arms Amendment Bill

On **10 April 2019**, the Arms Amendment Bill was passed into legislation following its third reading in Parliament. Work had also begun on the Arms Amendment No. 2 Bill, which would address issues about 'a gun register, the licensing regime, the system of Police vetting, and the 'fit and proper person' test, storage requirements and penalties', amongst other matters (Nash, 2019).

Week 1, Day 4 – Commemoration and Respect - PM Jacinda Arden: They are us

On that same **Monday following the attack**, Ardern (2019c) also said that the government's Department of Internal Affairs (DIA) would be working in conjunction with the Muslim community, iwi, local government, and the mayor of Christchurch, police, and other agencies so that New Zealanders would have 'the ability to commemorate as one the lives lost at Deans Avenue and Linwood mosques'. One such opportunity to reflect took place one week after the attack on Friday, 22 March when the nation stopped for the *Muslim Call to Prayer* at 1.30pm (Ardern, 2019d), at Hagley Park, opposite the Al Noor Mosque, a time when the nation stopped to mourn. A National Remembrance Service took place just one week later.
At that service, Ardern said:

"**We will remember the tears of our nation**, and the new resolve we have formed,"
Over the past two weeks **we have heard** the stories of those impacted by this terrorist attack.
They were **stories of bravery**.
They were stories of those
… who were born here, grew up here, or who had made New Zealand their home.
… who had sought refuge, or sought a better life for themselves or their families.
These stories, they now form part of our collective memories.
They will remain with us forever.
They are us.
But with that memory comes a responsibility.
A responsibility to be the place that we wish to be.
A place that is diverse, that is welcoming, that is kind and compassionate. Those values represent the very best of us.
But even the ugliest of viruses can exist in places they are not welcome.
Racism exists, but it is not welcome here.
An assault on the freedom of any one of us who practices their faith or religion, is not welcome here.
Violence, and extremism in all its forms, is not welcome here.
And over the last two weeks we have shown that, you have shown that, in your actions.
And we also ask that the condemnation of violence and terrorism turns now to a collective response.
The world has been stuck in a vicious cycle of extremism breeding extremism … and it must end."
We cannot confront these issues alone, none of us can.
But the answer to them lies in a simple concept that is not bound by domestic borders, that isn't based on ethnicity, power base or even forms of governance.
The answer lies in our humanity.' (Heard, 2019)

The Christchurch Call - Reforming Social Media to Counter Terrorism

The NZ Government expressed a view that the terrorist attack had 'made clear once again the harms that can be caused by terrorist and violent extremist content online, a threat that continues to evolve.' There was concern that the attack, having been livestreamed, had not only gone viral at the time, but continued to remain available on social media despite the measures taken by social media providers to remove it. Indeed, Prime Minister Ardern regarded the March 15 attack as shocking in its use of social media as a tool in the act of terror. She sought to address the matter through what became known as *The Christchurch Call* to eliminate terrorist and violent extremist content online (MFAT, 2019).
Ardern's vision, however, related to a global rather than a local response. Ardern found a willing political ally in French President Emmanuel Macron in bringing together political leaders and leaders from the social media and internet world to develop the *Call* as an action plan and to agree to implementing the call.

Week 9, Day 6 – The Christchurch Call, Paris Summit

On 15 May, two months to the day after the Christchurch terror attack, the *Call* brought leaders from ten countries and Facebook, Google, Instagram etc to Paris, where a commitment \was made to a set of collective actions:

- developing tools to prevent the upload of terrorist and violent extremist content;
- countering the roots of violent extremism;
- increasing transparency around the removal and detection of content, and
- reviewing how companies' algorithms direct users to violent extremist content.

Ardern said that implementing the Call was owed to 'those affected by the attacks in Christchurch, and other attacks in cities and towns around the world where terrorism and violent extremism have struck, to undertake this work' (Beehive, 2019).

Set alongside the Christchurch Call as a practical global response aimed at limiting the spread of offensive extremist views and the roots of violent extremism, Case A now provides opportunity to other matters deemed to be offensive in the context of the terror attack, and what a practical response to those matters might be.

The offense relates to the name, imagery and identity of the southern hemisphere's most successful professional sports team, the Canterbury Crusaders. Some had regarded the Crusaders name as being inappropriate at the time the franchise rugby team was formed in 1996, because of its connotations to the mediaeval geo-religious wars between the Christian Crusaders and the Muslim-faith based Saracens.

Whilst it had been previously claimed that the name was offensive to Muslims in general, and to the community at large, because of such connotations, the poignancy of the terror attack on the Muslin community had refocused attention on the Crusaders name.

THE CANTERBURY CRUSADERS - CASE B

The Canterbury Crusaders and SupeRugby

The Canterbury Crusaders are a professional rugby union organisation playing in the SANZAR sanctioned southern hemisphere SupeRugby competition. They were formed in 1996, as one of five regional franchise teams initially licensed by the New Zealand Rugby Football Union (NZRFU, now NZRU) governing body to compete in the original *Super 12* international interprovincial competition, involving the leading southern hemisphere rugby playing countries – South Africa, Australia and New Zealand (SANZAR). They have been the most successful team in the 25 years of the competition. In turn, the SuperRugby competition is widely regarded as the toughest competition in world rugby, and the Crusaders are recognised as one of the elite teams in global professional sport.

They are revered for creating an organisational culture that breeds success, winning nine of the thirteen championship finals they contested between 1996 and 2018; and for a development programme that has fostered character and community-mindedness amongst players. The organisation and its players are regarded as role models not just within its extended Crusaders community, but also globally.

The Crusaders region, encompassing the Canterbury, Nelson, Marlborough, Buller and West Coast provinces, also took great pride in their achievements and how the Crusaders conducted themselves on and off-the-field and in their support of local communities. They were widely thought of as exemplars of what could be achieved with shared values relating to excellence, vision and a strong team culture.

They are a pride of the community, whose successes have served to lift the spirits of people in Christchurch since the deadly earthquakes of 2011 devastated large regions of the city, and led to the death of more of than 150 people.
The advent of professional rugby in 1996 positioned the SupeRugby competition at a level above what was then the amateur National Provincial Championship (NPC) involving NZ's elite representative provincial teams. At that time, NPC team names simply reflected the names of their provinces, although most teams had nicknames and/or mascots.
However, multi-million dollar media-rights and sponsorship deals not only enabled an era of professionalization and professionalism in the conduct of rugby as a sport, but also in fostering commercialism and commercialisation in the conduct of all dimensions of the business of rugby.
Much consideration was given to the creation of franchise teams that had exclusive geographical territories and then names, logos and iconography that were readily identifiable with the franchise region, preserve existing traditions, and which reflected or had connotations with geographical, cultural or other demographic characteristics.
It was also considered important that names were readily identifiable in overseas rugby playing markets in order to attract spectators to matches and viewers. As a consequence, the initial franchise team names were double-barrelled place and nick names: namely Wellington Hurricanes, Waikato Chiefs, Otago Highlanders, Canterbury Crusaders and the Auckland Blues. The place names were those of the elite provinces well-known in international markets. The Highlanders name reflected the origin of the early Scottish settlers; the Chiefs name and associated iconography reflected the region's pervasive Maori heritage and culture; the Hurricanes name reflected Wellington's windy nature. By contrast, the Crusaders name, apart from its alliterative qualities, was said to be a 'nod' to the English heritage of Christchurch and Canterbury, albeit drawing on images of mediaeval England that afforded a link to 'the pageantry of knights and horsemen', and the iconography, later used in marketing, of shields, banners, flags and swords.
At the time, all five franchises were owned by the NZRU but licensed to a lead or host province in each region to organise on behalf of, and in harness with the other provinces in the region. The ownership model meant that the NZRU had the final decision in operating a centralised player drafts, player payment system, and names. Coincidentally, the NZRU CEO in 2019 had been the CEO of the Canterbury Rugby Football Union and the Crusaders in 1996 (TVNZ, 2019).

Immediate Reactions to the Terror Attack

Week 1, Day1 - The Impact on Sports Events

Rugby – New Zealand's national sport is rugby union, and the Canterbury Crusaders are the country's most successful professional rugby franchise. **On the day of the attack**, the Crusaders were in Dunedin, preparing to play the Otago Highlanders the following day. Both teams and the New Zealand Rugby Union (NZRU) agreed that it was inappropriate for the match to be played. It was cancelled out for respect for the Muslim community with the Crusaders team returning to Christchurch to provide support to their families, the victims and the community (Egan, 2019).

Other Sports - **Within hours of the attack**, the mainstream media were reporting the views of New Zealand CEO, David White, as saying that New Zealand sports hosting would change forever with the mosque shootings (Egan, 2015). White had immediately been in contact with his Bangladesh counterpart, Nizamuddin Chowdhury, and the decision to cancel the series of Test matches between NZ and Bangladesh, was quickly made, with "no other consideration" being given to alternatives.

The Bangladesh cricket team had been in Christchurch to play an international test match against the NZ Black Caps starting on the following day. They were on their way to the Masjid Al Noor Mosque for afternoon prayers when the Terror Attack happened. They fled from the area unharmed, and following the cancellation of the Test match, they flew home the next day.
White offered a view that the attack would "change the entire fabric of international sports hosting" and that the idea of even New Zealand being a 'safe haven' had disappeared. He suggested that there was a renewed need for an in-depth examination of security matters.
White also said that a decision had also been made to cancel the two remaining fixtures between the New Zealand women's development team and the Australia under-19 women's team scheduled for same weekend at nearby Lincoln's Bert Sutcliffe Oval.

THE CALL FOR THE CRUSADERS' NAME CHANGE

Week 1, Day1, Friday – Pressure bubbling in Social Media

Almost immediately after people became aware of the Attack, the web became active with calls for the Crusaders to change their name! It was suggested that not only did the Crusaders' name and other imagery have inappropriate connotations to violent religious wars between Christian and Muslim nations, but that, at a time, when Muslims in Christchurch had been harmed, the name would serve as a continual reminder of disrespect, on the one hand, and threat, on the other hand. As such, it was claimed the name was inappropriate.

1 Initial Response of the Crusaders What action should the Crusaders organisation have taken on first learning of the social media commentary?
2 Stakeholder Analysis In parallel, students should undertake a preliminary **stakeholder analysis**, identifying generic and specific stakeholders, their roles, stakes or interests, and their **attributes**. Such **stakeholder attributes** would include: their **power** or **influence** to impact others; their ability to act with **urgency** or **timeliness** and/or to effect urgent or immediate response from other; and the **legitimacy (**social, legal or otherwise **acceptance** of their views or actions. The Stakeholder Analysis should seek to identify **stakeholder dynamics** in terms of the emergence of 'new' stakeholders, or stakeholders whose attributes, say power or legitimacy have changed because of the Terror Attack
3- Crises and Crisis Management Students should briefly reflect on what they understand to be a crisis, and how crisis management situations can be distinguished from other management situations. In this case, students should become aware of how stakeholders associated with the Terror Attack, and whose interests/views have gained in legitimacy, may also have become and gained influence as Crusaders stakeholders, with an ability to 'coerce' a response – effecting a crisis in the Crusaders organisation

Week 1, Day 3, Sunday – The Crusaders' Response to the Call for a Name Change

A media statement from the Crusaders CEO, Colin Mansbridge, **on the Sunday following the attack**, expressed shock at the tragedy, stated that:

> 'Like all New Zealanders, the Crusaders team and organisation are deeply shocked by this tragedy and our thoughts are with the victims and their families. This is bigger than rugby and we're absolutely heartbroken for our wider community, which is where our thoughts are … the team and the wider organisation are united with our community in standing against such abhorrent acts … in Christchurch, and in standing in support of our Muslim community.'

The statement acknowledged and understood the concerns that have been raised about the Crusaders name, but that it was "a reflection of the crusading spirit of this community (and Canterbury rugby), and certainly not a religious statement.' (Crusaders, 2019a)

> 'What we stand for is the opposite of what happened in Christchurch yesterday (Friday); our crusade is one for peace, unity, inclusiveness and community spirit.
> In terms of the Crusaders name, we understand the concerns that have been raised. For us, the Crusaders name is a reflection of the crusading spirit of this community.
> In our view, this is a conversation that we should have and we are taking on board all of the feedback that we are receiving, however, we also believe that the time for that is not right now.'

Earlier on the same Sunday, Mansbridge had told TVNZ (2019), before issuing the statement, that the club would have a discussion about what to do next.

> 'There's a lot of emotion going around in terms of conversation at the moment. We've heard it, we've heard the feedback, and we do want to have a conversation about it.'
> 'At an appropriate time, we will thoroughly consider the issues that have been raised and our response to that," the Crusaders said in a statement. "That will include conversations with a range of people, including our Muslim community.'

Mansbridge said it was important to note that "things are still fairly raw".

> "I would say we are in a state of shock. We will acknowledge the feedback that we have had. It is appropriate. (But) rugby, in the context of what has happened, it is pretty hard to … elevate this conversation right at the moment.'

4 Crisis Management - Evaluating Actions by the Crusaders organisation

Consider the actions taken by the CEO, Mansbridge:

(i) on hearing of the Terror Attack
(ii) in explaining the Crusaders' concerns, values, and action plan on the Sunday following the Attack and the social media storm.

… and how they may reflect a crisis management plan to deal with such situations?

Consider the statements made, and actions taken after the Terror Attack, during the first week, Week 1, by the Prime Minister, Jacinder Arden, and the Labour/Greens/NZ First collation government.

Consider whether their statements of compassion and commitments to reform gun law, and their consequent legislative actions:

(iii) have created additional pressure for the Crusaders to act in terms of a possible name change or
(iv) have given direction to the Crusaders' decision-making

Week 1, Day 3, Sunday – Response to the Crusaders' Initial Statement

<u>The local/global media reaction to Mansbridge's statements and sentiments</u>

News that the Crusaders were contemplating a name change resonated throughout the sports world, and was reported in all major newspapers globally, reflecting not just interest in how NZ was handling the terror attack, but also the stature of, and interest in the Crusaders (TVNZ, 2019a; Stuff, 2019a)

Senior sports reporter, Tony Smith (2019) said that 'Mansbridge deserves credit for the sympathetic tone of his statement about the name issue. The very fact that the organisation is open to considering changing the name of the most successful team in professional rugby history speaks volumes for its leadership.'

He believed, like others, that the key issue, as Mansbridge highlighted, is how the Muslim community feel.

> "Would they be uncomfortable, after last Friday's tragedy, at their city's premier sporting team continuing to bear a name associated with wars waged against Muslims in the Middle East in the 11th to 13th centuries?

He suggested that the feelings of the Muslim community 'matters more than a moniker, adopted as much for marketing reasons and alliterative properties, as geographical location', and that no one doubted 'Mansbridge's sentiments.'
A view was offered that whilst no Crusaders player or coach, in his memory, had made a direct link about their rugby journey to the holy crusades, that 'perception, however, is important, and that there was a need to know what the Crusaders name means to Christchurch's growing Muslim community? He also reflected on an oft-stated view that 'was acceptable to the general populace in the mid-1990s may not be now, given New Zealand's more diverse demographics.' No-one can deny the mosque murders have changed the socio-political landscape forever.'

5 Crisis Management - Evaluating Actions by the Crusaders organisation
As in 4 above, does the extensive global interest in the Crusaders' potential name change and the initial response for their empathetic concerns about the Muslim community, reinforce the pressure for the Crusaders get things right – for the Muslim community, the Christchurch community and the rugby community stakeholders.

Week 3, Day 5 – The Crusaders' Commitment to Considering a Name Change

Nearly three weeks later, on 3 April, following extensive discussions with Crusaders' stakeholders, and with the governing body, the New Zealand Rugby Union (NZRU), Mansbridge explained the nature of their commitment (NZH, 2019; TVNZ, 2019b).

> 'This is an event that rocked our community and brought some important issues to the fore. One of the contentious issues that has been brought up in the aftermath of the Christchurch attacks is the name of our rugby team – the Crusaders.
> "Because of our desire to be the best we can be and to support our community, we are treating the question around the appropriateness of our brand extremely seriously.
> We are committed to undertaking a thorough process, taking into account all relevant opinions and, most importantly, we are committed to doing the right thing.'

Mansbridge said:

> 'Our challenge is that the name Crusader has come to mean something quite different to many of the team's supporters. This team is proud of the positive contribution it makes to the community. In recent times, the Crusaders region has been through a number of tragic and trying events and the team has played an important role in helping galvanise the community and raise spirits following significant events. Through these events, the Crusaders name has become more reflective of a positive Crusade."

New Zealand Rugby Chief Executive Steve Tew offered a view that:

> 'In the wake of the Christchurch attacks, it is apparent that the symbolism the club has used, combined with the 'Crusaders' name, is offensive to some in the community due to its association with the religious Crusades between Christians and Muslims.
> "One thing that has become very clear in the last two weeks is that there are divided opinions on the best way forward for the brand. We understand and appreciate the passionate feedback that we are receiving on both sides of the conversation, and at this stage we are committed to keeping an open mind until the independent research has been done."

Tew then outlined actions that had been decided, including engagement of the consulting organisation Research First, to seek feedback and provide recommendations on the Crusaders team name and brand name, specifically whether to retain the name but changing the branding and associated imagery; or undertaking a complete rebranding of name and imagery:

> 'Maintaining the status quo in terms of the Crusaders name along with the current imagery of knights on horseback is, in our view, no longer tenable because of the association with the religious Crusades that has now been drawn.
> That is therefore not one of the options that we will be considering.'

Tew said that any changes will be introduced in time for next season, 2020, and that the team would drop the horses and related entertainment for the rest of the current season.

6 Brand Management - Evaluating Actions by the Crusaders' organisation
Students may be introduced to issues of Branding, Brand Equity, Brand Management at this stage.
(ii) Students should engage in preliminary discussion of whether the Crusaders name is an element of brand identity, or constitute the brand itself.
Students would need to consider what, if any issues
(iii) may relate to and conflate the essential identity of the Crusaders organisation with the Crusaders brand and brand image/identity; and
(iv) may also impact brand equity;
They should also have preliminary discussions about whether existing elements of brand equity have been damaged by the negative connotations that have resurfaced with the Terror Attack.

When the first year of Super Rugby kicked off in 1996, the Crusaders title was chosen for the side representing the provincial unions from the top half of the South Island and the brand was built on a stylised version of medieval England, in a nod to the English heritage of Christchurch. The club's marketing over the past 23 years has incorporated the pageantry of knights and horsemen, and included symbols such as swords and banners.

Week 3, Day 7 – Response to the Crusaders' Commitment to Consider a Name Change

<u>The local and global media reaction to Mansbridge's statements and sentiments</u>

Following the joint Crusaders/NZR announcement that they would contemplate a change to the name and branding, there was an unfolding of media commentary. As much as people were polarised about whether to execute change or not (Dickens, 2019; Garner, 2019; Norquay,2019; Stuff, 2019c), there were also different views about whether and how any change would affect the Crusaders (Cully, 2019; Heard, 2019; Newstalk ZB, 2019; NZH, 2019; Troughton, 2019). Nevertheless, the Crusaders organisation were praised by the Prime Minister for contemplating change (Whyte, 2019).
Religious scholar, Troughton (2019) was unequivocal in stating that it was 'scarcely conceivable the name and brand will survive careful scrutiny and reflection. Ideally, the consultation should extend to input from other religious communities, including Christians.' His view was that 'the contemporary resonances of the Crusades remain steeped in images of violence and domination that are deeply disturbing and problematic.' However, whilst on the one hand, the Christian military expeditions at the time of the first millennium had sought to reclaim the Holy Land and other territories held by Muslim nations, on the other hand, Christian groups in NZ and the USA had often embarked on crusades for social improvement or social purity, or crusades for the 'battle of souls'. However, even such 'crusader' type branding has been subject to change – often to 'mission'-led brands.
Troughton has offered clear opinion that:

> 'while many supporters will use the Crusaders name with little thought to its origins and implications, it remains an anachronism and source of discomfort to many New Zealanders – Muslim, Christian, secular and others. If the Crusaders franchise and New Zealand Rugby chose, as they might, to consult with the Christian community, they are unlikely to find great enthusiasm for retaining the current name. Indeed, New Zealand Christians will probably be grateful to see it go.'

Senior sports reporter, Kevin Norquay (2019) supported change for similar reasons, attempting to put such change in the context of it not being of the great challenge that others perceive.

> 'There are many historically tarnished names no one sensible would consider naming a sports team after - the Jihadists for example, that being one Muslim equivalent of Crusaders.

> Adding Crusaders to that list would not be the end of the world, in fact it might help start a brave new one.'

By contrast, political commentator, Garner (2019), didn't perceive 'war and blood' in the Crusader imagery, he sees talent, playing by the rules, values, ethics … and winners. He sees 'a DNA and a record that refines the word Crusaders in our country.' Radio journalist, Andrew Dickens (2019), offered a similar view of the need for change, portraying the furore over the Crusader name as another example 'that virtue signalling and political correctness is alive and well in this country.'

7 Brand Management - Evaluating Responses to the Crusaders' Action Plan
Students should consider the range and credibility of the diverse opinions that have surfaced about a potential name change and rebranding, and how such commentary could feed into analysis of the name change/branding decisions

Week 4, Day 5 – Stakeholder Views – A Petition and Survey Responses

Public opinion seemed a little less divided as evidenced by a petition and a survey. The petition to retain the Crusaders name had accumulated 25,000 supporters by 9 April.(One News, 2019a). Then on 20 April, initial findings from a survey (conducted by One News Colmar Brunton) indicated that a majority (75%) of New Zealanders favoured retention of the Crusaders' name, with 14% opposed. The survey indicated that who were more likely to favour the Crusaders keeping their name included those living in Otago and Southland at the south of NZ's South Island, right-leaning National Party supporters and people aged 18-29.

8 Brand Management - Evaluating Actions by the Crusaders organisation
Students should be confronted with the challenge of balancing the weight of public opinion, in raw percentages, against the views of those who have a voice in the media, and those who have well-articulated views relating to the ethical, political, religious and marketing nuances. Students should identify the ethical, political, religious and marketing dimensions.

Renaming and Rebranding Experiences

In considering the matter of a name change, the Crusaders' dilemma may be set alongside a spectrum of similar dilemmas relating to name change that have been faced by other sporting and education-related organisations. They include some cases where name change and rebranding were considered to be worthwhile, and some cases, where no deliberate attempt had ever been made to create an official nickname.

Indeed, calls for name changes to official and unofficial mascots, logos, buildings and teams have become common place (SDUT, 2016), Students across the United States have been pressuring their colleges to update mascots, mottos and building names that they say are insensitive, and which demean, denigrate, impugn or alienate minorities or ethnic groups (AP, 2015; Gambino, 2014; Hoft, 2015).

For example, Susquehanna University dropped its nickname, the Crusaders, due to concern over potential offensiveness and negative connotations, and to ensure sure the campus remained 'welcoming and inclusive'. SU President L. Jay Lemons said that it was a Philadelphia sportswriter, not the college itself, that dubbed the athletic teams the "Little Crusaders" in 1924. The name stuck as "medieval iconogrophy" such as the Maltese cross came to accompany the nickname.

In the 1980's, Susquehanna formally dropped the Knight mascot and Maltese cross adopted as the "Crusader flag," to distance themselves from 'the medieval Crusaders iconography that had been present in an earlier time'. However, in the late 1990's, the image of the Crusader morphed 'strangely' into the form of a caped tiger, which then remained the mascot until 2014 when students began the discussion over the name and mascot again.

Similarly, following 'one or two' complaints, from its Muslim and Jewish membership, the Middlesex County Cricket team changed its name from the Middlesex Crusaders to the Middlesex Panthers (Cramb, 2009; Coc et al., 2016)). The President, at the time, said that the nickname was introduced because 'we wanted more kids to pick up a cricket bat, that's why it was devised. We were on a crusade to get children into cricket.

Some organisations, for example, the Washington Redskins have refused to bow to external pressure to change their name, including pressure to change on ethical grounds (Mintz, 2013). For example, the Oneida Indian Nation in upstate New York even launched an advertising campaign to urge the team to change its name, stating 'we do not deserve to be called redskins. We deserve to be treated as what we are — Americans." However, A Washington Post poll in 2016, found that nine out of 10 Native Americans were not offended by the Washington Redskins name, reflecting a similar survey in 2004. Washington's owner, Dan Snyder, has claimed the poll backs his stance resist change to the team's name"

> "The Washington Redskins team, our fans and community have always believed our name represents honor, respect and pride, … polling shows Native Americans agree. We are gratified by this overwhelming support from the Native American community, and the team will proudly carry the Redskins name."

Examples relating to education and sport include:

Sport	Organisation	Nicknames	Potential Name Change	Decision to Change
All Sports	Susquehanna University	Crusaders	Squirrels & River Hawks	Yes
	Alvernia University	Crusaders		
	University of North Dakota	Fighting Siouxa	Fighting Hawks	Yes
	Amherst College	Lord Jeff(rey Amherst)		Yes
Cricket	Middlesex Crusaders			
Basketball	Washington Bullets	Bullets	Washington Wizards	Yes
AFL	Houston Colt 45s	Colt 45s	Houston Astros	Yes
	Washington Redskins	Redskins		No
Baseball	Tampa Bay Devil Rays	Devil Rays	Tampa Bay Rays	Yes
	Atlanta Braves	Braves		No
	Cleveland Indians	Chief Wahoo mascot/logo		Yes

Week 9, Day 6 – The *Christchurch Call* Paris Summit

On 15 May, two months to the day after the Christchurch terror attack, the Christchurch *Call* brought leaders from ten countries and Facebook, Google, Instagram etc to Paris, where a commitment \was made to a set of collective actions:

- developing tools to prevent the upload of terrorist and violent extremist content;
- countering the roots of violent extremism;
- increasing transparency around the removal and detection of content, and
- reviewing how companies' algorithms direct users to violent extremist content.

9 Crusaders' Decision Making

Students should consider whether the success of the global initiative, described as the *Christchurch Call*,
(i) has created additional pressure for the Crusaders to act in terms of a possible name change or
(ii) has given direction to the Crusaders' decision-making.

10 Brand Management - Advice to Crusaders about a Potential Name Change and Branding

(i) Students should set out the pros and cons for name retention or name change as they impact brand identity, brand equity, and the Crusaders name as an element of equity and identity
(ii) Students can then provide advice to the Crusaders' CEO and the NZR, in relation to issues of brand identity, brand equity, and the Crusaders name as an element of equity and identity

Week 17, Day 2 – NZ Gun Buyback Scheme - Arms Amendment No. 1

On **13 July 2019**, the government gun buyback scheme began with the voluntary surrender of banned guns. Within one month, 10k guns had been surrendered with owners obtaining agreed compensation (Small, 2019).

Week 18, Day 6 – Gun Law Reform - Arms Amendment No. 2 Bill – Second Tranche

On **22 July 2019**, the government announced a second tranche of gun law reform to include the creation of a firearms register and more stringent licencing requirements (Fitzgerald et al, 2019).

SAMPLE OF CASE B QUESTIONS AND DISCUSSION POINTS - Managerial Issues and Tasks

General:

11a - Chronology of Events, Decisions and Dilemmas
Students should familiarise themselves with the both case narratives by creating a chronology of events, including the decisions made/dilemmas confronted by the major actors and stakeholders.

11b - Stakeholder Analysis
In parallel, students should undertake a stakeholder analysis, identifying generic and specific stakeholders, their roles, stakes or interests, and their attributes.
Such stakeholder attributes would include: their power or influence to impact others; their ability to act with urgency or timeliness and/or to effect urgent or immediate response from other; and the legitimacy (social, legal or otherwise acceptance of their views or actions.
The Stakeholder Analysis should clearly identify stakeholder dynamics in terms of the emergence of 'new' stakeholders, or stakeholders whose attributes, say power or legitimacy have changed over the time of the Affair.

11c - Crises and Crisis Management
Students should reflect on what they understand to be a crisis, and how crisis management situations can be distinguished from other management situations. In this case, students should become aware of how those stakeholders whose interests/views have gained in legitimacy, may also have gained influence, and an ability to 'coerce' a response – effecting a crisis in the primary case organisation.

Specific:

12 Crisis Management - Evaluating Actions by the Crusaders organisation
12a Consider the actions taken by the CEO:
(i) on hearing of the Terror Attack
(ii) as events unfolded over days and weeks.
1 and how they may reflect a considered crisis management plan to deal with such situations?
12b Given that considerable media coverage had been given to the actions of the Prime Minister and the Government, would it be reasonable for CEO to assume that a crisis/media storm surrounding the action on a name change had been averted? Say why.

13 Governance, Multiple Jurisdictions, Justice and Fairness, and Locus of Responsibility
Offer some brief comment on where the locus of responsibility should lie in the name change? – the Crusaders, NZR, *World Rugby*? That is, offer reasoned comment on whom should take responsibility for a name that has negative connotations for the religious minority subject to the terror attack, and for many others in the wider community..

14 Brand Management
What advice could you give the Crusaders' CEO and the NZRU, in relation to issues of brand identity, brand equity, and the Crusaders name as an element of equity and identity.
Establish the pros and concerns for name retention or name change.

15 Applicability to other Managerial or Professional Situations
a What are the implications, the ethics, crisis and brand management lessons, for those involved in management and governance?
b Can the situation be seen as presenting opportunities for the key people to accept responsibility for leadership, values and ethics; to be honest, to behave appropriately, to act with compassion and to convey integrity?
c Consider whether and why an inappropriate name, inappropriate to ethic or indigenous minorities minorities - may more or less acceptable in some contexts than in others.
d Provide a Summary of Implications for Management and Governance

16 On reflection, why did the Crusaders name attract initial media attention?
In particular, comment on how and why evolving interest from different stakeholder groups kept the spotlight on the Crusaders and the New Zealand Rugby Union.

17 Should the Crusaders organisation, its CEO and senior management; have been pressured following the terror attack, because of decisions made a quarter of a century earlier?

REFERENCES

All Right (2019). https://allright.org.nz/about-us/

Arden, J. (2019aa). PM statement on Christchurch shooting – 4pm 17 March. NZ Govt, 17 March 2019. https://www.beehive.govt.nz/release/pm-statement-christchurch-shooting-%E2%80%93-4pm-17-march

Arden, J. (2019e). New Zealand bans military style semi-automatics and assault rifles. NZ Govt, 21 March 2019. Retrieved from: https://www.beehive.govt.nz/release/new-zealand-bans-military-style-semi-automatics-and-assault-rifles

Ardern (2019c). PM statement on Christchurch mosques terror attack – 18 March. Retrieved from: https://www.beehive.govt.nz/release/pm-statement-christchurch-mosques-terror-attack-%E2%80%93-18-march

Ardern, (2019b). Statement from Jacinda Ardern on Christchurch mass shooting – 3.30pm 16 March 2019. Retrieved from: https://www.beehive.govt.nz/release/statement-jacinda-ardern-christchurch-mass-shooting-%E2%80%93-330pm-16-march

Ardern, J. (2019a). Statement from Jacinda Ardern on Christchurch mass shooting – 9am 16 March 2019. Retrieved from: https://www.beehive.govt.nz/release/statement-jacinda-ardern-christchurch-mass-shooting-%E2%80%93-9am-16-march

Ardern, J. (2019d). Nationwide reflection for victims of Christchurch terror attack announced. NZ Govt, 21 March 2019. Retrieved from: https://www.beehive.govt.nz/release/nationwide-reflection-victims-christchurch-terror-attack-announced

Associated Press (2014). NFL team defends 'Redskins' name after 49 senators call for change, The Guardian, 14 Juky 2017. Retrieved from: https://www.theguardian.com/world/2014/may/24/redskins-washington-nfl-senators-letter

Cox, JW, Clement, S & Vargas, T. (2016). New poll finds 9 in 10 Native Americans aren't offended by Redskins name. *The Washington Post, May 19 2016.* Retrieved from: https://www.washingtonpost.com/local/new-poll-finds-9-in-10-native-americans-arent-offended-by-redskins-name/2016/05/18/3ea11cfa-161a-11e6-924d-838753295f9a_story.html?noredirect=on&utm_term=.6aabd2ff0678

Beehive (2019). Christchurch Call to eliminate terrorist and violent extremist online content adopted. 16 May 2019. Retrieved from: https://www.beehive.govt.nz/release/christchurch-call-eliminate-terrorist-and-violent-extremist-online-content-adopted

Cramb, A. (2009). Middlesex Crusaders cricket team changes name after complaints from Muslims and Jews. The Daily Telegraph, 2 February 2009. Rerieved from: https://www.telegraph.co.uk/sport/cricket/4434777/Middlesex-Crusaders-cricket-team-changes-name-after-complaints-from-Muslims-and-Jews.html

Cooke, H. (2019). PM Jacinda Ardern asks public to not share video or manifesto from Christchurch shootings suspect, Stuff, 15 March 2019. Retrieved from: https://www.stuff.co.nz/national/politics/111327182/pm-jacinda-ardern-asks-public-to-not-share-video-or-manifesto-from-christchurch-shootings-suspect

Cox, JW, Clement, S & Vargas, T. (2016). New poll finds 9 in 10 Native Americans aren't offended by Redskins name. *The Washington Post, May 19 2016.* Retrieved from: https://www.washingtonpost.com/local/new-poll-finds-9-in-10-native-americans-arent-offended-by-redskins-name/2016/05/18/3ea11cfa-161a-11e6-924d-838753295f9a_story.html?noredirect=on&utm_term=.6aabd2ff0678

Creative NZ, 2019). Retrieved from: https://www.creativenz.govt.nz/development-and-resources/advocacy-toolkit/case-study-all-right-campaign-christchurch

Crusaders (2019a). Statement from Crusaders CEO. 17 March 2019. Retreieved from: https://crusaders.co.nz/news?start=20

Cully, P (2019). Super Rugby: Crusaders name change will be 'particularly challenging': expert . *Stuff,* 14:53, Apr 04 2019. Retrieved from: https://www.stuff.co.nz/sport/rugby/super-rugby/111789757/super-rugby-crusaders-name-change-will-be-particularly-challenging-expert

Dickens, A. (2019). Dickens: National, Crusaders show extent of a virtue signalling society. NewsTalk ZB, 4 April 2019, 12:19p.m. Retrieved from: https://www.newstalkzb.co.nz/on-air/andrew-dickens-afternoons/opinion/andrew-dickens-national-and-crusaders-show-virtue-signalling-rampant-in-society/

Egan, B. (2019). New Zealand sports hosting changes forever with Christchurch mosque shootings. Stuff, 15 March 2019. Retrieved from: https://www.stuff.co.nz/sport/cricket/111323282/christchurch-mosque-shootings-mosque-shootings-will-alter-new-zealand-sport-security?rm=m

Fitzgerald, K. Ensor, J. & D. Gooselink. (2019). Police fork out more than $400,000 to gun owners at first buyback. Newshub, 2 September 2019. Retrieved from: https://www.newshub.co.nz/home/new-zealand/2019/07/police-fork-out-more-than-200-000-to-gun-owners-at-first-buyback.html

Gambino, L. (2014). US patent office strips Washington Redskins of 'offensive' trademarks. *The Guardian*, 18 june 2014. Retrieved from: https://www.theguardian.com/sport/2014/jun/18/washington-redskins-lose-trademark-team-us-patent-office

Garner, D. (2019). Crusaders should hold the line, and keep their name. *Stuff*, 17.01 5 April 2019. Retrieved from: https://www.stuff.co.nz/sport/rugby/super-rugby/111832155/crusaders-should-hold-the-line-and-keep-their-name

Guardian (2016). Poll finds 90% of Native Americans are OK with Washington 'Redskins' name. The Guardian, 19 May 2016. Retrieved from: https://www.theguardian.com/sport/2016/may/19/washington-redskins-name-poll-native-american-opinion

Guardian (2019). Crusaders to review Super Rugby club's name and branding after Christchurch massacre. *The Guardian* 3 April 2019. Retrieved from: https://www.theguardian.com/sport/2019/apr/03/crusaders-to-review-super-rugby-clubs-name-and-branding

Heard, G. (2019). Jacinda Ardern: We will remember the tears of our nation. Stuff, March 29 2019. Retrieved from: https://www.stuff.co.nz/national/christchurch-shooting/111645738/jacinda-ardern-we-will-remember-the-tears-of-our-nation

Heard, G. (2019). The Crusaders, 'the Naki' and the power of names. *Stuff,* 17:00, Apr 04 2019. Retrieved from: https://www.stuff.co.nz/sport/opinion/111785689/the-crusaders-the-naki-and-the-power-of-names?rm=m

Hoft, J. (2015). Susquehanna University Drops 'Crusaders' Mascot – It's Too Offensive. *The Gateway Pundit,* 5 November, 2016. Retrieved from: https://www.thegatewaypundit.com/2015/11/susquehanna-university-drops-crusaders-mascot-its-too-offensive/

Kermeen, M. (2019). Super Rugby: Seeking Muslim feedback on Crusaders name 'not fair' so soon after terror attacks. Stuff, 3 April 2019. Retrieved from: https://www.stuff.co.nz/sport/rugby/super-rugby/111750672/super-rugby-crusaders-and-nzr-engage-research-company-on-branding-dilemma?rm=m

Kitchen, T. (2019). Christchurch mosque shooting victim backs Crusaders name change. *The Press*, 3 April 2019. Retrieved from: https://www.stuff.co.nz/the-press/news/111777244/christchurch-mosque-shooting-victim-backs-crusaders-name-change?rm=a

Lowe, A. (2019). Crusaders to consider changing name and branding after terrorist attack in Christchurch. *The Times*, 3 April 2019. Retrieved from: https://www.thetimes.co.uk/article/crusaders-to-consider-changing-name-and-branding-after-terrorist-attack-in-christchurch-220fgr3rc

Mann, N. (2019). Christchurch mosque shooting: Immediate ban for military-style weapons after massacre. The Times, March 21 2019, 12:00pm. Retrieved from: https://www.thetimes.co.uk/article/christchurch-mosque-attacks-immediate-ban-for-military-style-weapons-after-massacre-bq6w5gb2b.

McConnell, G. (2019). Face of empathy: Jacinda Ardern photo resonates with the world after terror attack. Stuff, 31 March 2019. Retrieved from: https://www.stuff.co.nz/entertainment/arts/111361303/face-of-empathy-jacinda-ardern-photo-resonates-with-the-world-after-terror-attack

MFAT (2019). Christchurch Call |to eliminate terrorist and violent extremist content online. Retrieved from: https://www.christchurchcall.com/

Mintz, S. (2013). Is it an Ethical Issue whether the Washington 'Redskins' Change its Name? Ethics Sage, 17 October 2913. Retrieved from: https://www.ethicssage.com/2013/10/is-it-an-ethical-issue-whether-the-washington-redskins-change-its-name.html

Nash, J. (2019). Third reading: Arms Amendment Bill. NZ Govt, 10 April 2019. Retrieved from: https://www.beehive.govt.nz/speech/third-reading-arms-amendment-bill

Newstalk ZB (2019a). Anger over Crusaders name change won't last long. Newstalk ZB, 4 April 2019. Retrieved from: https://www.newstalkzb.co.nz/on-air/early-edition/audio/mark-hutcheson-anger-over-crusaders-name-change-wont-last-long/

Newstalk ZB (2019b). Crusaders praised for review into name and branding. *Newstalk ZB, 4 April 2019. Retrieved from:* https://www.newstalkzb.co.nz/news/national/crusaders-praised-for-review-into-name-and-branding/

Newstalk ZB (2019c). Steve Tew backs Crusaders' name change review: It's the 'right thing. *Newstalk ZB, 4 April 2019. Retrieved fromL* https://www.newstalkzb.co.nz/on-air/mike-hosking-breakfast/audio/steve-tew-backs-crusaders-name-change-review-its-the-right-thing/

Newstalk ZB (2019). Retrieved from: https://www.newstalkzb.co.nz/on-air/early-edition/audio/mark-hutcheson-anger-over-crusaders-name-change-wont-last-long/

Norquay, K. (2019). Super Rugby: The time for the Crusaders to put their name to the sword has arrived. Stuff, 14:02, Apr 04 2019. Retrieved from: https://www.stuff.co.nz/sport/opinion/111790180/super-rugby-the-time-for-the-crusaders-to-put-their-name-to-the-sword-has-arrived

NZ Herald (2019b). Super Rugby: Crusaders can survive name change, says brand expert . NZ Herald, 4 Apr, 2019 7:45pm. Retrieved from: https://www.nzherald.co.nz/sport/news/article.cfm?c_id=4&objectid=12219437

NZ Herald (2019a). NZ Rugby: Crusaders name change likely before 2020 season. *NZ Herald*, 3 April 2019. Retrieved from: https://www.newstalkzb.co.nz/news/sport/nz-rugby-crusaders-name-change-likely-before-2020-season/

One News (2019a). Petition to keep Crusaders' name approaching 25,000 signatures. *1 News*, 9 April 2019. Retrieved from: https://www.tvnz.co.nz/one-news/sport/rugby/petition-keep-crusaders-name-approaching-25-000-signatures

One News (2019b). 'Striking' new poll indicates vast majority of Kiwis want Crusaders to keep their name. 1 News, 6.20pm 20 April 2019. Retrieved from: https://www.tvnz.co.nz/one-news/new-zealand/striking-new-poll-indicates-vast-majority-kiwis-want-crusaders-keep-their-name

Peters, W. (2019). Foreign Minister Statement to the Organisation of Islamic Cooperation Emergency Meeting. NZ Govt, 22 march 2019. Retrieved from: https://www.beehive.govt.nz/speech/foreign-minister-statement-organisation-islamic-cooperation-emergency-meeting

Reuters (2019a). Rugby: What's in a Name? Crusaders Ponder Change After Christchurch Shootings. *NY Times*, 2 April 2019. Retrieved from: https://www.nytimes.com/reuters/2019/04/02/sports/football/02reuters-newzealand-shooting-crusaders.html

Reuters (2019b). Shootings. 3 April 2019. Retrieved from: https:. //www.jpost.com/International/Whats-in-a-name-Crusaders-ponder-change-after-Christchurch-shootings-585532

Reuters (2019c). Shootings. *The Jurasalem Post.* 3 April 2019. Retrieved from: https:. //www.jpost.com/International/Whats-in-a-name-Crusaders-ponder-change-after-Christchurch-shootings-585532

RNZ (2019). 'This team is all we have' - Fans' uncertainty about changing Crusaders' name. RNZ, 7 April 2019. Retrieved from:: https://www.tvnz.co.nz/one-news/sport/rugby/team-all-we-have-fans-uncertainty-changing-crusaders-name

RNZ (2019x), What happened at Christchurch Hospital on the day of the Mosque shootings, 6 May 2019, Retrieved from: https://www.rnz.co.nz/programmes/christchurch-dilemmas/story/2018693736/what-happened-at-christchurch-hospital-on-the-day-of-the-mosque-shootings

SDUT (2016). Colleges update mascots, mottos, amid pressure from students. The San Diego Union-Tribune, 18 November 2016. Retrieved from: https://www.sandiegouniontribune.com/sdut-colleges-update-mascots-mottos-amid-pressure-from-2015nov18-story.html.

Shah, B. (2019). After March 15. RNZ Nights, 20 May 2019). Retrieved from: https://www.rnz.co.nz/national/programmes/nights

Small, Z. (2019). Second tranche of gun law changes: Firearms register, tighter licencing. Newshub, 22 July 2019. Retrieved from: https://www.newshub.co.nz/home/politics/2019/07/second-tranche-of-gun-law-changes-firearms-register-tighter-licencing.html

Smith, T. (2019). Why the Crusaders are right to consult the Muslim community over name change. Stuff, 19 March 2019.. Retrieved from: https://www.stuff.co.nz/sport/rugby/opinion/111372207/why-the-crusaders-are-right-to-consult-the-muslim-community-over-name-change. Stuff (2019). Christchurch shootings: Crusaders defend their name amid concerns it could be offensive.

Stuff (2019a), Christchurch shootings: Terrorist attack at mosques, 16 March 2019. Retrieved from: https://www.stuff.co.nz/national/111313938/live-terror-attack-video-christchurch-mosque-shooting-muslims-new-zealand

Stuff. (2019b). Crusaders considering name change in wake of Christchurch shootings. Stuff, 18 March 2019. Retrieved from: https://www.stuff.co.nz/sport/rugby/super-rugby/111352359/super-rugby-crusaders-consider-new-name-in-wake-of-shooting-tragedy?rm=m

Stuff (2019c). Crusaders name, swords and horses have got to go, says Muslim advocate. Stuff, 12:50, Apr 04 2019. Retrieved from: https://www.stuff.co.nz/sport/rugby/super-rugby/111788862/crusaders-name-swords-and-horses-have-got-to-go-says-muslim-advocate

Troughton, G. (2019). Playing in overtime: why the Crusaders rugby team is right to rethink brand after Christchurch attack. *The Conversation*, 2.46pm 4 April 2019. Retrieved from: https://theconversation.com/playing-in-overtime-why-the-crusaders-rugby-team-is-right-to-rethink-brand-after-christchurch-attack-114826

TVNZ (2019b). Crusaders CEO says organisation doesn't want to put 'pressure' on Muslim community about name change. TVNZ, 2 April 2019. Retrieved from: https://www.tvnz.co.nz/one-news/sport/rugby/crusaders-ceo-says-organisation-doesnt-want-put-pressure-muslim-community-name-change

TVNZ One News (2019a). Crusaders acknowledge requests for name change in wake of Christchurch shootings. TVNZ One News, 17 March 2019. Retrieved from: https://www.tvnz.co.nz/one-news/sport/rugby/crusaders-acknowledge-requests-name-change-in-wake-christchurch-shootings

Whyte, A. (2019). Real leadership' shown from Crusaders as rugby team weighs up whether to change its name, Jacinda Ardern says. 1 News, 8 April 2019. Retrieved from: https://www.tvnz.co.nz/one-news/sport/rugby/real-leadership-shown-crusaders-rugby-team-weighs-up-whether-change-its-name-jacinda-ardern-says

DECISION SCIENCES INSTITUTE
Impact of paradoxical thinking and individual perceptions on entrepreneurial intention

Jinyoung Kim
Korea University Business School
#611 LG-POSCO Bldg, 145 Anam-ro, Seongbuk-gu, Seoul, South Korea
jinyoungkim@korea.ac.kr / +82-10-4965-5869

ABSTRACT

Individual perceptions on internal resources and external environmental factors have garnered significant research interest recently as potential antecedents of individuals' entrepreneurial intentions. Prior studies also have emphasized the needs to investigate how individuals' diverse capabilities play their roles in an individual's decision-making to start one's own business. In this study, I investigate the role of an individual's paradoxical thinking capability in terms of its moderating effect on the relationship between individual perceptions, i.e., perceived individual capability, perceived opportunity, social perception of entrepreneurship, and perceived fear of failure, and subsequent entrepreneurial intentions. Analyses of survey responses of 10,000 individuals in South Korea reveal that paradoxical thinking positively moderates the effects of perceived capability, perceived opportunity, and fear of failure on entrepreneurial intention. On the other hand, paradoxical thinking negatively moderates the relationship between social perception of entrepreneurship and entrepreneurial intention. These findings suggest that while paradoxical thinking positively affects individuals' perceived feasibility, it simultaneously decreases individuals' perceived desirability because paradoxical thinking expands individuals' options to choose from, which, as a result, would decrease the desirability of an option of starting one's own business.

KEYWORDS: entrepreneurial intention, paradoxical thinking, individual perception

INTRODUCTION

Entrepreneurial activities have been gathering universal interests from various stakeholders that range from aspiring entrepreneurs in an individual level to governments in a societal level. These interests peak especially when economies are in bad situations such as financial crises, largely due to the distinguished impact of entrepreneurship: entrepreneurial activities are one of rare phenomena where micro-level activities have macro-level impacts on society as a whole (Davidsson, 2004). Understanding how and why these micro-level activities happen, even under the harsh environment of high failure rates, is important for both individuals who try to create new organizations and society as a whole which would benefit from those individuals who successfully initiate entrepreneurial processes to deliver their new ideas.

One observable indicator that has been found to be a single best predictor of behavior is an individual's intention (Kim & Hunter, 1993). Since the actual behavior is either rare or difficult to observe (Ajzen, 1991), prior studies have observed an individual's entrepreneurial intentions to understand the possible sources of one's decision to start a business. Prior studies have investigated entrepreneurial decision-making processes and organizational emergence from various perspectives, such as psychological, cognitive, behavioral, and strategic standpoints. However, our understanding of the possible origins of these micro-level activities is hardly complete, mostly because individual decision-makings and subsequent organizational emergence involve a series of perception-driven decisions (Bird, 1988; Gartner, 1985; Katz &

Gartner, 1988; Shapero & Sokol, 1982). Individuals' social and situational perceptions, in conjunction with their individual characteristics, lead to one's decision to start a new business. Naturally, prior studies have discussed individual perceptions of entrepreneurs in terms of risk perceptions (e.g., Gartner & Liao, 2012; Simon, Houghton, & Aquino, 2000), perceived opportunities (e.g., Krueger, 1993), or perceived environmental uncertainties (e.g., Milliken, 1987).

In this study, individual perceptions and their subsequent effects on their entrepreneurial intentions are investigated in terms of individual capability of thinking in a paradoxical way. Specifically, I focus on the individuals' paradoxical thinking capability that handles both perseverance and strategic flexibility with equal dexterity, and study the impact of this paradoxical thinking capability on the relationship between individual perceptions and their subsequent entrepreneurial intentions.

LITERATURE REVIEW

Intentionality is defined as 'a state of mind directing a person's attention toward a specific object or a path in order to achieve something' (Bird, 1988, p. 442). This rather broad definition of intentionality has been applied to and understood in the context of an entrepreneur who pursues entrepreneurial ideas. Entrepreneurial intention is an individual desire to become an entrepreneur (Jenkins & Johnson, 1997), more specifically defined as a self-acknowledged conviction about the likelihood of starting one's own business at some point in the future (Crant, 1996; Thompson, 2009). Prior studies that have focused on entrepreneurial intention have identified a wide range of possible antecedents: personality traits such as proactiveness (Crant, 1996; Zhao, Seibert, & Lumpkin, 2010), behavioral traits such as propensity to act (Krueger, Reilly, & Carsrud, 2000), individual predispositions such as proclivity for improvisation (Hmieleski & Corbett, 2006) or self-efficacy (Boyd & Vozikis, 1994; Piperopoulos & Dimov, 2015; Zhao, Seibert, & Hills, 2005), individual exposure to entrepreneurship (Bae, Qian, Miao, & Fiet, 2014; Krueger, 1993), or general individual factors such as gender or culture (Shinnar, Giacomin, & Janssen, 2012), to name a few.

According to the review article on entrepreneurial decision-making by Shepherd, Williams, and Patzelt (2015), perception-based variables have garnered more research interests recently mostly because of the subjective nature of individual perceptions, intentions, and subsequent actions as a form of decision-making. More specifically, a common theme that underlie the identified antecedents of entrepreneurial intention is how an individual perceives oneself as well as the surroundings of oneself. Endogenously, individuals perceive their resources (Kor, Mahoney, & Michael, 2007), capabilities (Gatewood, Shaver, & Gartner, 1995), identity (Fauchart & Gruber, 2011), personal motivations (Birley & Westhead, 1994), and potential goals. Simultaneously, individuals perceive exogenous factors such as environmental opportunities and threats (Krueger, 2003), social norms , and the level of uncertainties (Townsend & Hart, 2008). Constant interactions between perceptions of endogenous and exogenous factors are processed to form an individual's overall perceptions. Inherent heterogeneity in individuals' self-perceptions and environmental perceptions, consequently, drives the variance in the level of entrepreneurial intentions.

Prior studies in entrepreneurial intention have identified two critical dimensions of individual perceptions which impact one's entrepreneurial intention: perceived desirability and feasibility of a potential opportunity (Fitzsimmons & Douglas, 2011; Krueger et al., 2000). Perceived desirability is the degree to which an individual is attracted to an idea of the recognized opportunity. Perceived feasibility is the degree to which an individual feels capable of oneself to pursue the opportunity. These two specific types of perception have been empirically shown to have positive impacts on one's intention to start a business (Krueger et al., 2000). Prior studies have proposed various ways to conceptualize and measure the perceived feasibility and desirability constructs: a single-item measure such as 'how feasible or desirable' an entrepreneurial opportunity is (Krueger et al., 2000); self-efficacy as a perceived feasibility measure and perceived attractiveness of certain hypothetical career profiles as a perceived desirability measure (Fitzsimmons & Douglas, 2011); or more specifically, an entrepreneur's attitude toward risks as a proxy to measure one's perceived desirability (McMullen & Shepherd, 2006). Although both endogenous and exogenous factors which have been identified as possible antecedents of entrepreneurial intention as summarized above, empirical studies has not been incorporating both endogenous and exogenous factors so far. Therefore, in this study, both endogenous and exogenous factors are investigated to further understand the perceived feasibility and desirability.

Endogenously individuals perceive their ability to accomplish their intentions, and, at the same time, exogenously, they perceive entrepreneurial opportunities in their surroundings. Following the definition of the perceived feasibility, i.e., the degree to which an individual feels capable of oneself to pursue the opportunity, I focus on the perceived individual capability to pursue entrepreneurial opportunities as the endogenous factor, and the perceived availability of entrepreneurial opportunities as the exogenous factor of the perceived feasibility. Models and theories that explain the entrepreneurial intention model, such as Ajzen's (1991) theory of planned behavior or Shapero and Sokol's (1982) entrepreneurial event model, argue the positive impact of the perceived feasibility and subsequent entrepreneurial intention. Subsequent studies that have empirically tested the models show the supporting evidence of the positive relationship between the perceived feasibility and entrepreneurial intention (e.g., Krueger & Carsrud, 1993; Kautonen, van Gelderen, & Fink, 2015). Hence I hypothesize the positive relationships between the perceived capability as well as the perceived availability of entrepreneurial opportunity and the entrepreneurial intention.

Hypothesis 1. An individual's perceived capability to pursue entrepreneurial opportunities has a positive influence on his or her entrepreneurial intention.

Hypothesis 2. An individual's perceived availability of entrepreneurial opportunities has a positive influence on his or her entrepreneurial intention.

Following the definition of the perceived desirability, i.e., the degree to which an individual is attracted to an idea of the recognized opportunity, I focus on an individual's fear of failure as the endogenous factor, and the social perception of entrepreneurship in general as the exogenous factor of the perceived desirability. Fear of failure is the perceived possibility of experiencing failure and its consequences (Arenius & Minniti, 2005), which should be distinguished from a general risk aversion (Wyrwich, Stuetzer, & Sternberg, 2016). Fear of failure is a context-specific process which is affected by an individual's personal assessment and understanding of one's possibility and consequences of accomplishing his or her aspiration. Since each individual holds heterogeneous level of fear towards the possibility of failure, how individuals perceive their own fear of failure contributes to their intentions towards pursuing an entrepreneurial opportunity. However, an individual's endogenous fear of failure is not the only factor that affects one's assessment of desirability of the opportunity. Social perception of entrepreneurship in general, such as how the society where an individual belongs to views entrepreneurs and entrepreneurial processes, also affect an individual's aspiration to pursuing one's recognized entrepreneurial opportunity. Prior studies on social norms have found that social norms indeed affect an individual's entrepreneurial intention (e.g., Krueger & Carsrud, 1993; Kautonen, Tornikoski, & Kibler, 2011). Hence, I hypothesize that an individual's fear of failure negatively affects the entrepreneurial intention, whereas the positive social perception of entrepreneurship in general positively affects the entrepreneurial intention.

Hypothesis 3. An individual's fear of failure has a negative influence on one's entrepreneurial intention.

Hypothesis 4. An individual's social perception of entrepreneurship in general has a positive influence on one's entrepreneurial intention.

These hypothesized relationships between individual perceptions and entrepreneurial intentions are likely to be contingent upon individual characteristics. Prior studies that have focused on the entrepreneurial intention trace back to a few "epicenter" studies (e.g., Ajzen, 1991; Shapero & Sokol, 1982), one of which is Bird's (1988) intentionality model. The intentionality model conceptualizes a wide range of both individual and environmental attributes, e.g., personal history, personality, social and political contexts, that are potential antecedents to intentionality. While the intentionality model is fundamentally a behavioral model, Bird (1988) identified two distinctively different types of cognitive processes: (1) a rational, analytical, and cause-effect thinking process, and (2) an intuitive, holistic, and contextual thinking process. Although individual and environmental attributes have been investigated extensively so far (e.g., Boyd & Vozikis, 1994; Crant, 1996; Mueller & Thomas, 2001), two competing cognitive processes have been relatively under-investigated, mostly due to the inherent difficulties with measuring and collecting an individual's cognitive processes.

Bird (1988) argues that two competing cognitive processes, i.e., the analytical thinking which searches and processes information in a systematic and comprehensive way, and the intuitive thinking which searches and processes information in a heuristic and satisficing way, are equally critical for entrepreneurial intention. Moving away from the conventional conceptualization of "either/or" model, Bird's (1988) intentionality model implicitly assumes the presence of paradoxical thinking. Paradoxical thinking is an individual's ability to juxtapose and integrate contradictions, such as actively processing opposite ideas or concepts simultaneously (Rothenberg, 1979). Paradoxes, which are defined as "contradictory yet interrelated elements that exist simultaneously and persist over time" (Smith & Lewis, 2011, p. 382), are prevalent in our everyday decision-makings: we try to be efficient, yet flexible at the same time; we aspire to have "big picture" visions, yet be detail-oriented at the same time. Embracing and effectively handling, not avoiding or eliminating, paradoxes is a common task for individuals in the field of creativity (Rothenberg, 1979) or innovation (Martin, 2007), where a creative resolution is not only generated by processing opposing ideas but also superior to each idea (Ingram, Lewis, Barton, & Gartner, 2016). Logically challenges of paradoxes become more significant for individuals who aspire to be entrepreneurs, as their decision-makings need to accommodate and embrace competing yet interrelated concerns.

Departure from the formal "either/or" logic to adopt "both/and" one enables individuals to acknowledge the paradoxical nature of entrepreneurial ideas which require entrepreneurs to be efficient with limited resources, yet, at the same time, flexible under the high level of uncertainties they constantly face throughout their entrepreneurial journeys. More specifically, resource constraints require entrepreneurs to be efficient and effective in combining existing resources to create new resource bundles (Baker & Nelson, 2005). Concurrently, high level of uncertainty imposes a critical challenge to entrepreneurs: they need to be adaptive to environmental contingencies, and moreover, proactive with exploiting the recognized contingencies in a beneficial way (Sarasvathy, 2001). As a result, entrepreneurs are required to handle seemingly intractable paradoxes of systematically executing tactical movements efficiently while flexibly accommodating to future possibilities. These types of paradoxical tensions demand paradoxical thinking because an individual capability to process contradictory problems and generate superior solutions is more likely to show superior performance. Individuals who are capable of conducting paradoxical thinking, therefore, are expected to handle the paradoxical nature of pursuing entrepreneurial ideas in a superior way. Paradoxical thinking enables individuals to better process contradictory aspects of entrepreneurial ideas, and consequently, assists individuals to realize their entrepreneurial intention through perceived feasibility of their ideas. In other words, the individual capability of paradoxical thinking is likely to facilitate the process of transforming perceived opportunities and individual capability into individuals' intentions to pursue their ideas.

Hypothesis 5a. Paradoxical thinking positively moderates the relationship between the perceived individual capability and entrepreneurial intention.

Hypothesis 5b. Paradoxical thinking positively moderates the relationship between the perceived opportunities and entrepreneurial intention.

Individuals who are adept at paradoxical thinking, on the other hand, should be able to generate greater number of potentially desirable options to choose from, which may make an option of starting one's own business less desirable.

Hypothesis 5c. Paradoxical thinking negatively moderates the relationship between the social perception of entrepreneurship and entrepreneurial intention.

Hypothesis 5d. Paradoxical thinking negatively moderates the relationship between the fear of failure and entrepreneurial intention.

METHODS

In this study, I analyzed the nation-wide, individual-level Entrepreneurship Survey of South Korea 2017. The Korean Entrepreneurship Foundation conducts an annual survey to measure entrepreneurial aspects of general public in South Korea, by collecting survey responses from 10,000 South Korean individuals between August and October, 2017. Demographic profile and entrepreneurship-related experiences of survey respondents are summarized in Table 1.

Table 1. Survey Respondents

Gender	Male	5,055
	Female	4,945
Age	13-19	1,401
	20-29	1,651
	30-39	1,754
	40-49	1,836
	50-59	1,824
	60-69	1,534
Entrepreneurial experience	Yes	1,127
	No	8,873
Entrepreneurial education	Yes	4,672
	No	5,328

The dependent variable of entrepreneurial intention is measured by a survey item, "I have a plan to start my own business within three years." The perceived individual capability is measured by a survey item, "I have relevant knowledge and capability to start my own business." The perceived opportunity is measured by a survey item, "I believe abundant opportunities are available for me to start my own business." The social perception of entrepreneurship is measured by three survey items: (1) "Entrepreneurs are respected and well regarded," (2) "Becoming an entrepreneur is considered a good career choice," and (3) "Stories of successful entrepreneurs are readily available in media." Individual fear of failure is measured by a survey item, "Fear of failure makes me hesitant to start my own business."

Following the prior studies that create a multiplicative measure to capture a construct which contains two contradictory aspects, such as organizational ambidexterity measure (Gibson & Birkinshaw, 2004; Jansen, Simsek, & Cao, 2012; Kauppila & Tempelaar, 2016), I have created a multiplicative measure by multiplying two survey items, which measure (1) an individual's mindset to persevere through challenges without compromising, and (2) an individual's mindset to pivot, or adapt, flexibly according to environmental contingencies. This multiplicative interaction between individual perseverance and pivot adequately capture an individual's "both/and" thinking of planning and changing. Perseverance is measured by a survey item, "I am inclined stick to my original plan and achieve my goal regardless of challenges." Pivot is measured by a survey item, "I am inclined to change my plan and try new paths when I face challenges."

All variables are measured by a single survey item except the social perception of entrepreneurship, as they are theoretically simple and concrete constructs. Although single-item measures are commonly considered to show lower content validity, sensitivity, and reliability (McIver & Carmines, 1981), this methodological issue is prominent with complex and multidimensional constructs. When the construct is relatively simple and concrete, single-item scales have been argued to be appropriate measures (Freed, 2013; Grapentine, 2001). Accordingly, prior studies empirically show that single-item scales that measure theoretically simple constructs show identical performance with multi-item scales (Bergkvist & Rossiter, 2007; McKenzie & Marks, 1999) or even superior performance than multi-item scales (Bush et al., 2010).

In order to control for individual characteristics and entrepreneurial experiences, six control variables are added: age, current income, personal experience in entrepreneurship, personal experience in entrepreneurial education, family members' experience in entrepreneurship, and one's perceived creativity.

RESULT

Table 2. Means, standard deviations, and correlations among variables (N=10,000)

	Mean	S.D.	1	2	3	4	5	6	7	8	9	10
1. Entrepreneurial intention	3.02	1.66										
2. Entrepreneurial experience	.11	.32	.192***									
3. Entrepreneurial education	.47	.50	.056***	.220***								
4. Age	3.56	1.64	.067***	.227***	.117***							
5. Income	1.75	.79	.152***	.267***	.112***	.290***						
6. Creativity	3.82	1.42	.179***	.030***	.009	-.102***	.028**					
7. Perceived capability	3.46	1.53	.549***	.212***	.035***	.081***	.170***	.149***				
8. Perceived opportunity	3.72	1.43	.395***	.115***	.061***	.049***	.093***	.211***	.374***			
9. Social perception of entrepreneurship	4.51	1.00	.189***	.065***	.048***	-.051***	.051***	.196***	.188***	.185***		
10. Fear of failure	4.52	1.47	-.071***	-.015	.017	-.024*	.002	.052***	-.018	.047***	.364***	
11. Paradoxical thinking	16.49	5.46	.132***	-.005	-.012	-.007	.022*	.237***	.154***	.140***	.189***	.050***

**** p<.001 ** p<.01 * p<.05*

Table 2 displays means, standard deviations, and correlations among all variables. Table 3 shows the results of the hierarchical regression analysis. As the first step of the hierarchical regression analysis, Model 1 contains controls. Model 2 includes the main effects of individual perception, i.e., perceived capability, perceived opportunity, social perception of entrepreneurship, and fear of failure, and paradoxical thinking. Model 3 includes the interactions between individual perceptions and paradoxical thinking.

The analysis result of Model 2 show that the hypotheses regarding the direct effects of individual perception on entrepreneurial intention are supported. The change in R2 from Model

1 to Model 2 for the main-effect model indicates a significant increase in R2 above and beyond the control variables. The analysis result of Model 3 show that the interaction effects between individual perception and paradoxical thinking are indeed statistically significant as hypothesized. Figure 1 shows the interaction effects between individual perception and paradoxical thinking on entrepreneurial intention.

Table 3. Hypothesis testing using hierarchical regression analysis

	Model 1	Model 2	Model 3
Entrepreneurial experience	.154***	.055***	.061***
Entrepreneurial education	.007	.008	.008
Age	.020	.006	.009
Income	.099***	.037***	.035***
Creativity	.173***	.058***	.054***
Perceived capability (PC)		.427***	.351***
Perceived opportunity (PO)		.200***	.143***
Social perception (SP)		.095***	.170***
Fear of failure (FF)		-.109***	-.397***
Paradoxical thinking (PT)		.013	-.290***
PC x PT			.105***
PO x PT			.083***
SP x PT			-.152***
FF x PT			.443***
R^2	.077	.367	.377
R^2 change		.290	.010
F	167.172***	578.520***	430.894***

**** p<.001   ** p<.01   * p<.05*

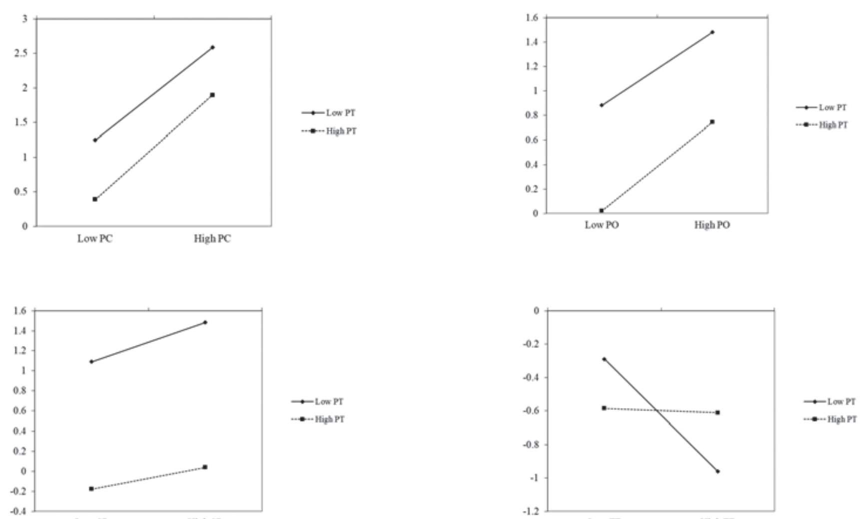


Figure 1. Interaction effects

DISCUSION

In this study, I investigate the role of an individual's paradoxical thinking capability in terms of its moderating effect on the relationship between individual perceptions, i.e., perceived individual capability, perceived opportunity, social perception of entrepreneurship, and perceived fear of failure, and subsequent entrepreneurial intentions. Analyses of survey responses of 10,000 individuals in South Korea reveal that paradoxical thinking positively moderates the effects of perceived capability, perceived opportunity, and fear of failure on entrepreneurial intention. On the other hand, paradoxical thinking negatively moderates the relationship between social perception of entrepreneurship and entrepreneurial intention. These findings suggest that while paradoxical thinking positively affects individuals' perceived feasibility, it simultaneously decreases individuals' perceived desirability because paradoxical thinking expands individuals' options to choose from, which, as a result, would decrease the desirability of an option of starting one's own business.

REFERENCE

Ajzen, I. 1991. The theory of planned behavior. ***Organizational Behavior and Human Decision Processes***, 50(2): 179-211.

Arenius, P., & Minniti, M. 2005. Perceptual variables and nascent entrepreneurship. ***Small Business Economics***, 24(3): 233-247.

Bae, T. J., Qian, S., Miao, C., & Fiet, J. O. 2014. The relationship between entrepreneurship education and entrepreneurial intentions: a meta-analytic review. ***Entrepreneurship Theory and Practice***, 38(2): 217-254.

Baker, T., & Nelson, R. E. 2005. Creating something from nothing: resource construction through entrepreneurial bricolage. ***Administrative Science Quarterly***, 50(3): 329-366.

Bergkvist, L., & Rossiter, J. R. 2007. The predictive validity of multiple-item versus single-item measures of the same constructs. ***Journal of Marketing Research***, 44(2): 175-184.

Bird, B. 1988. Implementing entrepreneurial ideas: the case for intention. ***Academy of Management Review***, 13(3): 442-453.

Birley, S., & Westhead, P. 1994. A taxonomy of business start-up reasons and their impact on firm growth and size. ***Journal of Business Venturing***, 9(1): 7.

Boyd, N. G., & Vozikis, G. S. 1994. The influence of self-efficacy on the development of entrepreneurial intentions and actions. ***Entrepreneurship Theory and Practice***, 18(4): 63-77.

Bush, S. H., Parsons, H. A., Palmer, J. L., Li, Z., Chacko, R., & Bruera, E. 2010. Single- vs. multiple-item instruments in the assessment of quality of life in patients with advanced cancer. ***Journal of Pain and Symptom Management***, 39(3): 564-571.

Crant, J. M. 1996. The proactive personality scale as a predictor of entrepreneurial intentions. ***Journal of Small Business Management***, 34(3): 42.

Davidsson, P. 2004. ***Researching Entrepreneurship***. New York, NY: Springer.

Fauchart, E., & Gruber, M. 2011. Darwinians, communitarians, and missionaries: the role of founder identity in entrepreneurship. ***Academy of Management Journal***, 54(5): 935-957.

Fitzsimmons, J. R., & Douglas, E. J. 2011. Interaction between feasibility and desirability in the formation of entrepreneurial intentions. ***Journal of Business Venturing***, 26(4): 431-440.

Freed, L. 2013. ***Innovating Analytics: How the Next Generation of Net Promoter Can Increase Sales and Drive Business Results***: John Wiley & Sons.

Gartner, W., & Liao, J. 2012. The effects of perceptions of risk, environmental uncertainty, and growth aspirations on new venture creation success. ***Small Business Economics***, 39(3): 703-712.

Gartner, W. B. 1985. A conceptual framework for describing the phenomenon of new venture creation. ***Academy of Management Review***, 10(4): 696-706.

Gatewood, E. J., Shaver, K. G., & Gartner, W. B. 1995. A longitudinal study of cognitive factors influencing start-up behaviors and success at venture. ***Journal of Business Venturing***, 10(5): 371.

Gibson, C. B., & Birkinshaw, J. 2004. The antecedents, consequences, and mediating role of organizational ambidexterity. ***Academy of Management Journal***, 47(2): 209-226.

Grapentine, T. H. 2001. A practitioner’s comment on Aimee L. Drolet and Donald G. Morrison’s “Do we really need multiple-item measures in service research?”. ***Journal of Service Research***, 4(2): 155-158.

Hmieleski, K. M., & Corbett, A. C. 2006. Proclivity for improvisation as a predictor of entrepreneurial intentions. ***Journal of Small Business Management***, 44(1): 45-63.

Ingram, A. E., Lewis, M. W., Barton, S., & Gartner, W. B. 2016. Paradoxes and innovation in family firms: the role of paradoxical thinking. ***Entrepreneurship Theory and Practice***, 40(1): 161-176.

Jansen, J. J., Simsek, Z., & Cao, Q. 2012. Ambidexterity and performance in multiunit contexts: cross‐level moderating effects of structural and resource attributes. ***Strategic Management Journal***, 33(11): 1286-1303.

Jenkins, M., & Johnson, G. 1997. Entrepreneurial intentions and outcomes: a comparative causal mapping study. ***Journal of Management Studies***, 34(6): 895-920.

Katz, J., & Gartner, W. B. 1988. Properties of emerging organizations. ***Academy of Management Review***, 13(3): 429-441.

Kauppila, O. P., & Tempelaar, M. P. 2016. The social‐cognitive underpinnings of employees’ ambidextrous behaviour and the supportive role of group managers’ leadership. ***Journal of Management Studies***, 53(6): 1019-1044.

Kautonen, T., Tornikoski, E. T., & Kibler, E. 2011. Entrepreneurial intentions in the third age: the impact of perceived age norms. ***Small Business Economics***, 37(2): 219-234.

Kautonen, T., van Gelderen, M., & Fink, M. 2015. Robustness of the theory of planned behavior in predicting entrepreneurial intentions and actions. ***Entrepreneurship Theory and Practice***, 39(3): 655-674.

Kim, M. S., & Hunter, J. E. 1993. Attitude‐behavior relations: a meta‐analysis of attitudinal relevance and topic. ***Journal of Communication***, 43(1): 101-142.

Kor, Y. Y., Mahoney, J. T., & Michael, S. C. 2007. Resources, capabilities and entrepreneurial perceptions. ***Journal of Management Studies***, 44(7): 1187-1212.

Krueger, N. 1993. The impact of prior entrepreneurial exposure on perceptions of new venture feasibility and desirability. ***Entrepreneurship Theory and Practice***, 18(1): 5-21.

Krueger, N. F. 2003. The cognitive psychology of entrepreneurship. In Z. J. Acs, & D. B. Audretsch (Eds.), ***Handbook of Entrepreneurship Research***: 105-140. New York, NY: Springer.

Krueger, N. F., Reilly, M. D., & Carsrud, A. L. 2000. Competing models of entrepreneurial intentions. ***Journal of Business Venturing***, 15(5-6): 411-432.

Martin, R. L. 2007. ***The Opposable Mind: How Successful Leaders Win Through Integrative Thinking***.

McIver, J. P., & Carmines, E. G. 1981. ***Undimensional Scaling***: Sage.

McKenzie, N., & Marks, I. 1999. Quick rating of depressed mood in patients with anxiety disorders. ***British Journal of Psychiatry***, 174: 266-269.

McMullen, J. S., & Shepherd, D. A. 2006. Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. ***Academy of Management Review***, 31(1): 132-152.

Milliken, F. J. 1987. Three types of perceived uncertainty about the environment: state, effect, and response Uncertainty. ***Academy of Management Review***, 12(1): 133-143.

Mueller, S. L., & Thomas, A. S. 2001. Culture and entrepreneurial potential: a nine country study of locus of control and innovativeness. ***Journal of Business Venturing***, 16(1): 51-75.

Piperopoulos, P., & Dimov, D. 2015. Burst bubbles or build steam? Entrepreneurship education, entrepreneurial self-efficacy, and entrepreneurial intentions. ***Journal of Small Business Management***, 53(4): 970-985.

Rothenberg, A. 1979. ***The Emerging Goddess: the Creative Process in Art, Science, and Other Fields***. Chicago: University of Chicago Press.

Sarasvathy, S. D. 2001. Causation and effectuation: toward a theoretical shift from economic inevitability to entrepreneurial contingency. ***Academy of Management Review***, 26(2): 243-263.

Shapero, A., & Sokol, L. 1982. The social dimensions of entrepreneurship. In C. Kent, D. L. Sexton, & K. H. Vesper (Eds.), ***The Encyclopedia of Entrepreneurship***: 72-90.

Shepherd, D. A., Williams, T. A., & Patzelt, H. 2015. Thinking about entrepreneurial decision making: review and research agenda. ***Journal of Management***, 41(1): 11-46.

Shinnar, R. S., Giacomin, O., & Janssen, F. 2012. Entrepreneurial perceptions and intentions: the role of gender and culture. ***Entrepreneurship Theory and Practice***, 36(3): 465-493.

Simon, M., Houghton, S. M., & Aquino, K. 2000. Cognitive biases, risk perception, and venture formation: how individuals decide to start companies. ***Journal of Business Venturing***, 15(2): 113-134.

Smith, W. K., & Lewis, M. W. 2011. Toward a theory of paradox: A dynamic equilibrium model of organizing. ***Academy of Management Review***, 36(2): 381-403.

Thompson, E. R. 2009. Individual entrepreneurial intent: construct clarification and development of an internationally reliable metric. ***Entrepreneurship Theory and Practice***, 33(3): 669-694.

Townsend, D. M., & Hart, T. A. 2008. Perceived institutional ambiguity and the choice of organizational form in social entrepreneurial ventures. ***Entrepreneurship Theory and Practice***, 32(4): 685-700.

Wyrwich, M., Stuetzer, M., & Sternberg, R. 2016. Entrepreneurial role models, fear of failure, and institutional approval of entrepreneurship: a tale of two regions. ***Small Business Economics***, 46(3): 467-492.

Zhao, H., Seibert, S. E., & Hills, G. E. 2005. The mediating role of self-efficacy in the development of entrepreneurial intentions. ***Journal of Applied Psychology***, 90(6): 1265-1272.

Zhao, H., Seibert, S. E., & Lumpkin, G. T. 2010. The relationship of personality to entrepreneurial intentions and performance: a meta-analytic review. ***Journal of Management***, 36(2): 381-404.

DECISION SCIENCES INSTITUTE

Cross-Efficiency Evaluation-Based Multiple Criteria Data Envelopment Analysis Approach to Humanitarian Supply Chain Network Design

Jae-Dong Hong
Industrial Engineering
South Carolina State University
Orangeburg, SC 29117
Email: jhong@scsu.edu

ABSTRACT

This paper proposes a new procedure of integrating data envelopment analysis (DEA), multiple criteria data envelopment analysis (MCDEA), multiple objective programming (MOP), and the cross-efficiency (CE) method to evaluate and rank the decision-making units (DMUs). Through a case study of designing the humanitarian supply chain network (HTSCN) system, we demonstrate that the integrated method performs well in terms of designing HTSCN system effectively and efficiently.

KEYWORDS: Multiple Criteria Data Envelopment Analysis, Cross Efficiency Evaluation, Multiple Objective Programming, Humanitarian Supply Chain Network

INTRODUCTION

Data envelopment analysis (DEA) is one of the methodologies that have been widely used to evaluate the relative efficiency of a set of peer organizations called decision-making units (DMUs) that use multiple inputs to produce multiple outputs. The DMUs in DEA to be assessed should be relatively homogeneous. As the whole technique is based on a comparison of each DMU with all the remaining ones, a considerable large set of DMUs is necessary for the assessment to be meaningful (Meza & Jeong, 2013; Ramanathan, 2006). DEA eventually determines which of the DMUs make efficient use of their inputs and produce most outputs and which do not. DEA produces a single, comprehensive measure of performance for each DMU. The best ratio among all the DMUs would identify the most efficient DMU, and every other DMU would be rated by comparing its ratio to the best one. However, a weakness of the DEA-based assessment is that a considerable number of DMUs out of the set of DMUs to be rated are classified as efficient so that it may suffer from a lack of discrimination particularly.

To improve the discriminating power of DEA methods, Li and Reeves (1999) propose a multiple criteria DEA (MCDEA) model under the framework of multiple objective linear programming (MOLP). The MCDEA model involves a broader definition of relative efficiency that the classical one introduced by Charnes et al. (1978). In other words, several different efficiency measures are defined under the same constraints. Each measure serves as a criterion to be optimized. They (1999) claim that efficiency criteria that are more restrictive than the classical one will yield fewer efficient DMUs and will allow less flexibility for input/output weight distribution. But, they (1999) do not explain how to rank DMUs but merely show which DMUs are more efficient than other DMUs by solving the MCDEA sequentially with one objective out of three objective functions. For example, a DMU, which is consistently efficient regardless of the objective function, is considered to be most efficient. Then, if several DMUs turn out to be consistently efficient after evaluated by MCDEA, the question of how to rank them in terms of efficiency remains.

DEA allows each DMU to be evaluated with its most favorable weights due to its nature of the self-evaluation. Thus, to maximize the self-efficiency, the DEA model may even ignore unfavorable inputs/outputs. That is the inherent deficiency stemming from the self-efficiency. To remedy this deficiency, the cross-efficiency (CE) evaluation method is suggested by Sexton et al. (1986) as a DEA extension to rank DMUs with the main idea of using DEA to do the peer evaluation, rather than DEA's pure self-evaluation. Due to its enhanced discriminating power, the CE evaluation has found a significant number of applications in the DEA literature (Wang and Chin, 2010).

To remedy the weaknesses/deficiencies of DEA and MCDEA, this paper proposes a new procedure of integrating MCDEA with CE evaluation method for a humanitarian supply chain network (HTSCN) design problem in a pre-disaster scenario. The HTSCN design problem consists of finding the optimal emergency response facility (ERF) locations and allocation scheme of humanitarian supplies through ERFs, where all ERFs are under the risk of disruptions. For a case study, we use Hong and Jeong's (2018) HTSCN design problem to demonstrate the applicability of the proposed procedure. This paper demonstrates that the innovative integrated procedure can be applied for various supply chain network design problems with multiple performance measures.

MULTIPLE CRITERIA DEA MODEL

Charnes et al. (1978) establish a CRS (Constant Returns to Scale) m-DEA model to find an efficiency score for DMU_k, which is formulated as the following LP problem:

$$Max\, h_k = \sum_{r=1}^{s} u_r y_{rk}\,, \tag{1}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ik} = 1, \tag{2}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \forall j, \tag{3}$$

$$u_r, v_i \ \ge 0, r = 1 \dots, s; i = 1 \dots, m$$

where
Ω = number of DMUs being compared in the DEA analysis
h_k = efficiency rating of the DMU_k being evaluated by DEA
y_{rj} = amount of output r generated by DMU_j , *j= 1, 2, …, k, …* Ω
x_{ij} = amount of input i used by DMU_j
i = number of inputs used by the DMUs
r = number of outputs generated by the DMUs
u_r = coefficient or weight assigned by DEA to output r
v_i = coefficient or weight assigned by DEA to input i

Li and Reeves (1999) propose the following MCDEA model using the LP model in (1)-(3):

$$Max\ h_k = \sum_{r=1}^{s} u_r y_{rk}, \tag{4}$$

$$Min\ M_k, \tag{5}$$

$$Min\ D_k = \sum_j d_j \tag{6}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ik} = 1, \tag{7}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j \leq 0, \qquad \forall j, \tag{8}$$

$$M_k - d_j \geq 0, \forall j\ , \tag{9}$$

$$u_r, v_i \ \geq 0, r = 1 \ldots, s; i = 1 \ldots, m; \text{ and } d_j \geq 0, j = 1,\ 2,\ \ldots,\ k,\ \ldots,\ \Omega.$$

Note that the three performance measures for each DMU would be classified as one output, h_k, to be maximized and two inputs, M_k and D_k, to be minimized. Thus, this paper proposes a multiple objective programming (MOP) model to MCDEA to combine these three performance measures into one objective. Solving the MCDEA model by MOP for a given set of weight set assigned to each performance measure would yield one DMU with two inputs, M_k and D_k, and one output, h_k. Thus, regardless of the number of performance measures for a supply chain system, solving the MCDEA by MOP will yield DMUs with two inputs and one output, so that the cross-efficiency DEA method can be applied effectively to rank the multiple efficient DMUs.

We propose the following procedure of integrating method with DEA, MCDEA, MOP, and CE method:

(i) Apply DEA to identify efficient DMUs.
(ii) For those efficient DMUs identified by applying DEA, apply MCDEA model and solve it by MOP with various sets of weights assigned one output and two inputs
(iii) Apply CE DEA method to rank those efficient DMUs.

HUMANITARIAN SUPPLY CHAIN NETWORK

Humanitarian supply chain (HSC) plays a critical role in providing disaster relief items such as first aids, drinking water, food, and daily commodities to alleviate the suffering of people. In 2017, the U.S. experienced a historic year of weather and climate disasters. In total, the U.S. was impacted by 16 separate billion-dollar disaster events including three tropical cyclones, eight severe storms, two inland floods, a crop freeze, drought, and wildfire. During 2018, the U.S. experienced a very active year of weather and climate disasters (see Figure 1). In total, the U.S.

was impacted by 14 separate billion-dollar disaster events: two tropical cyclones, eight severe storms, two winter storms, drought, and wildfires. The past three years (2016-2018) have been historic, with the annual average number of billion-dollar disasters being more than double the long-term average. The number and cost of disasters are increasing over time due to a combination of increased exposure, vulnerability, and the fact the climate change is increasing the frequency of some types of extremes that lead to billion-dollar disasters (https://www.climate.gov/news-features/blogs/beyond-data/2018s-billion-dollar-disasters-context). In early 2019, a wild March headlined by abnormal warmth in the US's coldest state and a destructive and costly "bomb cyclone" in the central US. By a landslide, Alaska posted its warmest March on record, and the powerhouse storm in the central US became the second billion-dollar weather disaster of 2019.

In this respect, the HTSCN design has become an important strategic decision due to the major damage inflicted by several natural disaster events. The weather-related emergencies have brought issues of emergency relief planning again. Indeed, after emergencies, it is critical through emergency response facilities (ERFs) to distribute humanitarian aid to the affected areas efficiently and effectively for saving human lives and alleviating suffering, and for a rapid recovery. Van Wassenhove (2006) emphasize that since the disaster relief is 80% logistics, it would follow that the only way to achieve this is through slick, efficient and effective logistics operations and more precisely, supply chain management. Logistics planning in emergencies involves the quick and efficient distribution of emergency supplies from the emergency response facilities to the affected areas via supply chains. The ERFs considered in this paper are three distinctive ones. They are (i) Central Warehouses (CWHs) or Distribution Warehouses (DWHs), where emergency relief commodities are stored, (ii) intermediate response facilities termed Relief Distribution Center (RDC) or Commodity Distribution Point (CDP), where people can more effectively gain access to relief goods, and (iii) neighborhood sites (NBSs) in need of humanitarian items. The distribution framework of HTSC is depicted in Figure 2 (see Hong and Jeong, 2019).

The HTSC design problem is an inherently strategic and long term in nature. The primary objective of the strategic level is to strengthen emergency preparedness as well as to select the most cost/distance-effective location of CWHs and CDPs among a set of candidate locations, to establish the distribution of emergency supplies throughout the HTSC, and to assign NBSs to CDPs and CDPs to CWHs. Determining such locations/allocations is a critical area in the design of an effective HTSC. However, traditional cost-based facility location models implicitly assume that located facilities will always be in service or be available and do not consider an associated risk of disruption. All facilities are susceptible to disruptions due to natural disasters, accidents, breakdowns, weather, or strikes. Effects of disruptions could be aggravated as a result of a lack of flexibility and interdependency in the HTSC.

CASE STUDY AND OBSERVATIONS

This paper applies the proposed procedure using a case study considered by Hong and Jeong (2019). Using major disaster declaration records in South Carolina (SC), they (2019) consider forty-six counties are clustered based on proximity and populations into twenty counties. Then, they choose one city from each clustered county based on a centroid approach and assume that all population within the clustered county exists in that city. The distance between these cities is considered to be the distance between counties. The database also provides a list of counties where a major disaster was declared. They assumed that when a major disaster is declared, the emergency facility in that county is disrupted and shut down. Based on the historical record and the assumption, the risk probability for each neighborhood (a county or a clustered county) is calculated in Table 1. The potential five locations for CWHs were selected based upon population, the proportion of area that each location would potentially cover, and the proximity to Interstate Highways in SC.

Figure 1: U.S. 2018 Billion-dollar weather and climate disasters

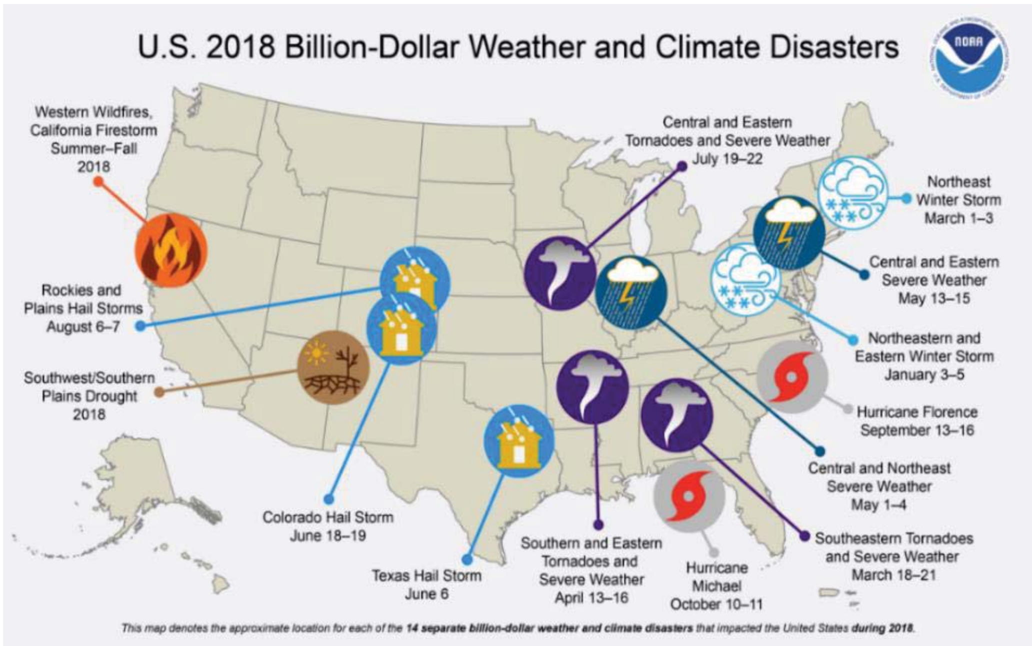


Figure 2: Distribution Framework of Humanitarian Supply Chain

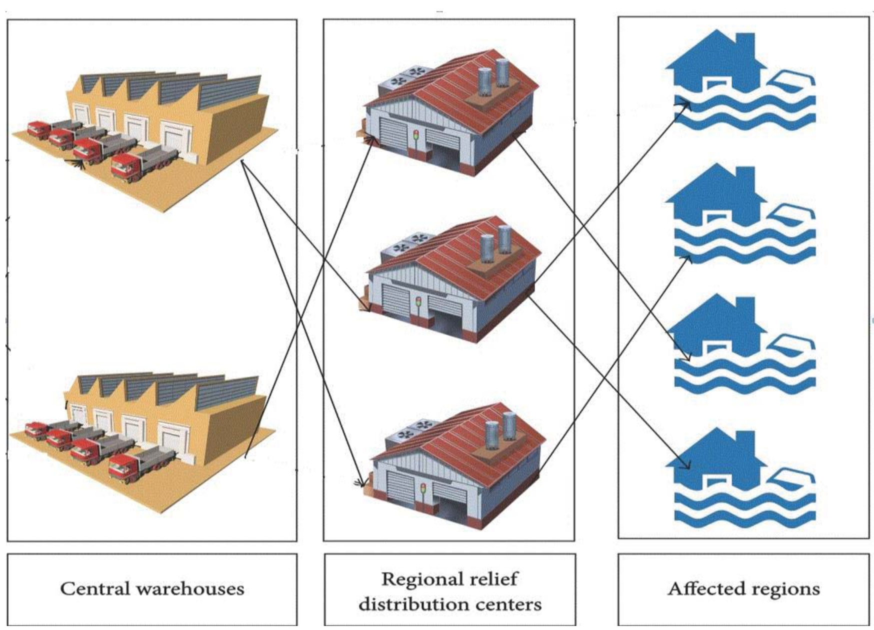


Table 1: Data for Locations of ERFs

No	City	County	Population (K)	Risk Probability
1	Anderson	Anderson/Oconee/Pickens	373	0.125
2	Beaufort	Beaufort/Jasper	187	0.063
3	Bennettsville	Marlboro/Darlington/Chesterfield	96	0.375
4	Conway	Horry	269	0.375
5	Georgetown	Georgetown/Williamsburg	93	0.438
6	Greenwood	Greenwood/Abbeville	92	0.125
7	Hampton	Hampton/Allendale	33	0.188
8	Lexington	Lexington/Newberry/Saluda	318	0.313
9	McCormick	McCormick/Edgefield	35	0.250
10	Moncks Corner	Berkeley	178	0.313
11	Orangeburg	Orangeburg/Bamberg/Calhoun	123	0.375
12	Rock Hill	York/Chester/Lancaster	321	0.313
13	Spartanburg	Spartanburg/Cherokee/Union	367	0.313
14	Sumter	Sumter/Clarendon/Lee	157	0.375
15	Walterboro	Colleton/Dorchester	135	0.250
16	Aiken†	Aiken/Barnwell	184	0.313
17	Charleston†	Charleston	350	0.250
18	Columbia†	Richland/Fairfield/Kershaw	461	0.375
19	Florence†	Florence/Dillon/Marion	203	0.438
20	Greenville†	Greenville/Laurens	521	0.125

†potential locations for CWH

They consider four goals as major performance measures. The first goal is to minimize the total logistics cost (*TLC*). The second goal is to minimize the maximum coverage distance (MCD), which attempts to minimize the longest delivery distance between ERFs. The third goal is to maximize the expected amount of satisfied demands (*ESD*), whereas the fourth goal is to maximize the covered demands in case of emergency (*CDE*). The goal programming model for the case problem is solved for various values of the weight, $\boldsymbol{\alpha}=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, where each weight assigned to each goal changes between 0 and 1 with an increment of 0.1.

There are 286 configurations for each model arising out of the combinations of the setting of $\boldsymbol{\alpha}$ under the condition $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$. After 286 runs, we reduce 286 configurations into 71 consolidated configurations, based upon the values of the four-performance metrics. Each of the 71 configurations is considered as a DMU, which represents the optimal locations of ERFs and their supply chain schemes. First, the DEA model is applied to find efficient DMUs with a perfect efficiency score (ES) of 1.000 and 27 DMUs are found to be efficient. Second, the efficient 27 DMUs are solved using MCDEA by MOP model. As expected, we see several efficient DMUs but there is no way to rank them. Thus, the CE method is applied to find the most efficient DMU, and DMU #81 is identified as No. 1 efficient one. Figure 3 depicts the network scheme of DMU #81. As shown in Figure 3, the most efficient HTSCN, DMU #81, finds {Greenville, Charleston} for the CWH locations and {Anderson, Lexington, Spartanburg} for CDPs covered by the CWH {Greenville}, and finds {Walterboro, Conway} for CDPs covered by the CWH {Charleston}. DMU #81 generates ($327,393, 107, 2854, 2827) as the four-performance metrics.

SUMMARY AND CONCLUSIONS

Several problems have appeared as DEA has been applied to a wide variety of evaluation areas. DEA evaluates DMUs in terms of self-evaluation, which allows each DMU to rate its efficiency score with the most favorable weights to itself. Consequently, the problems related to weak discriminating power have arisen as the applications of DEA advance, since multiple DMUs frequently turn out to be efficient. Lack of discrimination power is the major weakness for DEA. To remedy this weakness and increase the discrimination power, the cross-efficiency (CE) evaluation methods and the multiple criteria DEA (MCDEA) model have emerged. But there is no explicit procedure of ranking the efficient DMUs by either of them.

This paper proposes an innovative procedure of ranking efficient DMUs by integrating these three methods, remedying the inherent weaknesses of three DEA methods. In the first step, DEA is applied to obtain efficient DMUs. In the second step, a MOP model is applied for each efficient DMUs found in the first step to solve the MCDEA model for various values of the weights assigned to the three criteria. In the third step, for each solved MCDEA model, the CE DEA methods are applied to obtain the cross-efficiency score (CESs) for each efficient DMUs and each solved MCDEA model. Fourthly, we rank the efficient DMUs based on the CESs obtained in the third step.

To demonstrate the proposed procedure, this paper uses HTSCN design problem, whose goal is to relieve and minimize the effects of a disaster. To design more balanced HTSCN schemes, a GP model is applied to generate various HTSCN schemes. Then, the proposed procedure to evaluate many HTSCN schemes is implemented. Through the case study, we can rank the efficient HTSCN schemes and identify the most efficient scheme. This paper demonstrates the applicability of the new proposed approach and shows that the proposed approach would be used as an important tool for designing various supply chain network systems efficiently.

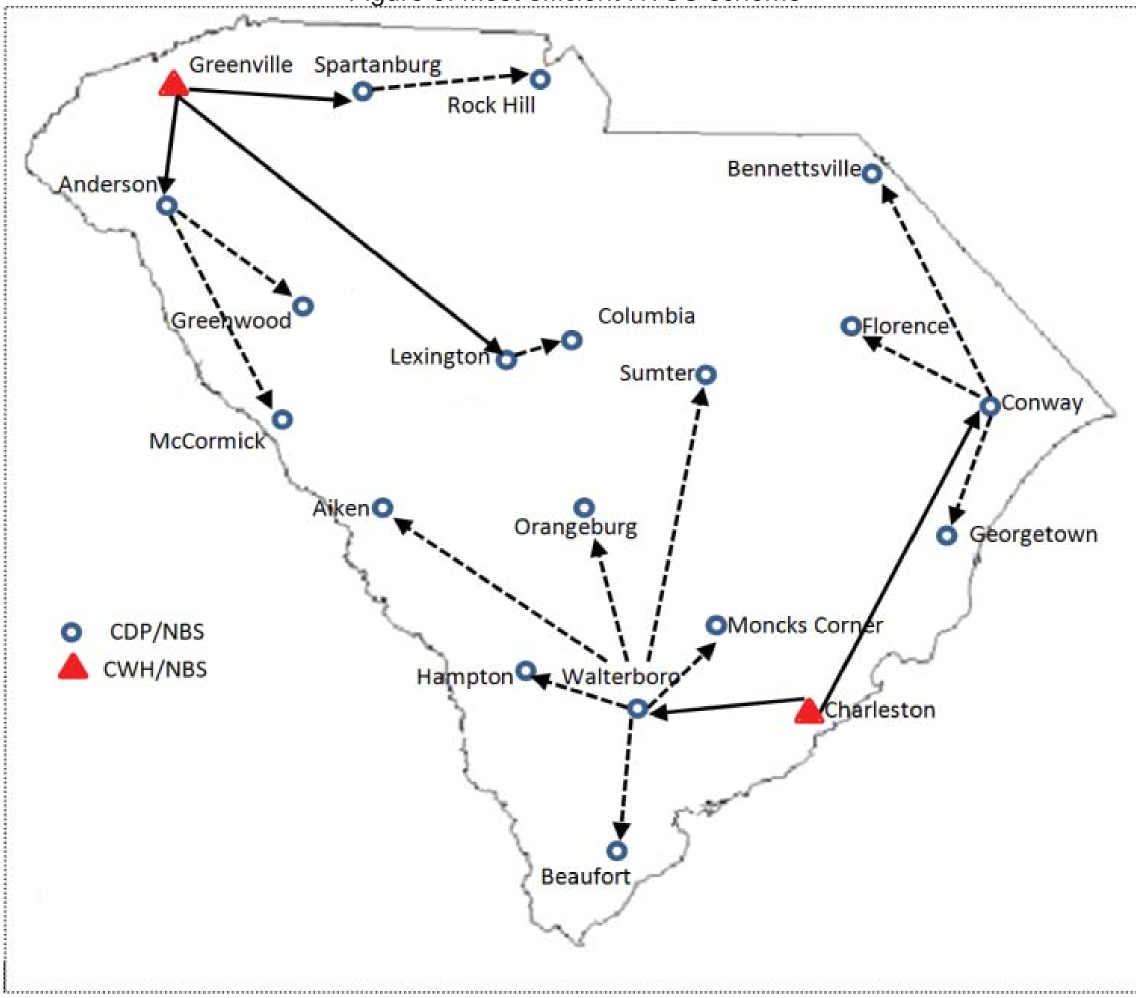
ACKNOWLEDGMENTS

This material is based upon work that is supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, Evans-Allen project number SCX-313-04-18.

REFERENCES

References available upon request.

Figure 3: Most efficient HTSC scheme



DECISION SCIENCES INSTITUTE

Preparedness for Natural Disasters in Houston: A Survey on Hurricane Harvey

Aishwarya Swamy
University of Houston – Clear Lake
adavayyaswama3525@UHCL.edu

Vikas Raghav
University of Houston – Clear Lake
raghavV5125@UHCL.edu

Emre Kirac, PhD
University of Houston – Clear Lake
kirac@uhcl.edu

ABSTRACT

In this study, a survey on Hurricane Harvey is designed to analyze the preparedness of Houston for a Natural Disaster. The project implements a survey research methodology where the data is gathered by asking questions to individuals. The questionnaire consisted of eleven close-ended questions, and the responses were recorded through online forms and face-to-face interviews. The study illustrates the significance of disaster management, previous approaches used in the field of disaster management, and analyzes the survey responses using graphical and tabular representation. The results of the study help understand the challenges and opportunities in disaster preparedness.

KEYWORDS: Hurricane Harvey; Natural Disaster; Survey research; Disaster management; Preparedness

INTRODUCTION

Hurricanes are the most destructive kind of natural disasters and have devastating impacts on the environment. Hurricanes are large, strengthening tropical storms that can pack wind speeds of over 157 mph and unleash more than 2.4 trillion gallons of rain a day (Nunez, 2014). When Hurricanes make a landfall, they carry heavy rain, strong winds and heavy waves that damage the environment and result in loss of lives.

Hurricane Harvey was a category four storm that made landfall in Texas on August 25th, 2017 (National Weather Service, 2017). Harvey made landfall three times in six days, affecting several residents of Houston (Amadeo, 2017). It incurred a loss of $125 billion in damages due to the catastrophic rainfall-triggered flooding (Amadeo, 2017). Houston received more than 50 inches of rainfall, leaving one third of Houston completely flooded (National Weather Service, 2017).

Natural disasters occur with minimum amount of warning and are difficult to prevent, but the risks involved in any disaster can be avoided by being prepared. Lack of preparedness before disaster leads to loss of life and property, and this study enhances the importance of being prepared before any disaster to avoid risks. This paper discusses the importance of efficient use of the resources and how to take measures to serve the affected community, based on the survey result analysis of Hurricane Harvey on the population of Houston. The government, emergency management agencies and the volunteers play a major role in disaster management to ensure adequate help to the affected population. Disaster management involves four phases: mitigation, preparedness, response and recovery, and, in this study, we analyze and discuss these three phases. The goal of this paper is to address the key issues in analyzing the preparedness of Houston to deal with a natural disaster, based on the Hurricane Harvey survey results.

This study applies survey research methodology, which benefits in understanding the preparedness of the government, communities and various organizations in Houston for any natural disaster. The study proposes solutions to reduce risk of damages and loss of lives due to natural disasters and discusses how to be prepared and efficiently to use resources during a disaster. The survey questions also aim to provide analysis on the significance of volunteering and on how various social media platforms can be used to assure help to the affected community, laying emphasis on digital volunteering. We believe that this paper provides a base to various other researches in the field of disaster management and an analysis on Hurricane Harvey.

LITERATURE REVIEW

In recent years, there have been several studies in the field of disaster management and how to efficiently deal with disasters; these studies provide a basis of understanding to our research. Each study differs in specifications. They provide analogies on disaster management and act as resources for our project.

The first study, "*Six in ten Americans are ready to deal with a natural disaster*" by Jeffery M. Jones, is a survey paper that provides guidelines to our research (Jones, 2007). This paper implements the survey research methodology to record the response to Hurricane Katrina. Telephone mode of survey is used in this paper to record responses, which are depicted graphically. In comparison, one key similarity is that our study utilizes survey research methodology, as well, and the key difference in our study is that we implement more than one mode of survey.

Like Harvey, Hurricane Katrina also had a huge impact on the population of New Orleans, and it is significant to understand the effects of another hurricane before analyzing Hurricane Harvey. The study conducted by Narayan Sastry discusses the effects of Hurricane Katrina by collecting data from the residents of New Orleans (Sastry, 2009). The paper lays emphasis on a pilot study, which is a collection of data on the after-effects, and the data is collected through a questionnaire. In comparison, the similarity is that our paper also implements a questionnaire for analysis of Hurricane Harvey.

Hurricane Andrew caused severe damage to the population of Florida, as the authors Stanley K. Smith and Christopher McCarthy discuss in their paper, which implements a telephone survey research methodology by random digit dialing (Smith & McCarty, 1996). The results discuss the damage caused, insurance settlements of the residents, evacuation of the residents, return status and living arrangements of the residents. This study inspires us as an implemented survey research methodology to understand the effects of Hurricane Harvey.

Fuqing et al. discuss the evacuation decisions taken by the residents of Texas Gulf Coast prior to Hurricane Rita (Fuging et al, 2007). They implement a face-to-face survey methodology to analyze the societal impacts of Rita. The survey results discuss the evacuation decision prior Hurricane Rita as well as the influence of Hurricane Katrina for the decision to evacuate. In comparison, our project implements the face-to-face survey methodology to analyze the preparedness of Houston and the effects of Hurricane Harvey.

Disaster preparedness plays a significant role in the disaster management cycle, and one of the objectives of our survey is to understand the preparedness of Houston for a natural disaster. Burke et al. provide insights on the preparedness of the vulnerable populations that are susceptible to the effects of natural disasters (Burke et al. 2012). Survey research methodology is implemented in their paper to assess the risks involved, and the survey results depicted Hurricanes as the top concern for the residents of North Carolina. Transportation and language were found to be of major concern for the vulnerable population, and the paper suggests that the emergency alerts, news, disaster threats and other information regarding the disaster should be provided in the population's native language. In comparison, the paper implements a survey research methodology to analyze the preparedness of the vulnerable population of North Carolina and provides great insights on the language barrier issue when dealing with the preparedness of a natural disaster (Burke et al. 2012). The study conducted by Anthony A. Peguero also enhances the barriers of disaster preparedness and the significance of Hurricane mitigation for the citizens to preserve safety (Peguero, 2006). The survey in this study is conducted through random digit dialing in Florida. The survey results discuss the various sources for information related to hurricanes, and the result and discusses the problems faced by the Latino community in Florida. In comparison, the survey research methodology gives an understanding and analysis on mitigation and preparedness for a Hurricane.

Surveys are designed to illuminate a specific problem through focused data analysis, and, in our project, we implement the survey research methodology to collect information from individuals, focusing on the problem of dealing with natural disasters. The study conducted by Kelley et al. (2003) provides a checklist of good practice in the conduct and reporting of survey research. The paper also discusses the processes of data collection, data analysis and reporting. This study provides a guide to our study and assists in understanding the implementation of research methodology.

It is significant that the effects of natural disaster on surveys are understood and analyzed before conducting the survey. Harter et al. (2006) illustrate an overview of the effects of disasters on surveys and their direct human causes. They discuss the effects of disaster such as injuries, death, refugees, loss of cooperation, inaccessibility and failure of necessary technology, that have an impact on the survey components and these components alter the survey response rates and eligibility rates. The study provides an understanding of the sensitivity of conducting the survey and the measures to be taken for an on-going survey, providing an understanding on survey research methodology.

Whittaker et al. (2015) insight on informal volunteerism of citizens and lays emphasis on growth of digital volunteerism in emergency and disaster management. This study provides a base to our project as our survey questions aim at understanding the role of informal volunteerism and awareness of digital volunteering during a disaster. The growing platform of digital volunteering is significant to reach out to the affected community, and, in our project, one of the survey questions aims at understanding the awareness of digital volunteering and its importance. Park & Johnston (2017) discuss the conceptual framework of digital volunteer networks in response to disasters and crises, as well as the challenges of digital volunteer networks for crisis response. This study helps our project explore the aspects of digital volunteering in preparedness for a natural disaster.

However, the information processed through digital volunteering platforms should be structured accordingly, and the data collected should be distributed to all the volunteers. It is important to consider all the aspects of data collected online, and the study conducted Starbird (2011) enhances the use of crowdsourcing for various information processing tasks during disasters and mass emergencies. This study provides a basis for our project on understanding data collection for digital volunteerism.

Social media platforms such as Facebook and Twitter can be used to help the affected community, as they cater to the response stage of disaster management. Li & Goodchild (2012) discuss the benefits and challenges of using social media in emergency management. The study provides an understanding on issues related to data access, data quality, information synthesis, privacy and equity. In comparison, our survey includes questions related to use of social media during a disaster.

Various-level government collaboration is significant in implementing the life cycle of disaster management, and, in the study, we also discuss the importance of collaboration of government with citizens as first responders. Lord (2003) discusses the framework of integrated emergency management, which allows various levels of government to work together to mitigate, respond and recover from disasters. In comparison, our questionnaire analyzes the role of government in response to a disaster. One such organization, Federal Emergency Management Agency (FEMA), took an initiative to create awareness for preparedness for earthquakes in the central United States. The paper, '2011 FEMA Central States Disaster and Earthquake Preparedness Survey Report' by FEMA, provides an analysis on earthquake preparedness in central United States (FEMA, 2011). The survey conducted in this paper evaluates the effectiveness of the Earthquake outreach activities and provides recommendations for increasing preparedness. The survey targeted the households in eight different states through telephone mode of interview, for which random digit dialing was implemented. The results of the survey indicate that the Earthquake Outreach program was effective and that outreach initiatives prepare communities to avoid risks. In comparison, the survey report in this study provides insights to our project for analysis of the results and our project uses two modes of survey, whereas, in this survey, report telephone mode of survey is implemented.

SURVEY RESEARCH METHODOLOGY

This methodology is a systematic process of gathering information by asking questions to individuals. The methodology was implemented to our project considering the following process:

Designing Survey

Here, goals are defined, target population is considered, and the timeline of the survey is defined. In this phase of methodology, we defined the objective of our project, and the objective of the survey was designed to understand the preparedness of communities and government during a natural disaster, the usage of social platforms during a disaster, the awareness of digital volunteering, and lastly, the collaboration of government and citizens as first responders during a disaster. The residents of Houston were considered as target population for our survey in order to understand the effects of Harvey and their preparedness for a natural disaster. Considering the timeline, the survey was completed within the timeline of 1-2 months.

Developing Questions

In this step of the process, closed-ended questions are developed, which includes response alternatives to all of the questions. The survey consisted of eleven closed-ended questions and one open-ended question to record the comments or suggestions of the respondents. The appendix of this paper consists of the survey form. Most of the questions in this survey were developed as close-ended question to avoid time consumption and to make the questionnaire effortless for the respondents.

Testing and Training

In this phase, the questions are reviewed, tested and revised multiple times. To ensure the validity of our survey questions, the sample survey form was distributed to a few students to obtain feedback. The feedback was obtained from these sample responses to make the required changes to the survey, and the final revised edition of the survey form was distributed to the communities in Houston.

Collecting Data

In this step, the survey is conducted using various modes such as telephone, email, online and interviews, and the response rate is recorded. In the project, two modes of data collection was implemented: online survey responses and face-to-face interview responses. Online survey responses were recorded using Google Forms, and face-to-face responses were collected by targeting the affected areas of Houston. The total number of responses collected is 120; 73.34% responses were recorded through google forms and 26.67% of the responses were recorded through face-to-face responses.

RESULTS

The survey responses were obtained through Online forms and Face-to-Face interviews, and a total of 120 responses were obtained. 73.34% of responses were recorded through Google Forms, and 26.67% of the responses were recorded through face-to-face responses.

Table 1: Number of responses obtained

Total responses	**120**
Online	88
F2F (face-to-face)	32

The following tables depict the results obtained for each question.

1. **Did you have the following minimum disaster supplies on-hand before Hurricane Harvey?**

Table 2: Response concerning minimum disaster supplies

	Online	**F2F**	**Total**	**%**
Sufficient supply of Food and Water	77	17	94	78.3%
Supply of personal care items and hygiene products	60	14	74	61.7%
A first aid kit	44	14	58	48.3%
Supply of Medicine	48	13	61	50.8%

Flashlight	51	18	69	57.5%
Local maps	16	13	29	24.2%
Cell phone with backup battery	46	11	57	47.5%
None of the above	7	4	11	9.2%

2. Prior to Hurricane Harvey, did your emergency evacuation plan include the following?

Table 3: Response concerning evacuation plan

	Online	**F2F**	**Total**	**%**
Evacuation Route	21	9	30	25.0%
Fuel in the vehicle tank	41	20	61	50.8%
First aid kit	30	12	42	35.0%
Emergency contact list	32	13	45	37.5%
I did not have any emergency evacuation plans	47	9	56	46.7%

3. **If needed, were you able to reach 911 at the earliest?**

Table 4: Response concerning 911

	Online	**F2F**	**Total**	**%**
Yes	20	14	34	28.33%
No	6	12	18	15.00%
I did not need to call 911	62	6	68	56.67%
			120	100.00%

4. **If needed, did you post a request for help on social media?**

Table 5: Response concerning social media

	Online	F2F	Total	%
Yes	13	12	25	20.83%
No	13	15	28	23.33%
I did not need help	62	5	67	55.83%
			120	100.00%

5. **Do you think Police departments, fire departments and rescue squads assured help and support as first-responders to the people of Houston?**

Table 6: Response concerning first responders

	Online	F2F	Total	%
Yes	79	22	101	84.17%
No	9	10	19	15.83%
				100.00%

6. **Do you think the Emergency Management Agencies should use social media as a platform to assure help for the people of Houston during a disaster?**

Table 7: Response concerning Emergency Management Agencies

	Online	F2F	Total	%
Yes	81	25	106	88.33%
No	7	7	14	11.67%
			120	100.00%

7. **Did you offer any kind of help during Hurricane Harvey as a volunteer?**

Table 8:Response on total initiatives as volunteer

	Online	F2F	Total	%
Yes	35	14	49	40.83%
No	53	18	71	59.17%
			120	100.00%

8. **Do you think enough citizens took initiative as volunteers for search and rescue operations during Hurricane Harvey?**

Table 9: Responses for citizens as volunteers

	Online	F2F	Total	%
Yes	73	18	91	75.83%
No	15	14	29	24.17%
			120	100.00%

9. **Would you trust help from anyone apart from friends or family during the disaster?**

Table 10: Response concerning trusting of help during a disaster

	Online	F2F	Total	%
Yes	78	21	99	82.50%
No	10	11	21	17.50%
			120	100.00%

10. **Are you aware of digital volunteering?**

Table 11: Response concerning awareness of digital volunteering

	Online	F2F	Total	%
Yes	64	17	81	67.50%
No	24	15	39	32.50%
			120	100.00%

11. **Do you expect government organizations to collaborate with citizens as first-responders during the disaster response?**

Table 12: Response concerning government organization collaboration

	Online	F2F	Total	%
Yes	79	20	99	82.50%
No	9	12	21	17.50%
			120	100.00%

ANALYSIS AND DISCUSSION

The graphical representation of the result is provided for a better understanding and analysis of the data.

Mode of Survey: The data was collected using two modes of survey: online surveys and face-to-face interviews. The number of responses obtained for the online survey was greater than face-to-face interview responses as depicted in Figure 1, with a total 120 responses, 73.34% responses were obtained from online surveys, and 26.67% of the responses were obtained by conducting face-to- face interviews. Face-to-face interviews were very challenging, as it was comparatively difficult to obtain responses. The following are the disadvantages:

- The respondents were self-conscious about the responses
- Refusals due to time constraints

- Refusals due to language barriers
- Sensitivity of the topic and lack of interest in the survey, leading to refusals.

The responses obtained from online surveys were higher in number and had several advantages such as:

- Easier and faster due to the availability of the software that support online survey forms; our project implemented Google Forms to conduct online surveys
- Responses were obtained from various distant locations in Houston
- Automated data collection through Google Forms which reduced waste of time and efforts
- Facilitated the respondents to participate at their convenience (Wright, 2005).

In conclusion, to overcome the challenges of conducting face-to-face interviews, one should provide a comprehensive introduction and adequate information regarding the survey to gain the trust of the respondents. Rewards, expressing appreciation and sense of confidentiality should be provided to the respondents to increase the response rate through face-to-face interviews. [28]

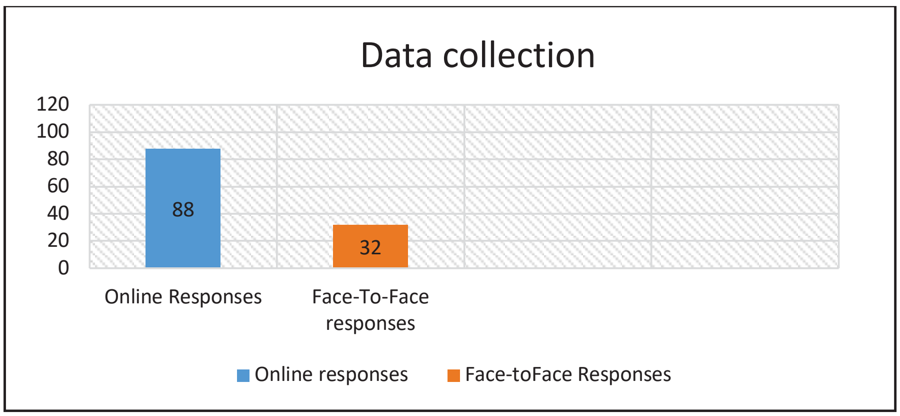


Figure 1: The collection of survey responses

Disaster Preparedness: The first question of the survey provides insight into the number of people that possessed the disaster supplies prior Hurricane Harvey. The objective of the project is to understand the preparedness of the people of Houston. Figure 2 depicts the number of people who had the minimum supplies before Hurricane Harvey.

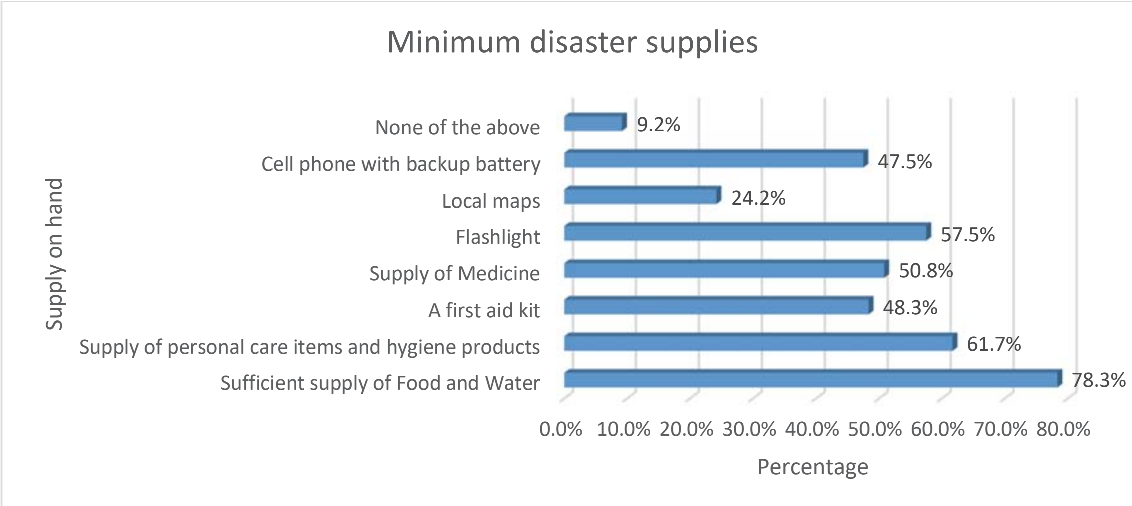


Figure 2:Minimum disaster supplies

Analyzing Figure 2, 9.2% of the respondents did not have any of the minimum supplies, and most of the respondents did not possess all the supplies listed in the question. Therefore, it is essential to create awareness regarding the minimum disaster supplies prior to any disaster among the communities, as disasters cause a lot of damage and can have long-last effects. Likewise, the second question also concerns disaster preparedness. Figure 3 depicts the number of respondents with an evacuation plan prior to Hurricane Harvey. Having an evacuation plan plays a major role in the preparedness phase of disaster management to avoid risks and damages, and 46.7% of the respondents did not have any emergency evacuation plan

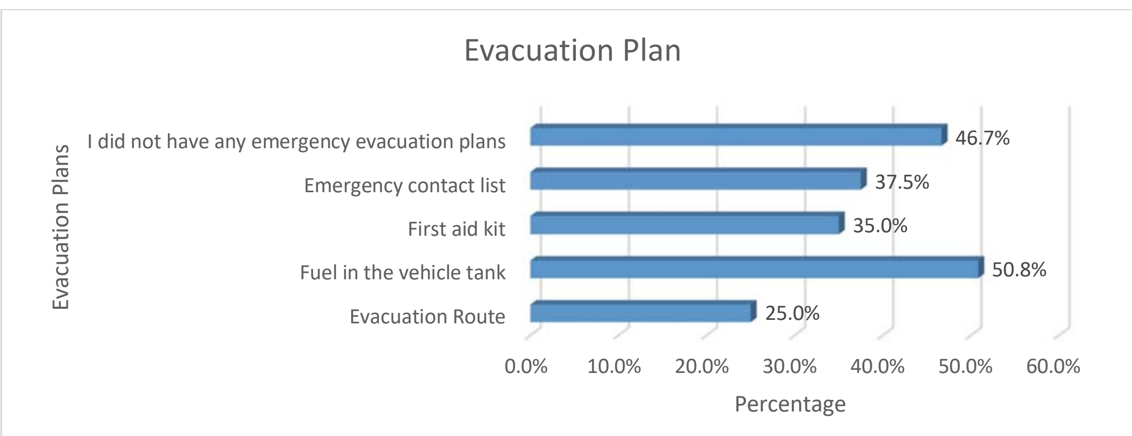


Figure 3: Responses concerning Evacuation Plan

Based on the survey result, we analyze that most of the respondents did not possess disaster supplies or an evacuation plan. However, it is important to create awareness about preparedness, and the following are proposed solutions for creating awareness: (Smith, 2008)

- The Emergency Management Agencies should conduct informational meetings and workshops at various localities of Houston to create awareness regarding the significance of planning and preparing prior any disaster.
- Local media should take initiative to inform and create awareness among the citizens regarding minimum disaster supplies and how to create an evacuation plan.
- Initiative should be taken to promote emergency preparedness at various levels of schools and colleges.
- Announcements should be made over television channels, radio stations and on various disaster management websites.
- With evolving technology, there is a drastic increase in the number of people that rely on social media platforms like Facebook, Twitter etc. for information and the use of social media platform to create awareness will prove successful. Therefore, social media should be better utilized by federal and state government agencies.
- Awareness booths should be set up at fairs and events; brochures containing the list of disaster supplies and evacuation plan should be distributed to the residents.

Emergency Number: The emergency number, 911, received around 80,000 calls in a 24-hour period during Hurricane Harvey (Kelly, 2017). Most people were unable to get through the lines as they were put on hold, and, based on our survey result in Figure 5, we observe that 15% of the respondents were unable to reach 911. It is important that every individual of the affected community is assured help, which is why it is necessary to create awareness about digital volunteering platforms and the use of social media platforms such as Facebook and Twitter to respond and ensure help during a disaster. The emergency number system should be advanced and updated to attend more number of calls and should provide a platform to process text messages from people to assure help to a greater number of people.

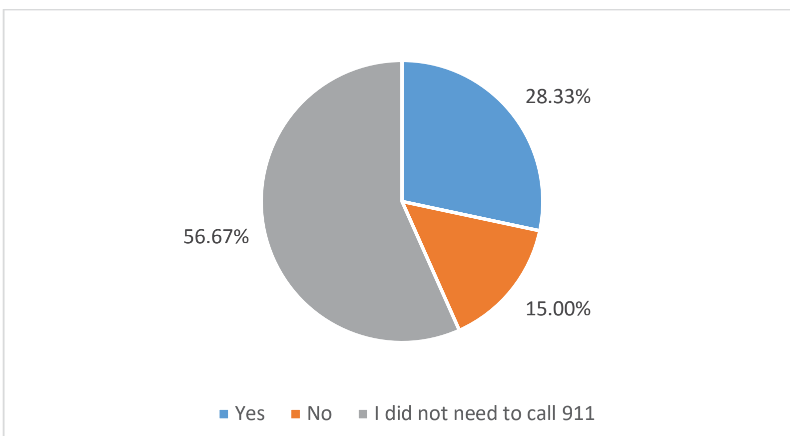


Figure 4: Response concerning 911

Social media in disaster management: With evolving technological innovations, social media plays a significant role in updating and keeping people informed. Use of social media platform in the field of disaster management can benefit a huge population and enhance the response phase. Based on the survey result, Figure 6 represents that 55.83% of the respondents used social media to post requests for help during Hurricane Harvey. Social media platforms such as Facebook and Twitter served as an alternative to 911 during Hurricane Harvey. The social media platform should be used by government organizations, Emergency Management Agencies and various other non-profit organizations to collect data and provide help to the affected community.

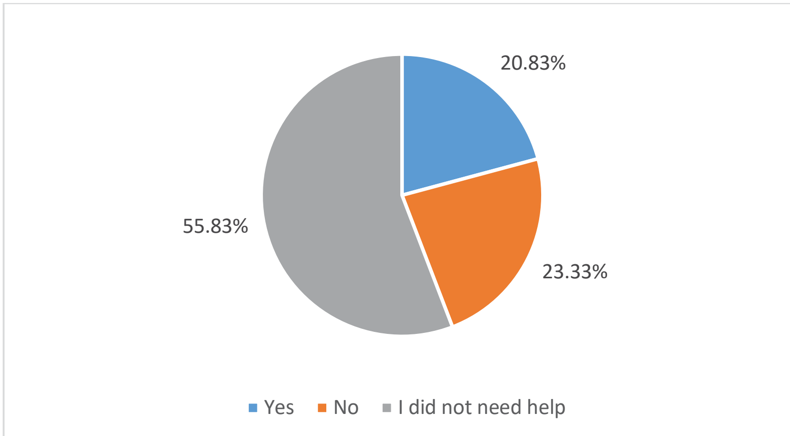


Figure 5: Response concerning social media

First-Responders: First-responders play a major role in disaster management, as they address the response phase. Considering the survey results, Figure 6 depicts that 84.17% of the responses show that the fire departments, police departments and rescue squad assured help and support to the people of Houston during Hurricane Harvey. The first-responders should be offered adequate training and guidelines to rescue and offer help during any disaster.

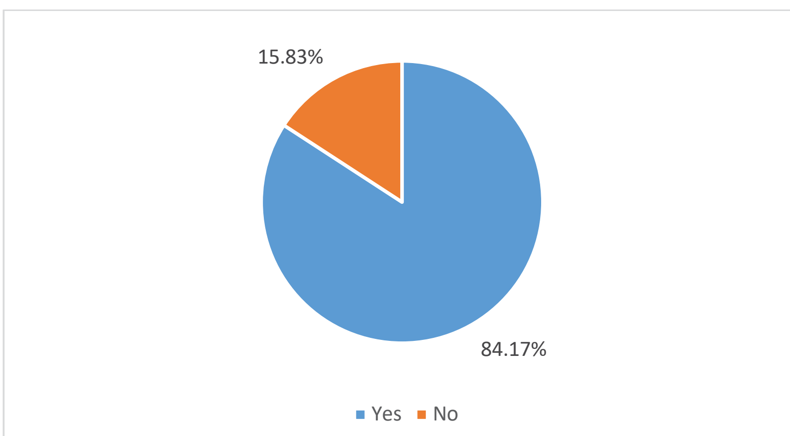


Figure 6: Response concerning first-responders

Emergency Management agencies: Federal Emergency Management Agency (FEMA), American Red Cross, the Federal Government of the United States and various other agencies played a significant role in reaching out to the affected communities during Hurricane Harvey. The Coast Guard deployed 2,060 personnel, 50 aircrafts and 75 boats to rescue 11,022 people and 1,384 pets. FEMA assigned 28 Urban Research and Rescue teams to assist local agencies, and there were various organizations that aided and supported the affected community during Hurricane Harvey (FEMA, 2017). Based on the survey results, Figure 8 represents that 88.33%

of responses suggest that the Emergency Management Agencies should also use social media platforms to provide more assistance and support to the communities of Houston. Social media platforms like Facebook and Twitter are being used by almost everyone today, and using these platforms to reach out to people will benefit the agencies.

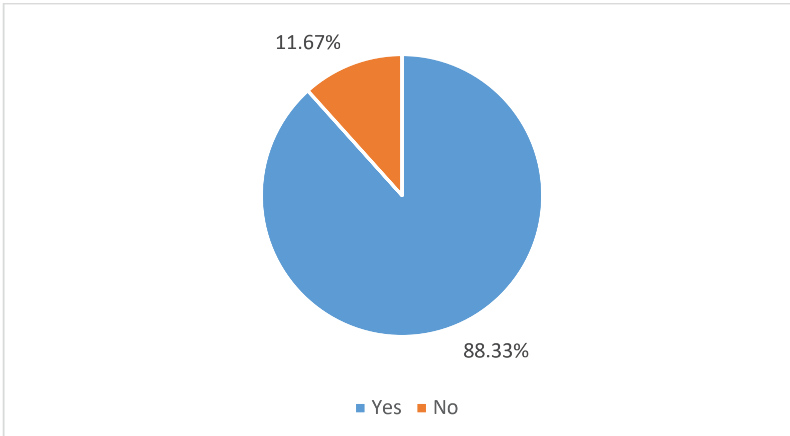


Figure 7: Response concerning Emergency Management Agencies

Informal Volunteerism: Citizen participation plays a vital role in disaster management, and the disaster relief and response phases rely largely on the citizens. In the survey, when asked if citizens participated as volunteers, the response depicted in Figure 8 illustrates that 59.17% of the respondents did not participate. It is essential to create awareness among the citizens regarding the importance of volunteerism and the evolving concept of digital volunteering.

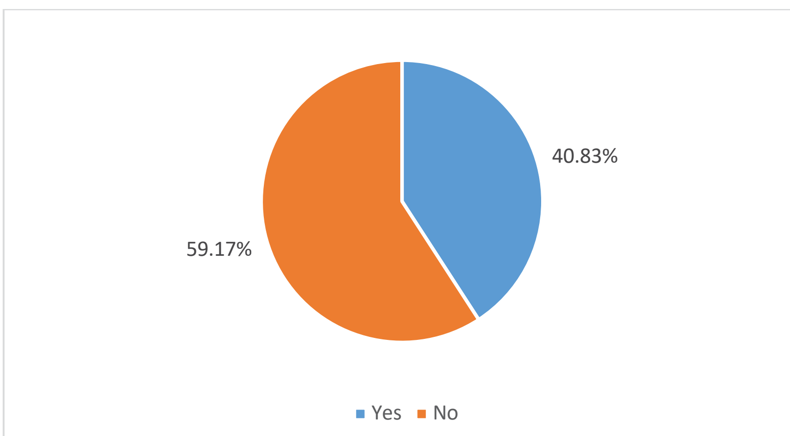


Figure 8: Responses on total no of initiatives as volunteer

The respondents were also asked if they thought enough citizens took initiative as volunteers during Hurricane Harvey, and Figure 9 illustrates that 75.83% of the respondents chose ‘YES’. The result illustrates that 24.17% thought that not many initiatives were taken for volunteerism, which is why it is important to create awareness regarding volunteerism, It is also significant that volunteers engage in their roles

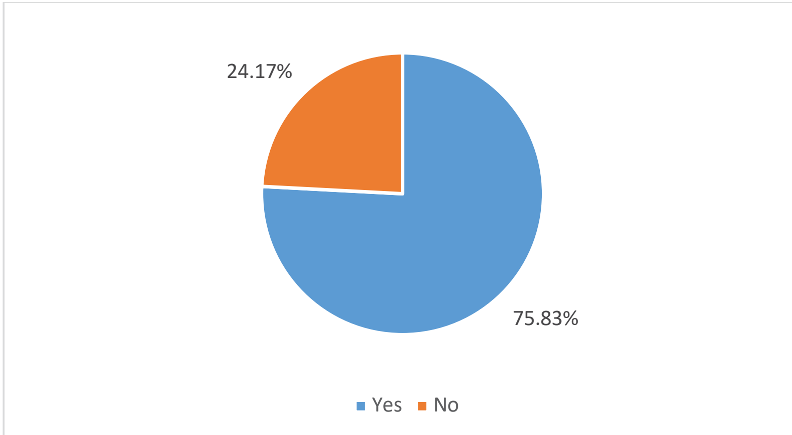


Figure 9: Responses for citizens as volunteers

Referring to Figure 10, 17.50% of the responses illustrated that help is not trusted from people other than family and friends, and 82.50% of the respondents' trust help from others. It is essential that the volunteers are trained to make sure that they gain trust of the citizens. The following are to be considered to create awareness and engage volunteers in their roles (CNCS, 2019):

- Create awareness among the citizens using various social media platforms, through T.V. and radio announcements, workshops and informational meetings at various locations in Houston.
- Define the roles and responsibilities of every volunteer and provide adequate training.
- Assign tasks to volunteers and provide them with necessary tools

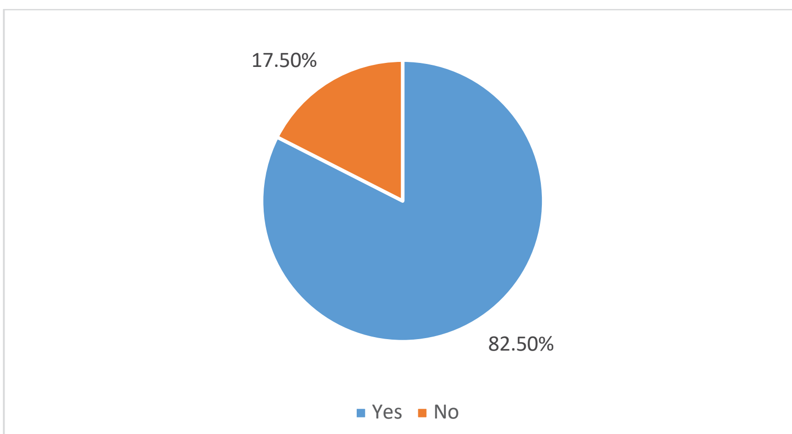


Figure 10: Responses to trusting help from volunteers

Digital Volunteering: With improving technological innovations, volunteering has been revolutionized, as it is more accessible and flexible. The digital volunteer networks allow citizens to contribute to disaster management, as digital volunteering utilizes social media and web-based mapping software to record and share emergency-related information that allows volunteers to respond to the affected communities. Figure 12 illustrates the awareness of digital volunteering among the citizens; 67.50% of the respondents are unaware of digital volunteering. It is important that citizens are aware of digital volunteering. With the help of digital platforms, every individual can contribute even from distant localities and can contribute at their convenience. ‘Crowdsourcerecue.com’ played a vital role in responding to the people of Houston during Hurricane Harvey. The platform connected over 12,000 professionals and volunteer rescuers with 35,000 people using its mapping and dispatching technology (https://crowdsourcerescue.com/). The platforms such as CodeCorps, All for Good, Volunteer-Match, American Red Cross and various other websites provide volunteer opportunities.

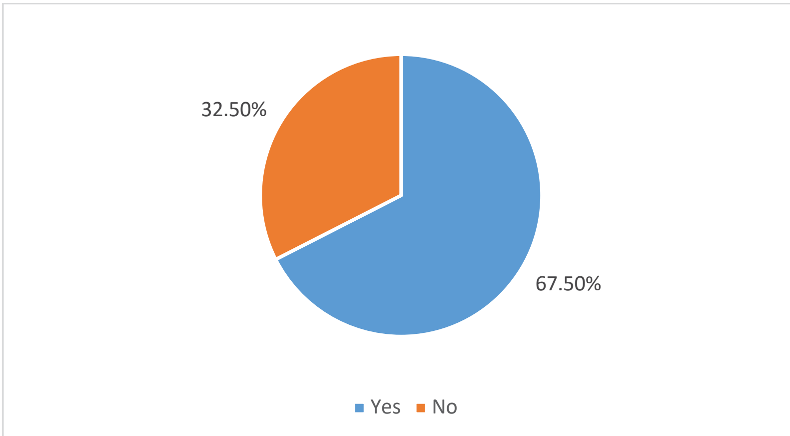


Figure 11: Response concerning awareness of digital volunteering

Collaboration of Government Organization: Government organizations should collaborate with citizens as first responders to create awareness and provide better support during any emergency. Figure 12 illustrates that 82.5% of the respondents expect government organizations to collaborate with the citizens as first-responders. Government organizations and volunteers play a significant role in disaster relief and response.

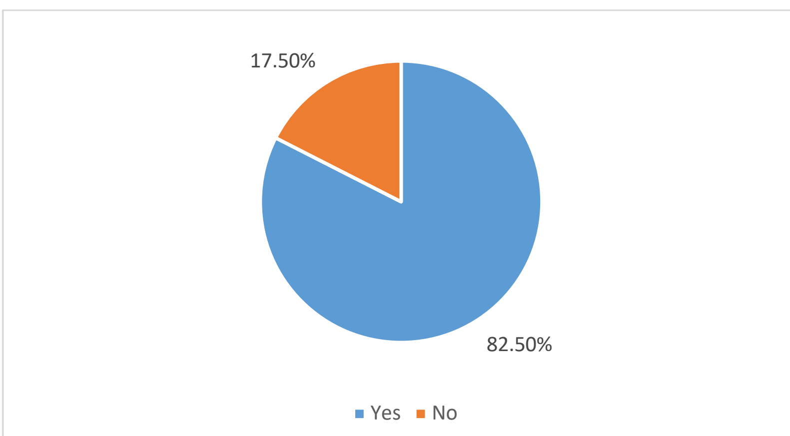


Figure 12: Collaboration

CONCLUSIONS

The objective of the study ‘Preparedness for Natural Disasters in Houston: A Survey on Hurricane Harvey’ is to understand the risks and damages associated with Hurricane Harvey and to propose effective solutions by analyzing the survey responses. The survey implemented online and face-to-face survey methodologies to understand the impact of Hurricane Harvey. The survey responses provided an insight into the preparedness of citizens, government organizations and first-responders for a natural disaster. It also provided an understanding of the awareness of digital volunteering and the role of social media in disaster management among the residents of Houston. Overall, the study provides solutions to reduce risks and damages caused by any natural disaster, information on how to effectively use resources, and suggestions on how to create awareness for disaster preparedness and digital volunteering.

APPENDIX

A Survey on Hurricane Harvey

<u>Instructions</u>:
☐ - Choose multiple options.
O - Choose only one options.

1. Did you have the following minimum disaster supplies on-hand before Hurricane Harvey?
   - ☐ Sufficient supply of Food and Water
   - ☐ Supply of personal care items and hygiene products
   - ☐ A first aid kit
   - ☐ Supply of Medicine
   - ☐ Flashlight
   - ☐ Local maps
   - ☐ Cell phone with backup battery
   - ☐ None of the above

2. Prior to Hurricane Harvey, Did your emergency evacuation plan include the following?
   - ☐ Evacuation Route
   - ☐ Fuel in the vehicle tank
   - ☐ First aid kit
   - ☐ Emergency contact list
   - ☐ I did not have any emergency evacuation plans

3. If needed, were you able to reach 911 at the earliest?
   - o Yes
   - o No
   - o I did not need to call 911

4. If needed, did you post a request for help on social media?
   - o Yes
   - o No
   - o I did not need help

5. Do you think Police departments, fire departments and rescue squad assured help and support as first-responder to the people of Houston?
   - o Yes
   - o No

6. Do you think the Emergency Management Agencies should use social media as a platform to assure help for the people of Houston during a

disaster?
- o Yes
- o No

7. Did you offer any kind of help during Hurricane Harvey as a volunteer?
- o Yes
- o No

8. Do you think enough citizens took initiative as volunteers for search and rescue operations during Hurricane Harvey?
- o Yes
- o No

9. Would you trust help from anyone apart from friends or family during the disaster?
- o Yes
- o No

10. Are you aware of digital volunteering?
- o Yes
- o No

11. Do you expect the government organizations to collaborate with citizens as first-responders during the disaster response?
- o Yes
- o No

REFERENCES

Amadeo, K. (2017). *Hurricane Harvey Shows How Climate Change Can Impact the Economy*. Retrieved April 11, 2018, from https://www.thebalance.com/hurricane-harvey-facts-damage-costs-4150087

Burke, S., Bethel, J. W., & Britt, A. F. (2012). *Assessing disaster preparedness among Latino migrant and seasonal farmworkers in eastern North Carolina*. International journal of environmental research and public health, 9(9), 3115-3133.

Corporation for National and Community Service (CNCS) (2019), Defining Volunteer Roles in Preparation for Disaster Response, https://www.nationalservice.gov/resources/disaster-services/defining-volunteer-roles-preparation-disaster-response

Crowd source rescue, retrieved on April 11, 2018 from, http://crowdsourcerescue.com/

FEMA, *2011 FEMA Central States Disaster and Earthquake Preparedness Survey Report*, https://www.fema.gov/media-library-data/20130726-1851-25045-5549/2011_fema_central_states_disaster_and_earthquake_preparedness_survey_report_508_compliant.pdf

FEMA, (2017), Historic Disaster Response to Hurricane Harvey in Texas, https://www.fema.gov/news-release/2017/09/22/historic-disaster-response-hurricane-harvey-texas

Harter, R., Petty J., and Kelly, J. (2006). Conducting Surveys When Disasters Strike. ASA Section on Survey Research Methods, 3113-3117.

Jones, J. (2007). *Six in 10 Americans Ready to Deal With a Natural Disaster*, Gallup News Service, https://news.gallup.com/poll/27583/six-americans-ready-deal-natural-disaster.aspx

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). *Good practice in the conduct and reporting of survey research*. International Journal for Quality in health care, 15(3), 261-266

Kelly, H. (2017). Harvey highlights issues of aging 911 tech, CNN Business, https://money.cnn.com/2017/09/01/technology/future/911-tech-harvey/index.html

Li, L., & Goodchild, M. F. (2012). *The role of social networks in emergency management: A research agenda*. In Managing Crises and Disasters with Emerging Technologies: Advancements (pp. 245-254). IGI Global.

Lord, A. F. (2003). *Integrated Emergency Management: The Roles of Federal, State, and Local Government with Implications for Homeland Security*. ARMY WAR COLL CARLISLE BARRACKS PA.

Nunez, C. (2014). *Hurricanes, cyclones, and typhoons, explained*. National Geographic. National Geographic Society, http://environment.nationalgeographic.com/environment/natural-disasters/hurricane-profile/

National Weather Service (2017). *Hurricane Harvey & Its Impacts on Southeast Texas (August 25-29, 2017)*. https://www.weather.gov/hgx/hurricaneharvey

Park, C. H., & Johnston, E. W. (2017). *A framework for analyzing digital volunteer contributions in emergent crisis response efforts*. New Media & Society, 19(8), 1308-1327.

Peguero, A. A. (2006). *Latino disaster vulnerability: The dissemination of hurricane mitigation information among Florida's homeowners*. Hispanic Journal of Behavioral Sciences, 28(1), 5-22.

Sasty, N. (2009), *Tracing the effects of Hurricane Katrina on the population of New Orleans: The displaced New Orleans residents pilot study*, Sociological Methods& Research, 23(1), 171-196

Smith B. (2008), 10 Ways to Increase Awareness of Preparedness. Emergency Management. https://www.ehstoday.com/fire_emergencyresponse/ehs_imp_82130

Smith, S. and McCarty, C. (1996). *Demographic Effects of Natural Disasters: A Case Study of Hurricane Andrew*. Demography. 33. 265-75. 10.2307/2061876.

Starbird, K. (2011). *Digital volunteerism during disaster: Crowdsourcing information processing*. In Conference on human factors in computing systems (pp. 7-12).

Whittaker, J., McLennan, B., & Handmer, J. (2015). *A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges*. International journal of disaster risk reduction, 13, 358-368.

Wright K. (2005). *Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services*, Journal of Computer-Mediated Communication, Volume 10, Issue 3, 1 April 2005, JCMC1034

Zhang, F., Morss, R. E., Sippel, J. A., Beckman, T. K., Clements, N. C., Hampshire, N. L., ... & Wang, S. (2007). *An in-person survey investigating public perceptions of and responses to Hurricane Rita forecasts along the Texas coast*. Weather and Forecasting, 22(6), 1177-1190.

DECISION SCIENCES INSTITUTE

Understanding Crowdfunding Behaviour of Citizens Post Disasters using Civic Voluntary Model

Abhishek Behl
Shailesh J Mehta School of Management
Indian Institute of Technology, Bombay, India
abhishekbehl27@gmail.com

Pankaj Dutta
Shailesh J Mehta School of Management
Indian Institute of Technology, Bombay, India
pdutta@iitb.ac.in

ABSTRACT

The study aims to understand the crowdfunding behavior of donors for supporting development of disaster affected areas from both technological and governance viewpoint in an Indian context. The results offer counter intuitive arguments to explain why and how facilitating conditions and recruitment network are insignificant in explaining crowdfunding behavior.

KEYWORDS: Crowdfunding; Civic Voluntary Model; Financial Donors; and Natural Disaster

INTRODUCTION

There has been a transcendental shift in the area of Humanitarian Supply Chain (HSC) and Disaster Relief Operations (DRO). The field has gradually gained attention by transdisciplinary researchers in the recent past. Statistical data related to disasters reveal two important trends: firstly, the rate of casualty deaths caused because of natural disaster has come down significantly and secondly, the rate of economic losses resulting due to catastrophes have gone up significantly(https://www.emdat.be/). Data from EM-DAT reports that between 1998 and 2017, countries with maximum likelihood probability of facing disaster reported $2.9 trillion in direct economic losses of which 77% of it was a resultant of climatic change. Out of the total losses recorded, one third of it has been reported by United States followed by China, Japan, India and Puerto Rico(UNISDR, 2018). Looking at the economies of United States, China and Japan, it can be figured out that they qualify these countries to lie in the developed nations while countries like India still struggle in position itself in the developing nations category. A two decade report published by UNISDR on disaster relief operations report India to be the only lower middle income country featuring in the top 10 countries affected highly by disaster(UNISDR, 2018). This makes India an interesting case and a geographic area of research by humanitarian researchers are moreover social scientists working in the area of financial and economic ecosystems linked with disaster.

India has a history of disasters with earthquake, floods and droughts being the most common ones and with the highest degree of disaster. Around 58% of the Indian landmass is prone to moderate to high intensity of earthquake of which 110 districts (just over one seventh of the total districts in India) fall in Zone 5 resembling the highest risk zone in the glossary of earthquake. Looking at the water based disasters, the figures are much more gruesome. The erosion data of Indian coastline shows that in the past 26 years (1990-2016) , one third of the coast has eroded.

Disaster although has not been a direct reason for it but climate change is, which resulted in 1.35 million people displaced from these affected areas in 2017, which indeed is an alarming rate. The displacement of people increases the population density of other places thus making them vulnerable to disaster in the long run(NDMA, 2019).

Complexity begins with multiple causes of disaster. The recent Kerala floods disaster is a fine example of the same. Extreme rainfall, rush of dam water and landslides each brought unique challenges to affected areas. Just weeks after the floods, parts of the state were facing a dry spell with groundwater levels falling. Thus, damage that look similar need to be addressed differently which makes every disaster an unique case in itself. Disasters are said to cost the global economy USD 520 billion dollars every year (Crunch, 2018). Considering that no economic loss data is available for nearly 87% of disasters in low-income countries (UNISDR), the actual figures may be much higher. The economic losses that do hit the headlines are often in terms of insured claims or large-scale infrastructural damage. These are clearly rising. However, what about the cost to those under the radar? Informal and uninsured losses don’t get captured or compensated. The massive floods across India in 2017, for example, were estimated by Munich Re to have cost the country USD 2.5 billion, yet insured losses were negligible(Kron, 2018). While economic and financial losses are inevitable, its recovery remains a constant challenge for the households and the society at large. India alone suffered a staggering economic loss of $79.5 billion due to climate related disasters in the past two decades (UNISDR, 2018). The recent Kerala floods in 2018 itself caused $4.25 billion economic loss which excludes the direct damage and business interruption costs(Aon catastrophic report, 2018). A careful assessment of monetary supply chain has been missing from the literature which would help in achieving financial stability to the affected people post natural disasters.

Systematic reviews and development of theoretical, conceptual and practical models related to humanitarian supply chain have been umpteen in number in academic literature(Banomyong et al., 2016). Studies have also used theoretical lens and cases study based analysis to understand various aspects of humanitarian supply chain (Dubey and Gunasekaran, 2016). However, there is a dearth of literature which explores the enablers of financial supply chain for relief victims in order to achieve financial normalcy. The stakeholders contributing towards financial normalcy includes government agencies, non-government organizations, banks, insurance firms, CSR activity done by companies and crowdfunding by agencies and individuals. There has been umpteen research done on understanding the role of government, NGO, banks and insurance firms regarding their contribution towards financial relief for the victims, while individual contribution made by the countrymen still needs attention(Burkart et al., 2016; Kunz and Gold, 2016). This also calls for addressing the topic of crowdfunding in the context of DRO and moreover the willingness of people to participate in it. The individual contribution made by countrymen are usually voluntary donations made through various channels like direct transfers to affected people (which is made when the affected people are related to the affected masses), indirect transfers made to the government routed through either from their firms which they work, payments made to the government on the disaster relief portals, donation made to relief agencies like red cross, etc. While some portion of these payments are made due to voluntary contributions, a large part of it is also involuntary as they are triggered by external agencies like NGO’s, state or central government, companies etc. (Bukart et al., 2016; Kabra and Ramesh, 2015). The existence of multiple players in the financial ecosystem contributing towards economic development sprouts issues related to coordination, abrupt pattern of flow of funds, disintegrated approach towards handling the demand and supply of monetary relief, unstructured disbursement of funds, unregulated and unplanned expenditure etc.

Studies confirm that there are multiple models and framework which are developed and adopted for disaster relief operations and humanitarian logistics (Habib et al., 2016), but the financial supply chain remains uncertain in general. While it is uncertain, the objectives of it is well defined when compared to disaster relief operations and humanitarian logistics which work on multi layered and multi actor approach. The problem becomes more relevant and important in an Indian context as it belongs to lower middle income bracket which doubles the impact on the affected audience. As government plays as an aggregator and facilitator for financial donations made through multiple agencies, its role in the financial supply chain becomes important. Both NGO and government has been instrumental in making their significant contribution in disaster planning and disaster relief, however the role of citizens has off late being less discussed in the contributing towards disaster relief operations. Literature confirms that citizens contribute both in cash and kind from remote locations and offer support to the disaster affected citizens by offering physical help as well. Furthermore, studies have also confirmed that it calls for multiple theories to explain this complex nature of human beings (Zolotov et al., 2018; Dwivedi et al., 2017).

While behaviour of an individual is a result of its reaction of stimuli and vice versa, it is important to reconsider looking at behavioural studies from multiple lens. The present study aims to understand the behaviour of citizens towards helping victims of a natural disaster by making financial contribution. Therefore, behaviour in the present study would not be an outcome of only pre constructed models in IS literature. Rather, it would an outcome of constructs borrowed from governance literature as well. While payments can be made directly or indirectly to the government to ensure economic development, the present study only limits to direct transfers made to the government by citizens and more importantly voluntarily.

The study therefore proposes to borrow constructs from both governance and IS literature which would help in explaining the voluntary behaviour. The study proposes to use unified theory of adoption and use of technology (UTAUT) from the information technology literature and civic voluntarism model (CVM) from political participation domain. UTAUT proposed by Venkatesh et al. (2003) explain the behavioural intention by integrating multiple theories from information systems (IS) literature and thus is considered better than the earlier theories like theory of planned behaviour, theory of reasoned action , diffusion of innovation theory, technology acceptance model etc. The present study aims to understand only the intentions and not the actual behaviour as in an Indian context e-governance for financial achieving financial resilience has not been implemented. The primary objective of using CVM is to identify and measure the effect of socio economic factors for civic engagement of citizens for achieving financial resilience for disaster affected people. The contribution of this study is to use political ideology as a construct in the model whose effect would be measured in the behaviour of the citizens.

The remaining paper is structured as follows. Section 2 discusses the theoretical underpinning and discusses how the proposed gap could be addressed using lens of Civic Voluntarism Model. Section 3 details the research design section and discusses a systematic approach for data collection. Section 4 reports the results and findings of the proposed hypothesis. Section 5 discusses key findings in the context of the research problem and how results of hypothesis testing have been used to extend the body of knowledge in explaining financial resilience of victims of natural disaster. Section 6 offers conclusion with theoretical and practical implications along with guidelines for future scope of research.

BACKGROUND OF RESEARCH

There is a steep rise in the requirements of funding for humanitarian and disaster relief operations in comparison to charity (Besiou et al., 2018). Most of the disaster relief operations performed involve essence of crowdfunding to sustain its activities. Crowdfunding is identified as “a collective effort by people who network and pool their money together, usually via the Internet, in order to invest in and support efforts initiated by other people or organizations” (Ordanini et al., 2011: 444).

Over the years, there has been a transition from physical channels of crowdfunding to internet based crowdfunding. Belleflamme et al. (2015) classified crowdfunding into four categories: lending based crowdfunding, equity based crowdfunding, donation based crowdfunding and reward based crowdfunding. Most of the crowdfunding processes have a defined financial goal which they intend to achieve in a stipulated time period. Donation based crowdfunding is the only category of crowdfunding which involves investments without any tangible or intangible returns. Donations made towards DRO is one of such sub type of donation based crowdfunding.
In a recent report published by Centre of Disaster Philanthropy (2018) discussed some of the key factors which resulted in a donation towards disaster relief operations by US citizens in 2017 and 2018. Some of the key factors were: intensity of disaster, media coverage of the disaster, connections of the donors with the disaster hit geographical area, relationship with the affected masses, reputation and efficiency of the recipient of the crowdfunded agency, frequency of disaster etc. Payment method also plays an important role in understanding the behaviour of donation. Study from the report also reveals that a total of 61% of the total households in US made a contribution towards disaster relief operations of which more than half of the donations were made either by cash or cheque. Other forms of donations also included donations made by organization website, direct wire transfers and crowdsourcing platforms.
Humanitarian supply chain differs from commercial supply chain in a lot of ways. One of the key differentiator is the degree of uncertainty involved in the process. In an uncertain environment, it becomes more important to discuss the degree of uncertainty in financial donations as it contributes towards the well-being of the affected masses. Burkart et al. (2018) highlighted that role of funding and its importance in a humanitarian context primarily due to shortfall of collected funds from global agencies. Most of the crowdfunding agencies use multiple strategies to enhance the amount of donations either by focusing on frequency of donations or by capturing larger masses. Studies also confirm that majority of the donations made towards crowdfunding are involuntary in nature while a handful of them are voluntary. While studies have not shown similar trends across different categories of crowdfunding (Belleflamme et al., 2015), it becomes important to understand the nature of voluntariness as a control variable towards behaviour of donors. Donation based crowdfunding unlike the remaining three types is a like a non-returnable boomerang and thus transparency also plays a vital role in the process of donation. Studied related to transparency in humanitarian operations are mostly linked to its accountability (Villa et al., 2017; Starr and Van Wassenhove, 2014; Haavisto and Goentzel, 2015; Beamon and Balcik, 2008). Mejia et al. (2019) studied the effect of operational transparency and conventional transparency on likelihood of receiving donations by private donors.

UNDERPINNING THEORIES

Citizen participation in democracy has a long history and scholars have theorized the phenomenon of citizen participation from different theoretical viewpoints (Kim et al., 2006). It is interesting that scholars have addressed key issues of citizen participation in democracy and politics but have missed some of the key aspects of continuous participation and lifelong participation. Unlike learning where the engagement is longer and systematic, citizen intervention is not that frequent and regular. This partially explains why some individuals sometimes participate while other are reluctant in participating in governance issues of the country (Pattie et al., 2003; Uthaman and Ramankutty, 2017). Some frequently used theories in the space of civic culture include the famous theory of polyarchal democracy, rational actor theory, civic voluntarisms model etc (Oni et al., 2017). While most of these theories discussed civic participation in the offline democracy, extensions of these were also tested for e-governance with technology often used as mediators or moderators or both in multiple studies. The current form of governance has majorly drifted towards the e-governance and thus it is important to not only check the effect of technology in these theories, it is also important to use multiple layers of theories drawn from different disciplines to understand the phenomenon of citizen behaviour (Khan and Krishnana, 2017; Ryzhov et al., 2015). Behavioural researchers have also claimed that behaviour is an outcome of a systematic thought process which is often driven by perception, experience, societal influence and other social traits (Blackburn, 2014). The changing nature of technology and evolving e-governance has although made the youth more inclined towards it but the technology has not engaged the older generation relatively. The launch of online portals for managing and governing services has a lower rate of adoption by older people as proposed by Xavier and Oliveria(2016) and one of the prime reason stated is their weak rate of communication with the government through online channels. Naher and Krimmer (2005) in their early research proposed Social Media Politics model which discussed a continuous cycle of political communication between media and different political and democratic actors. Studies also confirm that while certain hypothesis related to citizen engagement holds true in the developed nations, they fail in the developing nations (Uzoka, 2008; Jain and Suri, 2017). Recent proposition made by Kolmann and Kayser (2010) explored factors for citizen engagement in the German context by using political participation theory and information systems acceptance theory. It is also witnessed that studies have usually revolved around e-governance portals in developed nations like Germany, France, USA etc where political stability and technology led development has always been higher (Asongu and Nwachukwu, 2016). Also, the studies have generally studied practices related to e-governance for voting, community development, bill payment services, tax payment etc., while none of them has addressed governance related issues in times of uncertainty. Furthermore, studies have not explored the intervention of political culture in driving e-participation which by and large impacts the behaviour of the individuals (Haselip et al., 2015; Singh and Sahu, 2008; Kabra et al., 2017). Results from Serrat et al.(2016) claims that political culture has a positive impact on the participation of citizens , but they have not tested the hypothesis under uncertain or risk averse situations. The objective of the present study is explore the relations of behavioural enablers of e-participation of citizens towards achieving financial resilience post natural disasters. The study is guided by understanding the interdependency of two theoretical school of thoughts which are discussed in the section 2.1. The review of theories is discussed from the viewpoint of the present study and required hypothesis is proposed which would be tested with empirical data.

Unified Theory of Acceptance and Use of Technology (UTAUT)

There have been umpteen cases where various information systems theory have been used to explain the behaviour of the users (Al Athmay, 2015; Orong and Hernandez, 2019; Zhou et al., 2010; Abubakar and Ahmad, 2013). Most of the theories have behavioural intention being used as the outcome which have antecedents like perceived ease of use, perceived usefulness, attitude, subjective norm, societal influence and many more. Each of the theories viz theory of planned behaviour , theory of reasoned action, technology acceptance model, diffusion of innovation theory have used various permutations and combinations of the above mentioned constructs. Venkatesh et al. (2003) developed the unifying theory which integrated these into one theory UTAUT and resolved most of the shortcomings of the same. The theory since its inception has also been validated and extended by multiple authors in different contexts. Scholars have also challenged the use of only UTAUT to explain behavioural traits of the people and some additional theories like UMEGA have also been proposed which have also tried to integrate multiple theories to explain governmental actions (Dwivedi et al. 2017). The model has also helped in explaining e-governance amongst citizens and society. In the context of governance,

the model has been extended with moderating variables like age, gender, voluntariness, experience and culture (Venkatesh et al., 2012; Kahlilzadeh et al., 2017). It is because the theory measures behaviour which then leads to actual usage, the early studies have largely been empirical in nature without actually testing in with longitudinal data.
Studies have indicated that while UTAUT is good for explaining behaviour, it only marks partial behaviour and which calls for explaining the same with other organizational theories (Dwivedi et al., 2017; Dwivedi et al., 2018; Rodrigues et al., 2016). Although, studies have integrated UTAUT with different theories like resource based view theory, stakeholders theory, social capital theory etc, but a mixed theoretical model is missing in the context of explaining the behaviour of citizens when dealing with the government. The challenge becomes multiple as the responses of government affect the behaviour of citizens and it a reciprocatively impact on behaviour rather than unidirectional (Gundelach, 2016; Vatanasakdakul et al., 2017;). The study therefore aims to integrate CVM and UTAUT to explain the behaviour of citizens. The context of present study is unique as it aims to address the challenges of understanding the behaviour of citizens for making e-payment on government portals and contribute towards disaster relief funds.

Need for theory Integration

The central objective of the paper revolves around understanding the attitude of the citizens to participate towards offering payment on the government portals for development of disaster struck areas in the country. Thus, “attitude” towards the behaviour is defined as “an individual's positive or negative feelings about performing a behaviour” (Fishbein & Ajzen, 1975). “Actual Usage” in this study is operationally defined as continuous and consistent usage of crowdfunding platform by financial donors to make a contribution towards disaster relief operations after a natural disaster. As “Actual Usage” is a result of “Attitude” of human being which is empirically tested by earlier literature, it is important to discuss and debate how and why studying attitude is more important as compared to any other construct.

Attitude is measured and empirically validated by multiple studies in the past and forms a key construct in behavioural studies whether it is TRA, TAM, DOI or UTAUT. Attitude irrespective of its roots grounded in these theories have also been instrumental in explaining human behaviour (Hung, Chang, & Kuo, 2013; Akinyemi, Asani, & Adigun, 2013; Jahangir & Begum, 2007; Adesina & Ayo, 2010). Studies have also posted that attitude is a function of beliefs of individuals which are driven by self or by others and they lead to a particular behaviour. Similar connotations were also proposed by Davis et al. (1989) when perceived ease of use and perceived usefulness was used as drivers for attitude. With time immemorial, it has been validated by multiple studies that perception is different forms have positively impacted the attitude of an individual. An in depth analysis of the studies also confirm that most of the researchers have tested the model with people who are not exposed to any uncertain phenomenon (Shaw and Gupta, 2015; Zagefka and James, 2015). Also, a longitudinal study with uncertain situations could also change the attitude of the people. Thus, it is important to include constructs in theoretical models which would capture and measure the effect of participatory role of citizens and their political bend towards understanding their overall behaviour for e-participation of citizens post disasters.
The theoretical model can be referred from Figure 1 and the hypothesis development can be referred to as follows:

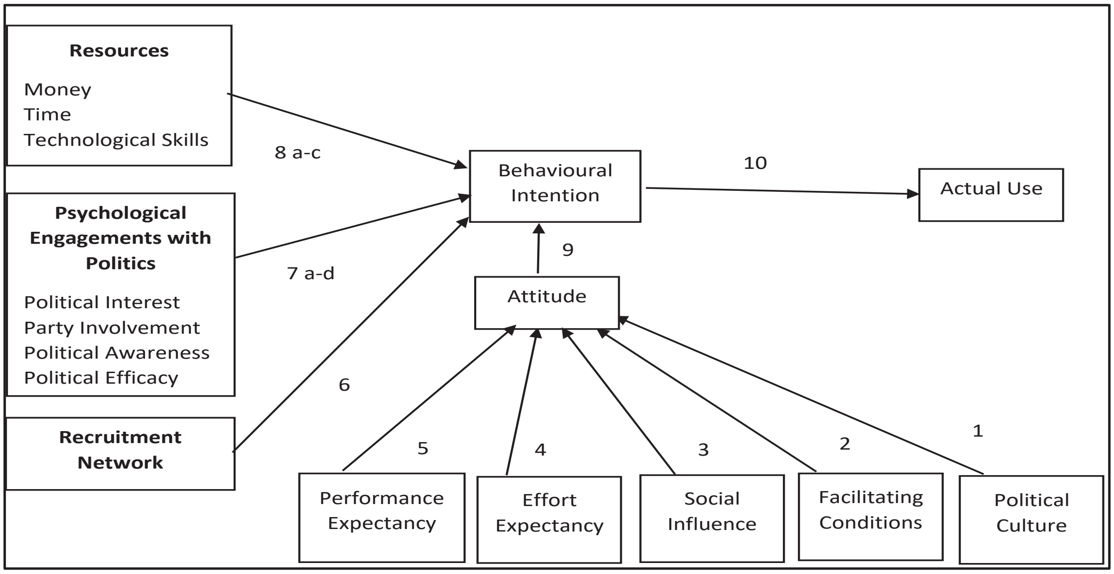


Figure 1: Theoretical Model for the study with proposed hypothesis

Hypothesis Development

It is also worthwhile to note that political beliefs unlike general beliefs are driven not only selfish motives but usually by community benefits. Most of the attitude and perception based studies conducted have measured these constructs at an individual level as the technology, service or process was individual driven. On the contrary, political actions are taken for the community and not for an individual and thus the drivers of belief could not be the same as explained by existing theories. The e-participation as a construct measured in this study would also not benefit the same person, but to a different section of the society through political intervention. While studies from Mizrahi and Vigoda-Godat (2009) confirms that citizen participation in public decision making has a positive impact on the trust on the government, but they have not discussed how trust would impact their behaviour towards participation which is a reverse phenomenon. There is a difference in opinion presented in the literature between the relationship between attitude and action of government in building trust of people and attitude and action of governance in building trust (Jones et al., 2015; Jacob and Singhal., 2017). There exists a hair line difference between the two as on one hand there is government which is a seasonal ruling party while on the other hand it is governance which is a phenomenon and could be run irrespective of the government in power (Paik, 2018; Strömblad and Bengtsson, 2017). Therefore, it is important to relate attitude of citizens with political culture rather than with political parties. The democratic political culture referred to in literature (Loveless, 2013; Tessler and Gao, 2008; Almond and Verba, 1963) is a construct which has found to have relationship with attitude of citizens and thus the present study hypothesizes:

H1: There is a positive effect of democratic political culture (PC) on citizen's attitude (AT) towards e-participation for financial resilience.

The other hypothesis were derived from existing studies on UTAUT which tests the key constructs to understand attitude. These relationships are discussed in literature umpteen number of times are therefore not discussed at length in this paper.

H2: There is a positive effect of facilitating conditions (FC) on citizen's attitude towards e-payment for financial resilience.
H3: There is a positive effect of social influence (SI) on citizen's attitude towards e-payment for financial resilience.
H4: There is a positive effect of effort expectancy (EE) on citizen's attitude towards e-payment for financial resilience.
H5: There is a positive effect of performance expectancy (PE) on citizen's attitude towards e-payment for financial resilience.

It is also discussed and well established in the information science literature that attitude of the respondent is linked with its positive behavioural intention towards using any technology. While this relationship is grounded in literature, it is more important that validate the same in the context of disaster relief operations. While usage of technology and attitude has mostly been studied in the context of technologies which are either less established or the response from the expected users were uncertain. The present study extends the contexts also towards understanding the same for a cause where the rate of returns of uncertain and at times negligible. It is in the light of the same that the study explores the behaviour of the respondents from civic voluntary model's viewpoint as well. It can therefore be hypothesized that:
H6: There is a positive effect of recruitment network (RN) on citizen's behavioural intention (BI) towards e-payment for financial resilience.
H7a: There is a positive effect of political interest (PIN) on citizen's behavioural intention towards e-payment for financial resilience.
H7b: There is a positive effect of political involvement (PIV) on citizen's behavioural intention towards e-payment for financial resilience.
H7c: There is a positive effect of political awareness (PAW) on citizen's behavioural intention towards e-payment for financial resilience.
H7d: There is a positive effect of political efficacy (PEF) on citizen's behavioural intention towards e-payment for financial resilience.
H8a: There is a positive effect of financial resources (FR) on citizen's behavioural intention towards e-payment for financial resilience.
H8b: There is a positive effect of time to engage (TE) on citizen's behavioural intention towards e-payment for financial resilience.
H8c: There is a positive effect of knowledge of using technology (KT) on citizen's behavioural intention towards e-payment for financial resilience.

It is debated that attitude of a respondent is used to explain his/her behaviour and behaviour invariably explains the actual usage of technology. Studies have tested these propositions in multiple scenarios and for multiple technologies. Most of these scenario are linked to the beneficiaries being the respondents themselves. However, literature pertaining to technology usage of e-governance and society indicates that citizen participation and their behaviour differs on various occasions and therefore is unpredictable. However, most of the studies have reported that attitude of citizen's positively influence their behaviour which in turn positive effects the actual contribution towards societal development. It can therefore be hypothesized that:

H9: There is a positive effect of attitude on citizen's behavioural intention towards e-payment for financial resilience.
H10: There is a positive effect of citizen's intention on making actual payment (AP) of financial resilience.

RESEARCH DESIGN

Survey Instrument

The proposed theoretical framework and hypothesis is a resultant of a three step approach which began with identification of constructs and operationally defining them from interdisciplinary literature in order to holistically define the scope, followed by identifying the items for measurement of each construct. Lastly, the final instrument was pre tested for content validity by experts after which the final instrument was used for data collection. We reviewed literature and scales used in crowdfunding and humanitarian supply chain related to objectives of the study. The literature was used to finalize the operational definitions of the study along with the first draft of the instrument with each construct and its corresponding items. Referring to the operationalizing concepts laid by of Dubey et al. (2018), we adapted the scale in the context of humanitarian supply chain and integrated it with crowdfunding. Each of the items was measured on a five point Likert scale to ensure higher variability among survey responses (Dubey et al. 2019; Dwivedi et al. 2013; Kim et al., 1999).

As a last step before proceeding with data collection, the instrument was pre tested for its content validity. The instrument was shared with eight experts and experienced researchers in order to share their feedback on the appropriateness of items, clarity in presentation and reducing the ambiguity in placement and formulation of the instrument (De Villis, 1991). Using Dillman (2007) guidelines, we further requested the experts to assess the instrument and validate that the items and constructs are indicating and appropriate for the context of the study. The comments of the panel of researchers were shared with each other and a unanimous suggestions were incorporated in the questionnaire to revise it. (Chen and Paulraj, 2004). As the instrument would be used for collecting data from financial donors in India, the instrument was also shared with the disaster wing of Oxfam over an email. The agency being a pioneer in involving poverty alleviation and financial aiding of affected people post disaster was used for the second round of content validation. We requested their feedback and required changes were made to make the instrument highly valid on content.

The final instrument was then used in the pilot study before actual data collection with the participants of World Disaster Congress held in India. The questionnaire was shared with the potential presenters in the international conference in order to ensure that the final instrument is valid and reliable. We received a total of 58 responses out of a total of 247 target respondents. The results of the pilot survey were useful in finalizing the structure of the instrument which was then used for the data collection process.

Data Collection

The study used individual financial donors made towards Kerala floods as unit of analysis for the study. The study used Kerala floods because of its degree of severity. Over 11 straight days of tempestuous rainfall, nearly 25 trillion litres of water fell on Kerala—an area of 38,800 square kilometres cramped with mountain ranges; third-highest population density in the country; and, 44 rivers with 61 dams—with apocalyptic fury (Venkatesh and Kuttappan, 2018). The state government has estimated the preliminary loss at Rs 20,000 crore, which is around 15 per cent of the state's GDP estimate for 2018-19. According to risk management agency, Care Ratings, floods have affected more than four million people, a significant percentage of them labourers. In August alone, people would lose wages worth Rs 4,000 crore. More than one million people are in relief camps that would take around Rs 300 crore a month to maintain. More than 12,000 kilometres of roads have been damaged hindering speedy relief and rebuilding operation. Kerala is not prone to catastrophe unlike north eastern states of India. Thus, it belongs to a geographic category of non-recurring disaster. The study gathered data by circulating questionnaire on the platform of Humanitarian ID. From the earthquake in Nepal in April 2015 to the hurricane in Haiti in October 2016. Humanitarian ID is the first choice to manage humanitarian contact lists in large scale emergencies. The list have more than 18,000 responders across the globe connected and working towards the disaster related operations. The questionnaire was posted as a weblink and the administrator was requested to spread this across the network. The respondents were also asked to repost the link in different channels. Data collection was done between July 2018 to September 2018 and two waves of postage of questionnaire was done over the span of 6 weeks. One of the pre requisite of submitting the questionnaire and checking the eligibility of respondents was submission of proof of payment made towards Kerala relief fund which are operationally defined as donors in this study. We received total of 439 responses of which only 354 were valid based on their eligibility of respondents and authenticity of the data. Data was then tested for missing values and the overall sample size of 321 valid data points were used for the study. Brief profile of the respondents can be referred to from the table below.

Demographic Variables		No. of respondents
Age	25-30 years	78
	30-35 years	62
	35-40 years	82
	40-45 years	25
	45-50 years	43
	50+ years	31
SEC group	A1	45
	A2	110
	B1	94
	B2	46
	C	13
	D	8
	E1	3
	E2	2
Geographical Distribution based on their native state	North India	52
	South India	129
	East India	36
	West India	69
	Central India	35
How much time have you stayed in Kerala?	Haven't stayed	123
	Less than 1year	34
	1 – 3 years	70
	3- 5 years	54
	5+ years	40
Education	No formal education	35
	Primary School Qualification	67
	Secondary School Education	58
	College Qualification (Diploma/ Certificate)	13

	Professional Qualification (CA/CS)	7
	Undergraduate Degree	104
	Post Graduate Degree and Above (Master's/PHD)	37
Gender	Male	179
	Female	142
	Others	0
Nature of Job	Not Working	106
	Part Time Working	27
	Full Time Working	188

Table 1: Demographic Profile of Respondents

The distribution of respondents from Table 1 depict that the their exists almost an inverse relationship between the age of the respondent and its financial contribution through e-governance portals. Earlier studies have also discussed similar trends wherein older generation are still not as much tech savvy as the younger generation is. This could be one of the reasons for the trend seen in this study as well. It is also seen that upper most consumption class was the mostly contributing towards financial well-being of the society. This although does not follows the trends of earlier studies wherein the richer class don't express their interest in societal development, but some of the recent studies have also confirmed the trend along the lines of the present study as well. Lastly, the geographical distribution of respondents was skewed towards Southern part of India. It is seen that most of the people who live in south Indian states have made the contributions. The study makes the results interesting as a sizable number of people who have not lived in Kerala have also contributed towards the development of the state.

As the data collection is empirical in nature, it is important to check the non-response bias in the data. We employed Armstrong and Overton (1977) suggestions for checking the responses received from non-respondents, early respondents and late respondents. By taking a sub-sample of 55 respondents, we tested that there is no significant difference between early and late respondents in our study for the constructs of the study. We also did not observe any significant difference between the profile of the respondents and non-respondents. Thus, our study is free from non-response bias.

The proposed hypothetical model was tested using Structured Equation Model (SEM). Unlike regression based methods, SEM allows simultaneous modelling of relationships among multiple independent and dependent constructs (Chin, 1998; Urbach and Ahlemann, 2010). SEM has its advantages in distinguishing between independent and dependent variable and more importantly between endogenous and exogenous latent variables (Hair et al., 2010). Preacher and Hayes (2017) distinguishes the use of covariance based SEM and partial least square based SEM and the present study fulfils majority of the criteria for PLS-SEM as it is more appropriate in an exploratory research and the developed framework is new. Smart PLS was used to perform the analysis and it takes care of the distribution of the data and does not care about its normalcy.

MEASUREMENT MODEL

The model was tested for its composite reliability followed by discriminant validity and confirmatory factor analysis. The results from Table 2 report that average variance extracted (AVE), construct reliability, Chronbach's alpha and range of factor loading of all the constructs and their respective items. As proposed by Chin (1998) and validated by multiple studies that any item in a questionnaire tends to be retained if its standardized loading should be greater than 0.7. The preliminary analysis dropped 7 items from the initial questionnaire and the remaining items with their responses were used for further analysis. A bootstrap analysis of 1000 resample gave significant results for the remaining items at 0.01 significance value. The results showed a considerate improvement in the construct reliability and convergent validity after dropping those seven items. The results of the final items and their corresponding values can be referred to from Table 2. Smart PLS 3.0 was used to derive the results and it was found that the construct validity of each of the items of every construct was found to be significant expect performance expectancy (PE) and Societal Influence (SI).

Constructs	**AVE**	**Composite Reliability**	**Cronbach's Alpha**	**Factor Loading Range**
Actual Payment (AP)	0.67	0.868	0.791	0.732-0.784
Behavioural Intention (BI)	0.721	0.921	0.838	0.714-0.821
Performance Expectancy (PE)	0.583	0.831	0.722	0.783-0.803
Effort Expectancy (EE)	0.741	0.853	0.845	0.696-0.752
Societal Influence (SI)	0.793	0.801	0.821	0.782-0.837
Facilitating Conditions (FC)	0.686	0.911	0.789	0.703-0.741
Political Culture (PC)	0.821	0.854	0.901	0.752-0.779
Attitude (AT)	0.703	0.795	0.776	0.762-0.793
Recruitment Network (RN)	0.592	0.742	0.428	0.784-0.721
Money (M)	0.662	0.884	0.689	0.741-0.752
Time (T)	0.690	0.715	0.862	0.803-0.852
Technology Skills (TS)	0.741	0.846	0.752	0.752-0.769
Political Interest (PI)	0.799	0.851	0.832	0.743-0.801
Party Involvement (PIV)	0.616	0.721	0.699	0.811-0.818
Political Awareness (PA)	0.729	0.852	0.724	0.704-0.736
Political Efficacy (PE)	0.635	0.782	0.741	0.751-0.773

Table 2: Reliability and validity results

It was also found that the AVE value of all the constructs was more than the threshold of 0.5. The convergent validity of the items was tested by examining the cross loadings and the results confirmed that there was no measurement items whose loading was more on any other construct which was not theoretically assigned to it. Hair et al. (2010) also proposed that the discriminant validity was examined by comparing square root of AVE of each construct with the correlation score of each pair of latent variable and it was found that the square root of AVE was greater for all the constructs. Table 3 can be referred to for the results. It can be observed that the square root of AVE score for the entire model is more than the threshold value of 0.7 as suggested by Hair et al. (2010). With the two satisfying criteria, we can confirm that these constructs share more variance with their respective measurement items and have lesser association with other constructs. The results fall in line with most of the previous studies which have tried and integrated multiple theories to test a phenomenon. The next step is to test the proposed hypothesis in this study.

	AVE	PC	FC	SI	EE	PE	RN	PIN	PIV	PA W	PEF	FR	TE	KT	BI	AT	AP
PC	**0.685**	0.73															
FC	**0.845**	0.14	0.84														

SI	**0.593**	0.32	0.23	0.78													
EE	**0.784**	0.21	0.26	0.34	0.73												
PE	**0.732**	0.17	0.31	0.21	0.22	0.85											
RN	**0.821**	0.31	0.14	0.27	0.24	0.26	0.81										
PI N	**0.694**	0.24	0.29	0.29	0.30	0.37	0.38	0.76									
PI V	**0.542**	0.15	0.36	0.31	0.17	0.42	0.29	0.36	0.69								
PA W	**0.674**	0.19	0.39	0.37	0.16	0.21	0.38	0.41	0.41	0.73							
PE F	**0.692**	0.30	0.25	0.14	0.29	0.27	0.36	0.13	0.36	0.06	0.80						
FR	**0.731**	0.05	0.41	0.17	0.26	0.31	0.28	0.16	0.29	0.42	0.37	0.71					
TE	**0.682**	0.25	0.10	0.20	0.25	0.28	0.31	0.17	0.31	0.21	0.27	0.29	0.79				
KT	**0.741**	0.21	0.09	0.29	0.10	0.18	0.16	0.25	0.26	0.27	0.25	0.12	0.23	0.75			
BI	**0.841**	0.31	0.15	0.37	0.34	0.13	0.14	0.26	0.31	0.31	0.13	0.31	0.25	0.16	0.88		
AT	**0.703**	0.23	0.19	0.31	0.17	0.41	0.23	0.31	0.13	0.33	0.16	0.22	0.21	0.20	0.32	0.72	
AP	**0.751**	0.22	0.26	0.17	0.27	0.30	0.26	0.39	0.25	0.43	0.35	0.24	0.11	0.34	0.21	0.29	0.86

Table 3: Correlation and AVE matrix between constructs

The hypothesis testing is done using multiple tests and one of the prominent ones is to check for the coefficient of determination of the model. The value of R square (coefficient of determination) is the measure of how much variation which is explained by the model. Three values for R square were calculated for attitude, behavioural intention and actual usage as all the three were dependent variables which were explained by constructs of UTAUT, CVM and the integrated model respectively. It was found that the overall exogenous factors explained 42% of the variation in the intention of use of e-payment services for financial wellbeing. While the value is acceptable by most of the published literature, it was worthwhile to discuss the reasons and justifications for the same which is done in the subsequent section. It was also found that the CVM model based independent constructs explained 31 % of the attitude towards use of e-payment services. Lastly, the behavioural intention only explained 21.2 % of the variation of the actual usage which confirms that intention does not explains the actual usage solely (Urbach and Ahlemann, 2010) and there might be other factors apart from the behavioural ones which could explain the phenomenon of usage.

Hypothesis Relationship	Path	T -statistics	Hypothesis Testing Results
PC → AT	0.367	5.452	Supported
FC → AT	-0.132	1.113	Not Supported
SI → AT	-0.112	0.258	Not Supported
EE → AT	0.212	4.235	Supported
PE → AT	0.298	3.163	Supported
RN → BI	-0.021	2.413	Not Supported
PIN → BI	0.201	2.311	Supported
PIV → BI	0.063	0.117	Supported
PAW → BI	0.189	0.995	Supported
PEF → BI	0.303	2.907	Supported
FR → BI	0.226	2.112	Supported
TE → BI	0.301	4.234	Supported
KT → BI	0.354	5.669	Supported
AT → BI	0.472	7.583	Supported

BI →AP	0.498	7.683	Supported
	R square	***F square***	
Attitude	***0.378***	***0.623***	
Behavioural Intention	***0.416***	***0.352***	
Actual Payment	***0.214***	***0.289***	

Table 4: R square and path co-efficient matrix (hypothesis testing)

The next step is to calculate path coefficients in order to find the degree of relationship between constructs (Wixom and Watson, 2001) which was carried out using bootstrapping with 1000 resampling process. The results of the bootstrapping process are presented in Table 4. The results were assessed by looking at the coefficient paths using the t-statistics. The process helped us in calculating the structural path coefficients of all the hypothesized relationships. The effect size of the overall model was then calculated using Cofen's F square value and are presented along with their respective value of their coefficient of determination in Table 4. The results show that political culture, performance expectancy and effort expectancy have higher effect on their endogenous LV which along with the distributed effect of the other constructs have multiplicated larger impact on the endogenous LV behavioural intention. Seven out of ten hypothesis were supported while the remaining six were not supported (Refer fig 2). The results from hypothesis testing reveal that there was a positive significant impact on attitude by political culture, performance expectancy and effort expectancy. The other UTUAT relationships of the second degree also have a positive impact on each other viz. attitude having positive impact on behavioural intention and it in turn positively impacting actual usage of e-payment portals by citizen for financial development of disaster victims. The results also revealed that constructs related to societal (facilitating conditions, societal influence) and psychological engagement (political efficacy, political awareness) have weak significant predictors to behavioural intention of citizens to make financial contributions. Lastly, it is seen that recruitment network, resource based constructs (technological skills, time, money) and part of psychological constructs (party involvement and political interest) have non-significant predictors for understanding behavioural intention of citizens making financial contributions for disaster victims. It is unlike the earlier studies which have used resource based view theory and have shown that resources play an important role in the development of the community, the present study indicates that it would not really impact the behaviour of citizens to make financial contributions.

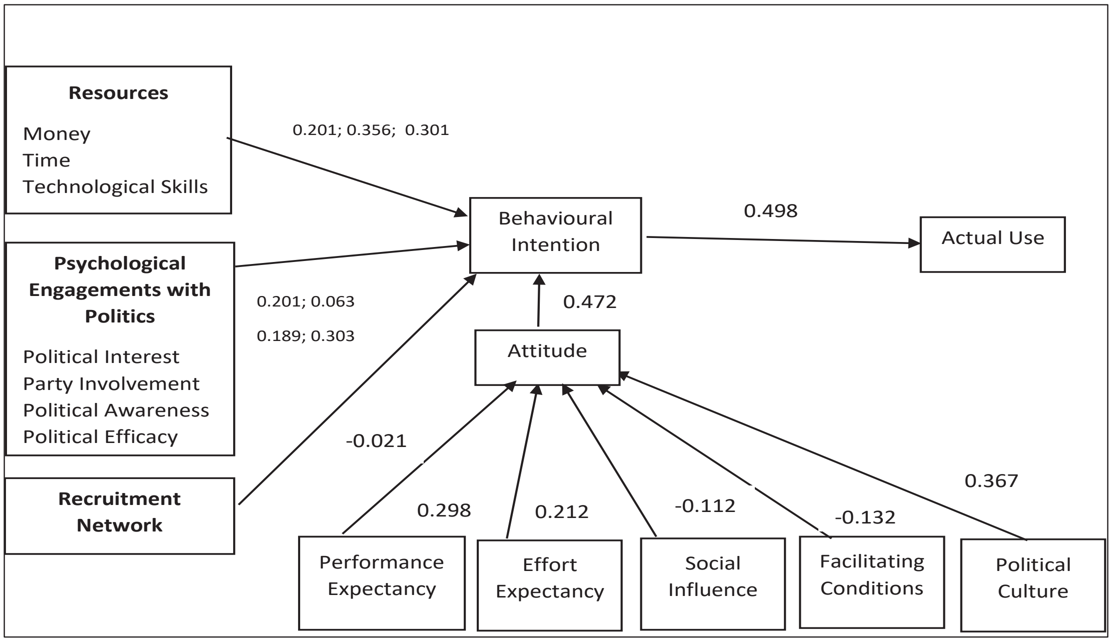


Figure 2: Results of hypothesis testing and structural model

Discussion of Results

Financial resilience of disaster affected victims is the key theme of the study and the study has understood the effect of political culture and perception of citizens for e-payment towards the financial development of victims. The participation of NGO and donations routed through them has been studied in the past while with the advent of technology, the study integrates the multiple theories from diverse literature to understand the attitude of citizens of the country who are involved into financial donations for emergency relief victims. Results reveal that effort expectancy, performance expectancy and democratic political culture have significant positive impact on attitude of the citizens for making e-payment towards Kerala disaster relief funds. The results reveal that H1 has the highest significant value amongst all the relationships. The study has extended UTAUT model with democratic political culture and its positive impact confirms that while measuring e-payment behaviour especially in the governance setting, it is important to include and test the effect of political culture (Abdulwahab, 2012). The results partially is in line with the cultural models in politics proposed by Naranjo-Zolotov et al.(2018) which confirms that participation of citizens are affected by governance of the country. As the payments were made to the state government of the country, the results may vary with donors perception towards other states of the country. The data had a mix of people who were and who were not the natives of Kerala which confirms that overall the effect of domicile does not really play an important role in the analysis. The study also presents that societal influence and facilitating conditions have played a non-significant role towards attitude of the donors. Most of the earlier studies (Hihorst, 2003; Poblet, 2010; Pakdebugree et al., 2011; Zhao et al., 2015) have confirmed that societal influence and facilitating conditions happen to be strong positive significant contributors towards attitude of the end users. But as the present study deals with behaviour of donors for a financial cause, both the constructs have not contributed significantly. As disaster is a seasonal or at times rare phenomenon and impacts the economy in the short run, the donors also are active for a short run. Thus, unlike studies which measure attitude for mostly a technology or a service which exists and will exist in present and will exist in future also, the present study measures the attitude which is also seasonal along with the disaster. The financial contributions made are based on an event and thus the attitude of donors are perpetual and may not be permanent (Lai, 2017). The study also confirmed that 64.3% of the respondents believe that making a payment through e-payment channels for KCMDRF would contribute towards financial resilience of the affected people. It also confirms that 81% believe that e-payment would help in improving status of life of the affected individuals and 72.4% believe that local and state governance play a vital role in achieving financial resilience in a disaster affected area. These results are mostly consistent with studies pertinent to e-governance (Mansoori et al., 2018; Becerra et al., 2015) in the past but are largely an outcome of role of citizens in e-democratic policies and practices made by the government. Citizens are also an equal and important stakeholder in achieving financial resilience as per the findings of this study.

While Weng et al. (2015) confirm that past experience contributes to citizen's perception towards e-governance, but this study does not confirm such results as the geography of Kerala has not faced floods of similar intensity in the past. One prime reason which is why citizen's democratic participation and engagement in high could be political image of the state government in the development of the state over the years. Report by Kerala government (2018) confirms that Kerala has been one of the most successful states in terms of its economy in the past 2 decades and that might have contributed towards positive perception of individuals towards its development. Choudhary and Neeli (2018) asserted that good governance also involves launching of schemes and policies are the right time and involve citizens wherever required. The present study adds to the findings as e-payment portals were launched and easily accessible to the local masses through easily accessible online channels (Strömblad and Bengtsson; 2017; Shi et al., 2010). Moreover, the government also promoted people to contribute for the development which would have acted as social stimulus for making financial donations. Under times of crises, citizen engagement is mostly done through online channels by government and similar instances were also seen with Kerala government which has a total of 762 tweets urging for making financial contributions which were retweeted by 1.3 lakh people (GOI, 2018). This ensures pro activism of government which would have ensured a positive attitude leading to positive behavioural intention and eventually payment. E-payment has an important construct called "privacy and security" which has been discussed by past literature but the present study did not incorporated it as the transactions were made through secure channels embedded in mobile apps like Google pay and therefore chances of commitment a fraud was negligible. While the results of delineated UTAUT model reveal that most of the constructs contribute towards attitude, the results also reveal that attitude also significantly positively transforms into behavioural intention of citizens (Munene et al., 2018; Meth et al., 2015). It can be confirmed with consensus that attitude has a positive impact on behavioural intention, but because behaviour was also measured from CVM model, it is important to look at the constructs of CVM to understand their effect on the behaviour of citizens. Citizen engagement through government is also a result of active discussion forums, chat sessions, inter citizen discussions (Ertas, 2015). The discussion also helps in arising political alignment of the citizen and vice versa (Hofman et al., 2016) but the study did not find any such evidence in this regard as the e-payment portals had no option of discussing with peers and getting to know their opinion. The government also did not displayed the total amount of money transacted to the public which probably is one reason why facilitating conditions and societal influence did not significantly contributed towards developing the attitude. Apriori to this, political beliefs of the public would have helped them to make payments for Kerala more than other states where floods caused human losses like Assam and Bihar. That's primarily one of the reasons why we could observe that the hypothesis proposing positive influence of political awareness towards

making financial contributions was supported in this study. Results also confirm that political interest of citizens positively impacts the intention to pay. Shi et al. (2018) asserted that political involvement of citizens and their know how of functioning of governance makes them more participative and in turn helps them to take decisions to support or not support government initiatives. In a recent study similar results were confirmed by Garrity(2015) which confirmed that digital media helps citizens to know about governance and eventually take actions against their activities. The participation of citizens are relatively more when their development is concerned rather than community development (Aldrich et al., 2015; Linnerooth-Bayer et al., 2015) while this study proves that there has been a significant positive impact of political interest of donors in making up their mind to make financial contribution.
Social Media channels have also helped to increase individuals digitally efficacy and contextually it has helped them to know about their government as well (Meth et al., 2015; AlAwadhi and Morris, 2008). Results from the study also confirms that political efficacy has a significant positive impact on e-payment behaviour of the citizens. One of the primary reason for the positive and significant impact could be the wider coverage of Kerala floods by media and corrective actions which government took for the disaster affected people. While political efficacy is not an after result of a short run coverage of Kerala politics by media, but over a prolonged effect of the how government initiatives have helped in the development of the state. This study results are in line with previous literature (Balckburn, 2014; Suzuki, 2015) as well which confirms that a well-read citizen and stronger media presence contributes to higher degree of political efficacy. With stronger positive political efficacy, they tend to contribute towards government initiatives for development of society and disaster is one such natural epidemic which needs political intervention to achieve both financial and economic resilience. Disaster has also helped economies grow, with Japan as a prime example. The growth is therefore dependent on strong governance and more than that citizen's faith on governance. E-payments by citizens for development of Kerala adds to similar theoretical reasoning with similar results.
Recruitment network plays a significant role in mobilizing citizens and inclining them towards political parties and their participation in political activities (Deen, 2015). Studies confirm that recruitment network has shown positive effect on citizen engagement by political propagandas organized by political parties (Oni et al., 2017; Paton and Johnston, 2017). However, this study has failed to support the hypothesis. It can be understood that the context of recruitment network remains insignificant in disaster relief operations as government does not organize any political briefings to involve citizens support. While financial funding has donors contribution as a major source, it is unlikely that government makes a point to use political pressure to draw money from people. The results support the study by Oni et al.(2017) which confirms that recruitment network comes in handy when there is a benefit of the government while the present study aims to understand the intentions of donors for the development of disaster hit community and government only acts as a mediator which is why the hypothesis has not been accepted. The results also confirm mildly significant values for sub constructs of resources of CVM. Literature has discussed that CVM resources factors comprises of three elements: money, free time and civic skill which help the citizens to participate in politics of a nation. As the respondents were spread across the country and donated significant amount of money for flood relief victims, corollary to the proposition we can confirm that those were economically sound, had enough time to take decision and were civically equipped to understand the dynamics of their contribution. The study although has not discussed whether the donations were made voluntarily or under some obligation as it would add to explanation of their behaviour to make e-payments on governance portals. People who belong to higher SEC level would be more involved into politically activity (Weng et al., 2015) and would have then made contributions for the development. The

demographic profile of the present study also confirms the same. That's probably one of the prime reason all the three sub hypothesis were accepted in this study.
It can be studied from past literature that most of the donors have either experienced similar situations earlier in their life or have been psychologically affected by the current one (Choudhary and Neeli, 2018). An in depth discussion with the respondents revealed that most of them have made financial contributions for disaster relief operations in the past. It is important to note that none of them made voluntary contributions in the past while a handful of those made it for out study which indicates that it would be worthwhile to explore the control variable of voluntariness in the present model for future study. 83% of the respondents were operating a smart phone from the past decade and were exposed to online money transactions through mobile apps and internet banking. 93% of the respondents have also used e-payment earlier to government in forms of utilitarian amenities and taxes. These could some reasons why resources and its sub constructs have received a positive significant impact on e-payment intentions of the respondents.
Figure 2 denotes the sig value of all the hypothesis. The overall model depicted that 3 out of 10 hypotheses were found to be insignificant. The findings also reveal that although resources construct had positive impact on behaviour, the impact was higher for constructs like performance expectancy, effort expectancy and political culture towards behavioral intention of citizens to make payments. Thus actual payment which is driven by intentions of citizens is a function of constructs drawn from information science theory (i.e. UTAUT) and governance theory (i.e. CVM) but behavioral traits are more driven by IS theories. The study recorded non-significant results of hypothesis and apart from intuitive and literature supported reasons which are mentioned above, there could be some challenges based on research design as well. The diversity of respondents in terms of their geographic location, their domicile, their SEC and most importantly the timing for making payment could lead to hypothesis not being accepted. Their interest in state democracy and their views about the same could differ from central government which could have also be reasons for insignificant hypothesis. CVM based studies in the past have also ascertained that role of citizens and their educational qualification and their experience plays a significant role in their participation. The present study has taken care of both the aspects but could have incorporated the effect of frequency and impact of floods on the psychology of the citizens in the country. Due to dearth of studies measuring the control effected these variables, the present study could be extended with inclusion of these control variables.

THEORETICAL AND PRACTICAL CONTRIBUTIONS

The study extends single stage theoretical viewpoints of explaining behavioural intention of viewpoints of citizens by integrating CVM and UTAUT. The study also extends UTAUT framework by adding political culture as a construct for explaining attitude and therefore the behaviour of citizens. The study discusses the concept of e-payment of on a government portal and therefore integration a governance and a information systems theory makes the theoretical model more robust. While earlier studies have tried to integrate CVM models with TPB and TAM models but they have also highlighted that the models could be extended with better theories. As UTAUT incorporates multiple IS theories together, use of UTAUT with CVM would produce robust results than its constituent theories. In a disaster related scenario, where e-payment practices are time bound, it becomes interesting to understand momentary behaviour of citizens rather then continuous behaviour which most of the earlier studies have discussed. The study also presents a contextual understanding of behaviour under disaster which is a seasonal event which causes disruption in financial and economic status of people. The understanding of politics and citizen's thoughts and views about the government could also explain their payment behaviour towards governance portal for development of society. Thus, the study tests the hypothesis from technological constructs and governance constructs to explain behaviour and therefore the contribution of citizens towards redevelopment of Kerala after the floods. The study also uniquely places itself in bringing forth concept of voluntary payment of citizens rather than forced and mandate payment like taxes and utility bills which also cater towards development of the country. While trust and security are key constructs which most of the technology based behavioural studies discuss, the present study omit them as the payment is recorded through UPI and bank transfer on government portals. The inclusion of attitude also adds an intermediary construct to explain behaviour which earlier theories like theory of planned behaviour used. The study however does not use moderating variables like gender or age or SEC levels as earlier studies have pointed out that these demographic factors have insignificant effects on developmental activities. The study can be further mediated with knowledge based constructs which could explain behaviour of citizens and their contributions for social cause and development. Results also reveal that societal influence and facilitating conditions which play an integral role in most of the behavioural studies have proved to have negligible effect on behaviour of citizens in this case. The social interaction of citizens making financial contributions is minimal which could be further controlled and its combined effect could be measured for understanding payment behaviour. Since political intervention is a mandatory activity in achieving financial resilience post natural disaster, the study could separate the role of non-government bodies and citizens for making financial contributions and extend this model further.

CONCLUSION AND LIMITATION OF STUDY

The study contributes to e-payment literature by integrating governance and information systems theory. The integration of constructs from both the theoretical literature brings a unique model in the context of financial relief operations for victims post disasters. The study tests the model in an Indian scenario where state government plays a different role from central government and thus the study brings fresh perspectives by discussing the local governance viewpoint in comparison to governance in general. Discussion of Kerala floods as a case study makes the results more interesting as the state is better in a lot of parameters like literacy and governance and yet faced one of the biggest floods in Indian history. The study collected behavioural data from people who contributed in rebuilding of Kerala by their financial contribution towards KCMDRF and the results highlighted some behavioural patterns which earlier studies on technology and governance has not recorded. The study found that three out of ten hypothesis found insignificant relationships and counterintuitive arguments were presented in the study to understand the plausible reasons behind it. The study tested a new belief which is political culture as a part of UTAUT theory which became an antecedent to attitude which would explain behaviour of citizens. The results contribute towards theory extension for both CVM and UTUAT which is first of its kind and more importantly in a disaster scenario. The study also highlights the importance of understanding state politics by citizens which could impact their decision for making a contribution. The study presents a unique amalgamation of theoretical lenses to understand how political know how, interest and alignment of citizens affect their decision making towards development of the state under crisis situation. Disaster relief operations have financial resilience as one of the key indicators and the study posits political culture which makes the study unique as although “political culture” is a construct developed over time but its effect needs to be captured for a short run payment window. The study also has some shortcomings and could be further developed with careful diligence in other disaster scenario. The results cannot be generalized as the intensity of disaster, political understanding of citizens for different states and more important the degree of financial contributions made by the citizens differ. The present study has used data from different parts of the country which could have given a heterogeneous responses while homogeneity could be considered by collecting responses from people with similar domicile, SEC levels and political interests. The study only measures behaviour of citizens contributing towards governance portals while it could be extended to NGO's and other private organizations also. It is also worthwhile to test the effect of voluntariness in making financial contributions to make the model robust.

REFERNCES

AbdulWahab, L. (2012). A Modified of the Unified Theory of Acceptance and Use of Technology (UTAUT) from Users' Perspective of Telecentre in Nigeria (Doctoral dissertation, Universiti Utara Malaysia).

Abubakar, F. M., & Ahmad, H. B. (2013). The moderating effect of technology awareness on the relationship between UTAUT constructs and behavioural intention to use technology: A conceptual paper. Australian Journal of Business and Management Research, 3(2), 14-23.

Al Athmay, A. A. A. R. A. (2015). Demographic factors as determinants of e-governance adoption: A field study in the United Arab Emirates (UAE). Transforming Government: People, Process and Policy, 9(2), 159-180.

AlAwadhi, S., & Morris, A. (2008, January). The Use of the UTAUT Model in the Adoption of E-government Services in Kuwait. In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual (pp. 219-219). Ieee.

Aldrich, D. P., Sawada, Y., & Oum, S. (2015). Community, Market, and Government Responses to Disaster. In Resilience and Recovery in Asian Disasters (pp. 1-16). Springer, Tokyo.

Asongu, S. A., & Nwachukwu, J. C. (2016). The role of governance in mobile phones for inclusive human development in Sub-Saharan Africa. Technovation, 55, 1-13.

Banomyong, R., Varadejsatitwong, P., & Oloruntoba, R. (2017). A systematic review of humanitarian operations, humanitarian logistics and humanitarian supply chain performance literature 2005 to 2016. *Annals of Operations Research*, 1-16.

Barkan, S. E. (2004). Explaining public support for the environmental movement: A civic voluntarism model. Social Science Quarterly, 85(4), 913-937.

Becerra, O., Cavallo, E., & Noy, I. (2015). Where is the Money? Post-disaster foreign aid flows. Environment and Development Economics, 20(5), 561-586.

Blackburn, S. (2014). The politics of scale and disaster risk governance: Barriers to decentralisation in Portland, Jamaica. Geoforum, 52, 101-112.

Burkart, C., Besiou, M., & Wakolbinger, T. (2016). The funding—Humanitarian supply chain interface. Surveys in Operations Research and Management Science, 21(2), 31-45.

Choudhary, C., & Neeli, S. R. (2018). Risk Governance for Improving Urban Disaster Risk Reduction. In Disaster Risk Governance in India and Cross Cutting Issues (pp. 211-224). Springer, Singapore.

Crunch, C. (2018). Economic losses, poverty & disasters: 1998-2017. *US Agency for International Development*.

Deen, S. (2015). Pakistan 2010 floods. Policy gaps in disaster preparedness and response. International journal of disaster risk reduction, 12, 341-349.

Dubey, R., & Gunasekaran, A. (2016). The sustainable humanitarian supply chain design: agility, adaptability and alignment. International Journal of Logistics Research and Applications, 19(1), 62-82.

Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, M. (2017). An empirical validation of a unified model of electronic government adoption (UMEGA). Government Information Quarterly, 34(2), 211-230.

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2017). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. Information Systems Frontiers, 1-16.
Dwivedi, Y. K., Shareef, M. A., Mukerji, B., Rana, N. P., & Kapoor, K. K. (2018). Involvement in emergency supply chain for disaster management: a cognitive dissonance perspective. International Journal of Production Research, 56(21), 6758-6773.
Eckstein, K., Jugert, P., Noack, P., Born, M., & Sener, T. (2015). Comparing correlates of civic engagement between immigrant and majority youth in Belgium, Germany, and Turkey. Research in Human Development, 12(1-2), 44-62.
Ertas, N. (2015). Political voice and civic attentiveness of public and non-profit employees. The American Review of Public Administration, 45(5), 607-626.
Friess, D. A., & Thompson, B. S. (2016). Mangrove Payments for Ecosystem Services (PES): A Viable Funding Mechanism for Disaster Risk Reduction?. In Ecosystem-Based Disaster Risk Reduction and Adaptation in Practice (pp. 75-98). Springer, Cham.
Garrity, D. M. (2015). Mobile Financial Services in Disaster Relief: Modeling Sustainability. In Technologies for Development (pp. 45-54). Springer, Cham.
Gundelach, B. (2016). Economic development and civic engagement in latin America: A comparative study. Nonprofit and Voluntary Sector Quarterly, 45(2), 238-260.
Habib, M. S., Lee, Y. H., & Memon, M. S. (2016). Mathematical models in humanitarian supply chain management: A systematic literature review. Mathematical Problems in Engineering, 2016.
Haselip, J., Hansen, U. E., Puig, D., Trærup, S., & Dhar, S. (2015). Governance, enabling frameworks and policies for the transfer and diffusion of low carbon and climate adaptation technologies in developing countries. Climatic Change, 131(3), 363-370.
Hilhorst, D. (2003). Responding to disasters: diversity of bureaucrats, technocrats and local people. International Journal of Mass Emergencies and Disasters, 21(1), 37-55.
Hoffmann, C. P., Suphan, A., & Meckel, M. (2016). The impact of use motives on politicians' social media adoption. Journal of Information Technology & Politics, 13(3), 239-256.
Jacobs, B. W., & Singhal, V. R. (2017). The effect of the Rana Plaza disaster on shareholder wealth of retailers: Implications for sourcing strategies and supply chain governance. Journal of Operations Management, 49, 52-66.
Jain Gupta, P., & Suri, P. (2017). Measuring public value of e-governance projects in India: citizens' perspective. Transforming Government: People, Process and Policy, 11(2), 236-261.
Jones, S., Manyena, B., & Walsh, S. (2015). Disaster risk governance: evolution and influences. In Hazards, risks and disasters in society (pp. 45-61).
Kabra, G., & Ramesh, A. (2015). Analyzing drivers and barriers of coordination in humanitarian supply chain management under fuzzy environment. Benchmarking: An International Journal, 22(4), 559-587.
Kabra, G., Ramesh, A., Akhtar, P., & Dash, M. K. (2017). Understanding behavioural intention to use information technology: Insights from humanitarian practitioners. Telematics and Informatics, 34(7), 1250-1261.
Khalilzadeh, J., Ozturk, A. B., & Bilgihan, A. (2017). Security-related factors in extended UTAUT model for NFC based mobile payment in the restaurant industry. Computers in Human Behavior, 70, 460-474.
Khan, A., & Krishnan, S. (2017). Social Media Enabled E-Participation: Review and Agenda for Future Research. e-Service Journal, 10(2), 45-75.
Kim, Y. I., Jang, S. J., & Johnson, B. R. (2016). Tying Knots With Communities: Youth Involvement in Scouting and Civic Engagement in Adulthood. Nonprofit and Voluntary Sector Quarterly, 45(6), 1113-1129.

Kirbiš, A., Flere, S., Friš, D., Krajnc, M. T., & Cupar, T. (2017). Predictors of Conventional, Protest, and Civic Participation among Slovenian Youth: A Test of the Civic Voluntarism Model. International Journal of Sociology, 47(3), 182-207.
Kunz, N., & Gold, S. (2017). Sustainable humanitarian supply chain management–exploring new theory. International Journal of Logistics Research and Applications, 20(2), 85-104.
Lai, C. H. (2017). A study of emergent organizing and technological affordances after a natural disaster. Online Information Review, 41(4), 507-523.
Linnerooth-Bayer, J., & Hochrainer-Stigler, S. (2015). Financial instruments for disaster risk management and climate change adaptation. Climatic Change, 133(1), 85-100.
Mansoori, K. A. A., Sarabdeen, J., & Tchantchane, A. L. (2018). Investigating Emirati citizens' adoption of e-government services in Abu Dhabi using modified UTAUT model. Information Technology & People, 31(2), 455-481.
Meth, S., Lee, K. Y., & Yang, S. B. (2015). Factors Influencing Facebook Users' Political Participation: Investigating the Cambodian Case. In PACIS (p. 44).
Munene, M. B., Swartling, Å. G., & Thomalla, F. (2018). Adaptive governance as a catalyst for transforming the relationship between development and disaster risk through the Sendai Framework?. International journal of disaster risk reduction, 28, 653-663.
Naranjo-Zolotov, M., Oliveira, T., & Casteleyn, S. (2018). Citizens' intention to use and recommend e-participation: Drawing upon UTAUT and citizen empowerment. Information Technology & People.
Okada, A., Ishida, Y., & Yamauchi, N. (2017). Effectiveness of social media in disaster fundraising: Mobilizing the public towards voluntary actions. International Journal of Public Administration in the Digital Age (IJPADA), 4(1), 49-68.
Oni, A. A., Oni, S., Mbarika, V., & Ayo, C. K. (2017). Empirical study of user acceptance of online political participation: Integrating Civic Voluntarism Model and Theory of Reasoned Action. Government Information Quarterly, 34(2), 317-328.
Orong, M. Y., & Hernandez, A. A. (2019). User Acceptance of Emergency and Disaster Response Mobile Application in the Philippines: An Investigation Based on the Unified Theory of Acceptance and Use of Technology Model. International Journal of Enterprise Information Systems (IJEIS), 15(1), 85-99.
Paik, S. (2018). Impact of Natural Disasters on Schools in India: Current Scenario in Terms of Policies and Practices. In Disaster Risk Governance in India and Cross Cutting Issues(pp. 295-312). Springer, Singapore.
Pakdeeburee, P., Denpaiboon, C., & Kanegae, H. (2011). Economic Valuation of the World Cultural Heritage for Promoting Community-based Flood Disaster Management: A Case Study of Ayutthaya Historical Park. Disaster Mitigation of Cultural Heritage and Historic Cities, 5, 247-254.
Paton, D., & Johnston, D. (2017). Disaster resilience: an integrated approach. Charles C Thomas Publisher.
Pattie, C., Seyd, P., & Whiteley, P. (2003). Citizenship and civic engagement: Attitudes and behaviour in Britain. Political Studies, 51(3), 443-468.
Poblet, M. (2010). Mobile governance: Empowering citizens to enhance democratic processes. In Proceedings of I-KNOW'10: 10th international conference on knowledge management and knowledge technologies (No. 10a).
Rodrigues, G., Sarabdeen, J., & Balasubramanian, S. (2016). Factors that influence consumer adoption of e-government services in the UAE: A UTAUT model perspective. Journal of Internet Commerce, 15(1), 18-39.
Ryzhov, I. O., Han, B., & Bradić, J. (2015). Cultivating disaster donors using data analytics. Management Science, 62(3), 849-866.

Scharf, K., Smith, S., & Wilhelm, M. (2017). Do disaster appeals reduce other donations? Evidence from the UK. Working paper, University of Birmingham.
Serrat, R., Petriwskyj, A., Villar, F., & Warburton, J. (2016). Learning through political participation: A case study of Spanish elders involved in political organizations. Adult Education Quarterly, 66(2), 169-187.
Shaw, J. D., & Gupta, N. (2015). Let the evidence speak again! Financial incentives are more effective than we thought. Human Resource Management Journal, 25(3), 281-293.
Shi, P., Li, N., Ye, Q., Dong, W., Han, G., & Fang, W. (2010). Research on integrated disaster risk governance in the context of global environmental change. International Journal of Disaster Risk Science, 1(1), 17-23.
Shi, P., Shaw, R., Ardalan, A., Chan, E. Y. Y., Choudhury, J. R., Cui, P., ... & Kasuga, F. (2018). Fourteen actions and six proposals for science and technology-based disaster risk reduction in Asia. International Journal of Disaster Risk Science, 9(2), 275-279.
Singh, A. K., & Sahu, R. (2008). Integrating Internet, telephones, and call centers for delivering better quality e-governance to all citizens. Government Information Quarterly, 25(3), 477-490.
Strömblad, P., & Bengtsson, B. (2017). Collective Political Action as Civic Voluntarism: Analysing Ethnic Associations as Political Participants by Translating Individual-level Theory to the Organizational Level. Journal of Civil Society, 13(2), 111-129.
Suzuki, M. (2015). Identifying roles of international institutions in clean energy technology innovation and diffusion in the developing countries: matching barriers with roles of the institutions. Journal of Cleaner Production, 98, 229-240.
Uthaman, V. S., & Ramankutty, V. (2017, March). e-Governance Service Quality of Common Service Centers: a Review of Existing Models and Key Dimensions. In Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance (pp. 540-541). ACM.
Uzoka, F. M. E. (2008). Organisational influences on e-commerce adoption in a developing country context using UTAUT. International Journal of Business Information Systems, 3(3), 300-316.
Van Niekerk, D. (2015). Disaster risk governance in Africa: A retrospective assessment of progress against the Hyogo Framework for Action (2000-2012). Disaster Prevention and Management, 24(3), 397-416.
Vatanasakdakul, S., Aoun, C., & Chen, Y. (2017). Chasing Success: An Empirical Model for IT Governance Frameworks Adoption in Australia. Science, Technology and Society, 22(2), 182-211.
Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.
Weng, W. W., Woo, C. K., Cheng, Y. S., Ho, T., & Horowitz, I. (2015). Public trust and corruption perception: disaster relief. Applied Economics, 47(46), 4967-4981.
Wenner, G., Bram, J. T., Marino, M., Obeysekare, E., & Mehta, K. (2018). Organizational models of mobile payment systems in low-resource environments. Information Technology for Development, 24(4), 681-705.
Wolfgang Kron. (2018, April 1). Asian Floods overshadowed by Houston Flooding. Munich RE. Retrieved from https://www. munichre.com/topics-online/en/2018/01/flood-disasters
Xavier, N., & Oliveira, T. (2016, September). Factors Affecting Behavioural Intention to Adopt e-participation: Extending the UTAUT 2 Model. In The European Conference on Information Systems Management (p. 322). Academic Conferences International Limited.
Zagefka, H., & James, T. (2015). The psychology of charitable donations to disaster victims and beyond. Social Issues and Policy Review, 9(1), 155-192.

Zhao, M., Wang, F., Zhao, D., & Wei, J. (2015). Using CSR Theory to Examine Disaster Aid Response to the Wenchuan Earthquake. Global Business and Organizational Excellence, 34(4), 50-59.

Zhou, T., Lu, Y., & Wang, B. (2010). Integrating TTF and UTAUT to explain mobile banking user adoption. Computers in human behavior, 26(4), 760-767.

Zolotov, M. N., Oliveira, T., & Casteleyn, S. (2018). E-participation adoption models research in the last 17 years: A weight and meta-analytical review. Computers in Human Behavior, 81, 350-365.

https://www.unisdr.org/we/inform/publications/61119

https://www.emdat.be/database

https://ndma.gov.in/en/

Decision Sciences Institute
Novel MINITAB macros for advanced statistical and plant genetic analyzes:
A case study with maize

Ali Arminian
Ilam University
Email: *a.arminian@ilam.ac.ir*

Karimi Zaman
Ilam University
Email: NA

Paulo Canas Rodrigues
University of Tampere, Finland, and the Federal University of Bahia, Brazil
Email: paulocanas@gmail.com

Ceyhun Ozgur
Valparaiso University
Email: Ceyhun.ozgur@valpo.edu

Abstract

In this paper, development of advanced and novel macros for MINITAB environment, described allowing users to perform certain statistical and experimental design analyses: (i) a curve-fitting macro examining 14 fundamental algorithms on a response-predictor dataset to extract the best relationship(s), (ii) a macro for comparing two experimental designs using relative efficiency, (iii) a macro analyzing Latin square designs with sub-samplings,(iv) a macro for orthogonal contrasts of treatments/groups, (v) a macro performing multivariate path coefficient analysis (MULTPATH), (vi and vii) macros for analyzing stability analysis (univariate and multivariate i.e. AMMI[1]) of genotypes in plant breeding . In a field experiment, 20 of recently-bred maize (*Zea mays* L.) genotypes were examined under 2 drought-stress environments along with the control, and grain yield (YLD), seeds per ear (SPE), and ear length (EL) were scored to identify the superior hybrids. Results showed a non-significant difference between, genotypes, blocks and stress levels for ear length trait as a multiple-observation Latin square design. Moreover, the relative efficiencies (RE) showed that LS design has a lower RE than its lower-level designs. Orthogonal contrast comparison indicated a significant difference between two groups of maize genotypes (9 genotypes vs 11 genotypes). According to the results of curve-fitting, our macro could unravel a specific kind of the exponential curves with the highest R-sq(adj) index (91.54%). Path coefficients analysis revealed that independent variables EL and SPE had no significant direct effects on the dependent variable YLD. The univariate stability analysis introduced genotype BARAKAT-74, GARZA, and DKC-6101, BOLSON were the most stable genotypes in drought stress conditions. Finally, the AMMI stability showed that AMMI2 model was the best model and it introduced the stable, high-yielding and moderate-tolerant maize genotypes in this experiment. The macros presented here could be utilized in various areas of biological sciences, especially plant breeding, to perform advanced statistical and experimental design analyses.

Keywords: AMMI model, curve fit, experimental designs, global and local macro, non-linear, path analysis.

Introduction

A crucial task in statistical analyses is finding the best fitting model for a dataset, which may include a dependent/response and one or more causative factor(s) or casual variable(s). In regression analyses, typically one of the variables, often called response or dependent variable, is of particular interest and may be denoted as y . The other variables, i.e., x_1, x_2, x_3, …., x_k, usually called explanatory, regressor, or independent variables, are primarily utilized to predict or explain the behavior of dependent variable *y* (Seber and Wild, 2003). Regression analysis reveals the changes in response variable with the value of a vector consisting of predictors (independents), where multiple linear regression and most of the experimental design models are special forms of a linear regression model (Olive, 2017). Sometimes, the appropriate mathematical form of the relationship of two variables (e.g., response and independent) is known (e.g., linear regression), except for some unknown constants or coefficients (called parameters). In most of the statistical analyses, particularly in biological sciences, in contrast to physical sciences, the underlying processes are generally complicated and not well understood. This means that we have little or no idea about the shape of the association, and the object is to simply find the best-fitting function. Accordingly, finding the best model fitting a dataset is of interest and is mainly the target of studies. Many statistical approaches are being used in various fields of studies, e.g., biological and agricultural trials. Data generated are subjected to statistical analyses to extract the best/real relationships among the variables; the relationship may be simple, or multiple, linear regression line; or it could be more complex, e.g., quadratic, cubic, logarithmic, natural logarithmic, inverse or other variants. One may wish to compare efficiencies of experimental designs via relative efficiency (RE); or compute path coefficients involving independent and dependent variables; or make orthogonal contrasts or group comparisons following analysis of variance. For some of these procedures, computer programs are being proposed, which can be implemented in MINITAB, which is simple and user-friendly statistical software, with a programming language interface. MINITAB (MINITAB, 2018) is powerful statistical (Karafilowska, 2004) software that has a wide range of statistical procedures and graphical capabilities. In MINITAB, macros for many statistical procedures, experimental designs and algorithms can be handled (Arminian *et al*., 2008). MINITAB macros are easily edited, customized, and maintained, and can be made more powerful with a higher-level MINITAB macro programming environment (Mathews, 2005).

Drought is detrimental environmental stress in agriculture, and therefore the goal of conventional plant breeding programs. A review of recent advances in drought tolerance suggests that, to a certain extent, selection for high seed yield under non-stress conditions directly improves yield under water constraints (Sserumaga et al., 2018). Maize (*Zea mays* L.), after wheat and rice, is the third most important crop in the world, has the highest cropping area in some parts of the world, such as Nigeria, which is due to its high productivity, wide adaptability and particularly simple cultivation, processing, storage and exchange (Oyekunle *et al.,* 2017).

In this paper, we present and describe some advanced macros for MINITAB, saved in the form of "*.mac" and apply them in the study the grain yield and some other traits of maize plant under drought stress. These macros are designated as 1) curve-fitting macro, 2) macros for comparing the relative efficiency (RE) of two experimental designs, followed by estimation of the number of replication required for the second design to be as efficient as the first (targeted) design, 3) a macro for performing multivariate path coefficient analysis (MULTPATH) and its components, 4) macros for computing stability analysis univariate including parametric and nonparametric (under work), and more important, multivariate methods of stability analysis especially AMMI (additive main effects and multiplicative interaction), in plant breeding. The statistical procedures being covered in the proposed MINITAB software are described as follows.

Materials and methods

This experiment was conducted in which, 20 of recently-bred maize hybrids (some as dual-purpose: grain and forage) were examined in a field experiment under 3 irrigation management regimes, i.e. control, drought-treatments at kernel milk stage (R3) and kernel dough stage (R4) according to Ransom (2013) and Nielsen (2018). The experiment was performed at the Institute of the agricultural research center of Ilam province, Iran, in the city of Mehran (32:33º, 38:53´, 22:58"N; 46º, 2:23´, 9:19"E) in the 2017-2018 cropping year. The design was a factorial experiment design with two factors*:* drought and genotypes in 3 blocks and 3 samplings per each block. Before planting, the soil was examined for texture, drainage, and percentage of mineral elements. It has a good-drainage. Organic humus, like compost, and livestock manure were added to the soil in fall to improve drainage and a nutrient-rich soil formation. Maize needs a high level of nitrogen for better growth, and therefore excess fertilizer is needed to be added during planting and during the growing season. Grain yield (tons ha^{-1}), ear length (EL: cm), and a number of kernels per rows of the ear (SPE) were measured for each sampling per block. The harvesting time occurred when most all of the maize planted has reached “black layer” or in other terms, physiological maturity. This is accompanied by developing a black layer at the tip of the kernels, where they attach to the cob.

MINITAB is one of the world's powerful, user-friendly, reliable and statistical quality control software that performs numerous statistical analyzes. It also has a fascinated environment for writing and executing programming functions as exec, global and local functions called macros. Unlike environments like R, the MINITAB macros are more accurate and simpler. To execute commands, you need to enable the command line prompt (MTB >), click anywhere within the session window to activate it, then follow the path: 'Edit> Enable macros' (or: Editor> show command line) and type: "MTB > %mymacro in front of the command prompt, followed by clicking the ENTER key. Then the macro is executed and the results are achieved in the session window. The “%” sign is needed in front of the commands to perform a macro (global or local) which is written by users. MINITAB programming capabilities enable the users to provide their powerful macros according to the fields of study. MINITAB has three macro types. The first one is the most simplified preliminary programming command called "exec" or executive. The 2nd and more complex is called "global macros," and the 3rd and the most advanced is called "local macros." MINITAB accepts data in simpler formats than other statistical software. There are particular ways to call/recall a dataset into MINITAB. The first and the simple procedure is entering (typing) the data directly into worksheet cells. The second way is to open and recall a dataset file that has previously been saved using MS Excel, Notepad or other environments, e.g., "*.txt" or "*.dat" file extensions. In the following, we describe the statistical methodologies and the design of the experiments discussed in this paper. In the results and discussion, we explain the results of each macro step-by-step.

1- Analysis of designed experiments with MINITAB

1-1- Latin square (LS) design with subsampling

Latin squares (LS) are types of designs sometimes called double grouping designs, in which the treatments are grouped into replicates in two different ways (Cochran and Cox, 1957). They are, as usual, the simplest form of row-column designs. They are used to compare *t* treatments in *t* rows and *t* columns, where rows and columns represent the two types of blocking factors (Hinkelmann and Kempthorne, 2008). Of course, “row” and “column” terms are optional and majorly used in agriculture sciences, and first to third factors in other branches like physics and chemistry. Latin squares have origins in agricultural experiments and have applications in the social and behavioral sciences (Richardson, 2018a&b). The statistical equation (model) of a Latin square design with sub-sampling is as follow:

$y_{ij(k)l} = \mu + R_i + C_j + T_{(k)} + e_{ij(k)} + \varepsilon_{ij(k)l}$ (or: $y_{ijl} = \mu + \rho_i + \gamma_j + \tau_k + e_{ij} + \delta_{ijl}$

$$e_{ij} \sim^{iid} N(0, \sigma_e^2)$$
$$\delta_{ij\ell} \sim^{iid} N(0, \sigma_d^2)$$
$$i = 1, \ldots, t,\ j = 1, \ldots, t,\ k = 1, \ldots, t,\ \ell = 1, \ldots, n$$

where $y_{ij(k)l}$ denotes the observation of *i*th row, *j*th column, *k*th treatment and *l*th sample; μ = grand mean, R_i = row effect, C_j = column effect, $T_{(k)}$ = treatment effect, $e_{ij(k)}$ = main error and $\varepsilon_{ij(k)l}$ = sampling error [i = 1, . . . , t; j = 1, . . . , t; k = 1, . . . , t; and ℓ = 1, . . . , n.] In addition, the number of degrees of freedom for the sub-sampling error = $t^2(n-1)$. Treatments, columns, and rows may be necessarily fixed or random. The computerized analysis of the Latin square design with sub-sampling is tedious and has been impossible until now, justifying our main reason for introducing a MINITABMINITAB macro for this kind of design. We provide here a method for analyzing this type of design in MINITAB. In this paper, we analyze the maize ear length trait (EL) as an LS design with subsampling to show the method of executing the relevant macro we wrote.

1-2- Relative efficiency (R.E.)

Latin square design with sub-sampling and rectangular, repetitive, split-factorial and factorial-split, split-split-plot and split-block designs, are generally used in agricultural studies performing via MINITAB. In this context, we wrote some complementary macros for DOE (designs of experiments) for estimating and comparing the relative efficiency (RE) of two experimental designs. In designed experiments, the RE of a given design relative to another design is specified as the number of replicates of the other design needed to gain the equal result as one replicate of the first (Oladugba *et al*. 2013). Suppose we have performed an RCBD (randomized complete block design) in the field, and we analyzed it as a CRD (completely randomized design), and we are interested to know which design would give higher efficiency. In this case, it can be concluded that computing relative efficiency is a tool by which determining the magnitude of the reduction in experimental error due to blocking is achieved as:

$$R.E. = \frac{(r-1)E_b + r(t-1)E_e}{(rt-1)E_e}$$

Where *Eb* and *Ee* stand for block and error mean squares, and also *t* and *r* stand for the number of treatments and replications (or blocks), respectively, in the RCB design (Gomez and Gomez, 1984). If the error *d.f.* is less than 20, the *R.E.* value should be multiplied by the adjustment factor *k* defined as:

$$k = \frac{[(r-1)(t-1)+1][t(r-1)+3]}{[(r-1)(t-1)+3][t(r-1)+1]}$$

Moreover, blocking is essentially affecting the difference in the magnitude of experimental error between a CRD and an RCBD; hence the value of the *R.E.* is indicative of the gain in precision due to blocking (Gomez and Gomez, 1984). It is assumed that the CRD is the simplest design among all existing designs. When only two factors are examined, then the CRD is equivalent to a *t*-test. If the RCBD design has RE greater than 100, then analyzing it as a CRD would need to have more replications to achieve the accuracy of an RCBD. Our written macros also compute the number of replications needed for the ‘lower’ (less efficient) design to be as accurate as the ‘upper’ (more efficient) design. These macros compare the relative efficiency of RCBD to that of CRD, and also of Latin square design (LS) to RCBD and CRD. We designated our macros as *RB2CD*, *LS2CD*, *LS2RBc,* and *LS2RBr*. The *RB2CD* macro compares RB design to CD, *LS2CD* macro compares Latin square design to CD, *LS2RBc* macro compares Latin square to RB when columns are supposed to be blocks, and *LS2RBr* macro compares Latin square to RB when rows are supposed to be blocks. In this paper, we compare the RE of the analyzing LS design with its equivalent RB and CD designs to show the method of executing the macros written.

1-3- Orthogonal contrasts (group comparisons) between factors (e.g., treatments in experiments)

Suppose, the null hypothesis is: $H_0 = \mu_1 = \mu_2 = \mu_3 = ... = \mu_k$. One may test it by means of one-way ANOVA. However, if one is interested in determining how precise differences existing between the groups, then rejecting the omnibus test requires not performing mean comparisons/contrasts. However, all the information required to compare the factors in the trial using simple pair-wise comparisons is not possible. In such a case, it is necessary to do single-degree-of-freedom, orthogonal contrasts. To do this, the ANOVA table information is needed, especially replication number and error mean square. The orthogonal contrast technique is mentioned as a simple and efficient approach to analyzing the experimental data in order to obtain, e.g., the main, interaction and nested effects, for group mean comparisons and/or specific residuals. In addition, it could be an alternative means of analyzing experimental data without a definite structure, such as experiments with additional treatments (Cristina and Nogueira, 2004).

Thus, the estimation of the significance of the comparisons is a valuable task, especially in agricultural trials, and preparing a good program that is simple and comprehensive in performing this kind of analysis is needed. In this context, we wrote a macro called "*contrast*, which simply performs any group comparisons and gives the user a P-value, whereby he/she is able to find that a specific group comparison is or is not significant. The most attractive part of this macro is where it estimates the orthogonal coefficients for the group comparisons automatically, which is a hard task. In many other programs, the user has to insert them manually into the program, which is time-consuming and tedious. Here we perform a contrast test with "*contrast*" macro we wrote for its functionality and outputs, comparing a group of factors with other groups (s), as come in the results and discussion part of this paper. This macro is the best program for doing orthogonal contrasts because it does not require any coefficient definition and insertion by the user.

2- Studying variables relationships using MINITAB

2-1- Curve fitting algorithms and MINITAB macro results

Considering a single independent (x) and a given response (y) variable, a simplified linear regression equation may be fitted as: $y_i = \beta_0 + \beta_1 x + \varepsilon_i$, where the model $y \approx \beta_0 + \beta_1 e^{\beta_2 x}$ is a nonlinear model, being nonlinear in β_2, where β_0 refers to intercept, β_1 and β_2 to slopes of the equation and e is natural logarithm (Seber and Wild, 2003). Even when a linear approximation exists and works fine, a nonlinear model may still be applied for clearly interpreting the parameters (Seber and Wild, 2003). There are numerous nonlinear models in statistics, and the above-mentioned one is a type that may be used instead. The algorithms discussed in this paper and supplemented in our "*curvfit*" macro, are applicable to all fields of science. Users may, however, for obtaining information about statistical curve fitting algorithms, refer to nonlinear materials in Seber and Wild (2003). Our macro advantage is that it seeks the most suitable curve fit algorithm between 15 algorithms implemented in the macro.

As Sokal and Rohlf (1995) denoted, linear regression is not always enough to assess the differences among some sample means. So instead of a linear trend, some relations in nature might best be expressed via fitting a curvilinear regression. It is said that the linear regression approach is not sufficient for most biological situations and majority of nonlinear approaches applied to biological or agricultural situations can be categorized into three types: a) exponential growth curves, b) negative exponential and c) Michaelis–Menten (1913) model which has been widely used for enzymatic and chemical kinetic reactions (Lawal, 2014), and also growth models (like logistic growth model and Gompertz growth model), which have main roles in biology. Simplifying the complicated curve fits or nonlinear forms is a good way to better investigate such trends. According to Johnson and Bhattacharyya (2010), such transformations are used to transform the response to approximate a linear relation. So, as they mention, the data could be transformed to a nearly straight-line relationship, or one can try to linearize the relation of two variables by plotting a transformed y response variable or some other transformed y with the independent x variable. They count some types of transformations or nonlinear models and their corresponding linearized forms including power and root forms (mentioned above in our macro), and concluded that sometimes a specific nonlinear association is mainly suggested and that while initial information regarding the form is not in hand, a study of the scatter diagram may often show the proper linearizing transformation (Johnson and Bhattacharyya, 2010). Data transformation in regression has been suggested by Zar (2014), as he mentions that transforming the independent variables has no effect on ”Y” distribution but transforming the “X” may generally linearize data.

2-2- Path coefficient analysis

Path coefficients analysis (Fig. 1), an extension of multiple regression analysis and a powerful tool for studying causal associations among a set of normal variables, has been detailed in the literature (Wright, 1934; Li, 1975; Dewey and Lu, 1959; Kang, 2015). A path coefficient analysis entails exogenous, endogenous, intervening or intermediate variables and an ultimate dependent variable. In the path model, a path coefficient resembles the direct effect of a variable (causal variable) (Pedhazur, 1997; Wuensch, 2016). In addition, for each path toward an endogenous variable, a path coefficient, i.e. “*pij”,* refers to the effect of an independent variable *i* on *j*th variable. Next, the square of a path coefficient gives the proportion of the affected variable's variance (Wuensch, 2016). The interrelationships between such variables are indicated using the one-headed arrow (direct path coefficient); two-headed arrows represent correlation coefficients between pairs of variables (for details see Arminian *et al*., 2008). Path coefficient analysis has been used in various fields of study, e.g., biological sciences, social studies, and agricultural fields, such as plant breeding. Plant breeders have, indeed, used path analyses to identify useful traits as selection criteria to improve crop yield (for example, see Sidwell *et al*., 1976; Kang *et al*., 1983; Board *et al*., 1997, 2003; Kang, 2015).

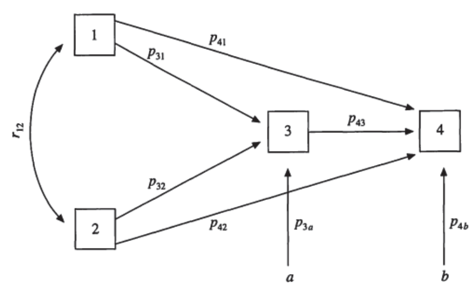


Fig. 1. A given path diagram for the relationship between a dependent (4), 2 exogenous (1 and 2) and 2 endogenous (3 and 4) variables, (courtesy of Pedhazur (1997). Path coefficient analysis is described by Wright (1934) and later by Li (1975). Single-headed arrows represent direct effects (Pij), whereas double-headed arrows represent correlation coefficients (rij) in the model.

The complete guide regarding path coefficient analysis is given above as reported by researchers (Pedhazur, 1997; Wuensch, 2016). The benefit of path coefficient analysis is that it allows decomposition of a correlation coefficient into its components, i.e., direct (the path coefficient, involving the direct effect of predictor variable upon its response variable) and, indirect effect (of a predictor variable on the response through another predictor) (Dewey and Lu, 1959). Carey (1998) in an introduction to multiple regression and path coefficient analysis with a new macro for SAS called PROC CALIS, mentions that the main advantage of path analysis is when there are two or more dependent variables are present which is so called as “multivariate multiple regression”. In addition, Carey (1998) indicated that path analysis can also be used to decompose the variance of a single dependent variable. The PATHSAS macro by Cramer *et al*. (1999) is another SAS macro that performs path analysis. Kang (2015) brought in details in a computer program called "Kang_SASPath" and gave an example of performing path coefficient analyses. This program also determines index weights for developing selection indices for plant breeding.

In agriculture, path analysis has been applied by plant breeders to help identify beneficial characters as selection criteria to improve crop yield (Dewey and Lu, 1959; Milligan *et al*., 1990). MINITAB "*MULTPATH.MAC*" macro gives descriptive statistics, path coefficient analysis (direct and indirect effects of the causative parameters influencing intervening variables or a final dependent one). The macro also tests the significance of each direct effect divided by its related standard error. The complete version of this macro is available in Arminian *et al.* (2008). We suggest applying and using this MINITAB macro and other macros introduced in this paper on earlier versions of MINITAB, like 13, but not 18. To run the macro, the procedure is as follows: suppose the executive file is available in the path *'D:\MINITAB14\Macros\multpath.mac'* and one dependent and two independent variables are in this order (dependents, independents: left to right, respectively) within the worksheet columns C1 to C3.

MTB>%multpath c1-c3.

```
*   Elements of m.3 matrix for 4 independents:
*
*      P1Y   r12P2Y   r13P3Y   r14P4Y  ..
*    r12P1Y     P2Y   r23P3Y   r24P4Y  ..
*    r13P1Y   r23P2Y     P3Y   r34P4Y  ..
*    r14P1Y   r24P2Y   r34P3Y     P4Y  ..
*    ......   ......   ......   ......  ..
*    Path equations specify the path model, through m.3
*    i.e e.g.: r1Y = P1Y + r12P2Y+ r13P3Y+ r14P4Y + ...
```

The above lines represent the framework of the basic matrix of path coefficient analysis. In our macro outputs (results and discussion part of this paper), the first rows of the results represent original data, followed by descriptive statistics, e.g., the number of observations, mean, SE mean, standard deviation, minimum, first and third quartile, median, min, and max. Correlation coefficients are accompanied by their respective significance levels. In the next part, macro performs linear multiple regression analysis for dependent(s) versus independent variable(s), respectively [when more than one dependent is present, the macro performs this analysis for all dependent-independents variable pairs. For this, regression equation, along with variance inflation factor (VIF), statistical tests, such as R-squared, and importantly analysis of variance table of regression are constructed Arminian *et al*., 2008). The term VIF or simply variance inflation factor distinguishes multicollinearity in regression analysis. When there is a correlation between predictors (i.e., independent variables) in a given model, this is known as multicollinearity, which can adversely affect the regression results. The VIF parameter estimates the inflation rate of the variance of a regression coefficient due to multicollinearity in the model. Mathematically, $VIF=1/(1-R^2i)$. Finally, Durbin-Watson statistics of the regressions are estimated. Most regression problems including time series data show positive autocorrelation. In this case, the Durbin-Watson statistic is a test that shows that the residuals from a linear regression or multiple regression are independent (Montgomery *et al.,* 2012). The lower values of this statistic are desirable.

3- Stability analysis in plant breeding using MINITAB

3-1- univariate parametric stability analysis of genotypes

We have written and evaluated two MINITAB advanced macros called "*MINISTAB.MTB*" for analysis of 20 parametric stability statistics, "*NONMINISTAB*" for nonparametric analyses, and "*AMMI.MAC*" for additive main and multiplicative interaction (AMMI) model, respectively, for use in plant breeding. The MINISTAB computes most sophisticated parametric stability statistics, such as Finlay-Wilkinson (1963), Eberhart and Russell (1966), and much more (Roemer, 1917; Plaisted and Peterson, 1959; Wricke, 1962, 1964; Hanson, 1970; Hussein *et al.*, 1970; Shukla, 1972; Freeman and Perkins in Freeman, 1973; Pinthus, 1973; Francis and Kannenberg, 1978; Lin *et al.*, 1986; Becker and Léon, 1988; Lin and Binns, 1988; Kang, 1998). Such univariate stability analysis is beneficial in plant breeding so that recently an R package (https://cran.r-project.org/web/packages/stability/stability.pdf) is being developed called "stability" to identify superior and stable genotypes under diverse environments. This package extracts and performs Eberhart and Russell ANOVA (1966), Fin-lay and Wilkinson (1963) Joint Linear Regression, Wricke (1962, 1964, 1965) Ecovalence, Shukla's stability variance parameter (1972) and Kang's (1993) simultaneous selection for high yielding and stable parameter. Of course, although this and other G×E analysis programs are being used extensively nowadays, they extract some of the stability parameters and not all of them. Our MINITAB macros for univariate stability part, extract 20 stability parameters which will be discussed later. Our target was to estimate all of these parameters because we see that many researchers are still using univariate parameters, although multivariate methods for GE like GGEBiplot and AMMI, may be the best choices or alternatives. For stability analyses, current maize data set (grain yield or YLD as tons ha^{-1}) was used. The dataset involved 20 genotypes and 3 environments (drought stress levels) with 3 replications.

3-2- multivariate parametric stability analysis of genotypes

Additive main effect and multiplicative interaction (AMMI) is a useful approach to perform stability analysis in plant breeding, which is done here using a macro called AMMI for MINITAB. Of course, some other AMMI programs are included in known softwares like GENSTAT or written like that developed by Rodrigues *et al*. (2015). The AMMI macro we provided for MINITAB environment is nearly a complete one, which provides almost all information a user needs for AMMI approach. Among the different approaches to investigate multivariate stability analysis, AMMI is a good choice and according to Gauch (2006), AMMI partitions the overall variation into main effects of genotype and environment, and also genotype by environment interactions. With our *AMMI* macro, 3 types of data could be computed: 1) genotype, environment, replication and trait(s), one by one, respectively, arranged left to right. 2) genotype, environment, trait(s), respectively, as above but without replication. 3) a two-way table of E×G (environment by genotype) interaction. When executing this macro, a window pops up prompting the user to enter some crucial info as follows:

```
…………..…
                AMMI.MAC
  Local macro performing Additive Main Effects and Multiplicative Interaction (AMMI)
analysis of an ANOVA data or an ExG table of means,
….
  Example calling statements:
  1: Two-way data serial (data in 3 columns: Gen, Env, and Response),
     e.g.,in C4-C6 columns. Then, at command prompt type, MTB> %AMMI C4-C6.
*****
  2: Data in 4 columns, e.g., Gen, Env, Rep & Response in C4-C7, as MTB>%AMMI C4-
C7
*****
  3: Two-way data matrix of E by G (E in rows and G in columns),
     E.g., if G1-G10 genotypes mean in C11-C20, MTB>%AMMI C11-C20.
*****
  Note: The order of data must be as above.
  To exit the macro at any stage, press CTL + Pause/Break buttons.
  Copyright(C) April 2018. A. Arminian. Ilam university, Iran.
  All rights reserved.     E-mail:   arminian_a@yahoo.com
…………………….
```

As shown, the program accepts 3 types of data as input: 1) Two-way data serial (data in 3 columns: Gen, Env, and Response, respectively, left to right order), 2: Data in 4 columns: Gen, Env, Rep & Response, e.g., in C4-C7, as MTB>%AMMI C4-C7, 3: Two-way data matrix of E by G (E in rows and G in columns), e.g., if G1-G10 genotype means are located in C11-C20 columns, the command must be executed as MTB>%AMMI C11-C20. The AMMI macro written here calculates more advanced functions than other AMMI programs, such as GENSTAT, including: ANOVA tables of stability and AMMI with Gollob (1968) and Cornelius (1993) *F* functions, descriptive statistics, Finlay-Wilkinson (Finlay and Wilkinson, 1963) regression output, nominal yield estimates, use of supplementary approach in stability analysis (Pacheco *et al.,* 2005), and also graphical representations of the stability parameters as biplots (Fig. 4) and surface plots as referred by Kang and Gauch (1996). A fictitious dataset consisting of genotypes (G), environments (E), and replication (courtesy of Gauch, 1992) is given in appendix no. 5.

Results and discussion

Latin square (LS) design with subsampling

For the maize data of this experiment, 3 of genotypes were examined in 3 columns (as drought stress) with 3 samples per each block. The map of the trial is as below (Table 1):

Table 1. map for the multi-observation (sub-sampling) Latin square design in this experiment

Control	Drought-1	Drought-2
BARAKAT 2 22, 22, 18	**BARAKAT-4 19, 21, 19**	**BOLSON 21, 18, 16**
BARAKAT 4 18, 20, 20	**BOLSON 19, 19, 20**	**BARAKAT- 16, 14, 19**
BOLSON 18, 20, 24	**BARAKAT-2 18, 15, 15**	**BARAKAT- 23, 19, 21**

According to table 2, five columns of data, represent the block, drought stress (column1), genotype (treatment) and sampling as factors and ear trait were enough to do a Latin square with sub-sampling. Figure 2 shows the method for entering data of map above into MINITAB.

Table 2. A part of the data set of maize for analyzing the Latin square design with subsampling

Genotype	ENV=Stress	Block	Sample	EL
BARAKAT-2	Control	1	1	22
BARAKAT-2	Control	1	2	22
BARAKAT-2	Control	1	3	18
BARAKAT-2	Control	2	1	19
BARAKAT-2	Control	2	2	16
BARAKAT-2	Control	2	3	20
BARAKAT-2	Control	3	1	21
BARAKAT-2	Control	3	2	21

The calculations are performed through MINITAB worksheet windows as STAT » ANOVA » General Linear Model » Model: LE, and Factors: Block, Stress, and Genotype. The results are:

Analysis of Variance

Source	DF	Seq SS	Adj MS	F	P
Genotype	2	26.741	13.370	2.63	0.097
Block	2	7.185	3.593	0.71	0.506
Stress	2	19.185	9.593	1.88	0.178
Error	20	***101.852***	5.093		

Total 26 154.963

For the analysis of the final variance, we must formulate another table with the sums (or averages) of the treatments over the samples and perform the analysis of variance to give the following results.

Analysis of Variance

Source	DF	Seq SS	Adj MS	F	P
Genotype	2	80.22	40.11	0.98	0.504
Block	2	21.56	10.78	0.26	0.791
Stress	2	57.56	28.78	0.71	0.586
Error	2	***81.56***	40.78		
Total	8	240.89			

The main target of such tables is to construct the sub-sampling error SS as: **101.852-(81.56/3=27.19)=74.66**.Then the final ANOVA table is:

Analysis of Variance:

Source	DF	Seq SS	Adj MS	F	P
Genotype	2	26.741	13.370	2.63	0.097
Block	2	7.185	3.593	0.71	0.506
Stress	2	19.185	9.593	1.88	0.178
Error	20	101.85	5.09		
&$\text{Error}_{\text{Exp}}$	2	***27.66***	13.59	3.28	0.061
$\$$$\text{Error}_{\text{samp}}$	18	74.66	4.15		
Total	26	154.963			

z&*Lack –of-Fit*
$*Pure Error*

For ease of use, we provided a local macro for lower versions of MINITAB, like 14, which can extract important parameters from the table of analysis of variance, such as the degree of freedom, sampling error, and MS. To execute this macro, as said, datasets must be located in 5 columns similar to Figure 2. To do this, activate the session window and in front of the command line prompt (MTB>) type: MTB > %latss c1-c5 (OR: Block-EL) followed by pressing the ENTER key. During the execution of the program, MINITAB asks the user 2 questions in which they should enter the number of treatments and the sum of squares of the two designs. The results are:

MTB > %latss c1-c5
Executing from file: D:\MINITAB14\MACROS\latss.MAC
Dataset:

Row	Genotype	Block	Stress	Sample	EL
1	BARAKAT-2	1	Control	1	22
2	BARAKAT-2	1	Control	2	22
……					
27	BARAKAT-74	3	Drought-2	3	21

ANOVA for original data (subsampled):

Factor	Type	Levels	Values
Genotype	fixed	3	BARAKAT-2, BARAKAT-74, BOLSON
Block	fixed	3	1, 2, 3
Stress	fixed	3	Control, Drought-1, Drought-2

<u>Analysis of Variance for EL, using Adjusted SS for Tests</u>

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Genotype	2	26.741	26.741	13.370	2.63	0.097
Block	2	7.185	7.185	3.593	0.71	0.506
Stress	2	19.185	19.185	9.593	1.88	0.178
Error	20	**101.852**	101.852	5.093		
Total	26	154.963				

S = 2.25668   R-Sq = 34.27%   R-Sq(adj) = 14.56%

* ANOVA for summated data over samplings:

Factor	Type	Levels	Values
Treat.	fixed	3	BARAKAT-2, BARAKAT-74, BOLSON
Block.	fixed	3	1, 2, 3
Column.	fixed	3	Control, Drought-1, Drought-2

<u>Analysis of Variance for Response, using Adjusted SS for Tests</u>

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Treat.	2	80.22	80.22	40.11	0.98	0.504
Block.	2	21.56	21.56	10.78	0.26	0.791
Column.	2	57.56	57.56	28.78	0.71	0.586
Error	2	**81.56**	81.56	40.78		
Total	8	240.89				

S = 6.38575   R-Sq = 66.14%   R-Sq(adj) = 0.00%

Enter number of TREATS, and SAMPLINGS, respectively:
DATA> 3 3

Enter ESS of the 1st and 2nd ANOVA tables, respectively:
DATA> 101.852 81.56

* --------------------------------
Lack-of-Fit SS: 27.1867

* --------------------------------
sub-sampling=Pure Error SS: 74.6653

* ---
Degree of freedom for sub-samp: 18.0000

* ----------------------------------
Sub-sampling Error MS: 4.14807

As you can be seen through the results above, the final ANOVA table for this experiment includes two types of errors called ErrorExp and Errorsamp. Interestingly when we analyzed the original data containing samples by MINITAB 18.1 and defining the ANOVA model as STAT » ANOVA » General Linear Model » Fit model: Response: LE. And the factors are Block, Stress, and Genotype. Then the constructed table is as follows:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Block	2	7.185	3.593	0.71	0.506
Stress	2	19.185	9.593	1.88	0.178
Genotype	2	26.741	13.370	2.63	0.097
Error	20	101.852	5.093		
Lack-of-Fit	2	27.185	13.593	3.28	0.061
Pure Error	18	74.667	4.148		
Total	26	154.963			

As could be seen from the results above, there is not significant difference between blocks, drought stress levels, and 3 genotypes. The current result is likely due to fewer genotypes used in this little Latin square design. As a result, the user can simply, analyze Latin squares with multi-observation or sub-sampling via MINITAB 18.1 as discussed above. So, “Error” term denotes the Experimental error and Pure-Error is the sampling error. This table contains all the information we achieved through our macro and some handy calculations. So, when the MINITAB 18.1 is present, the user should use it. We think no one used this ability of MINITAB 18 for doing Latin squares with sub-sampling. Of course, most of the reports, refer to lack-of-fit and pure error as components of the regression approach. But this can be used for analysis of latin square with sub-sampling or mult-observation design.

Relative efficiency (R.E.)

Suppose, we analyzed the above Latin square without sub-sampling or averaged over samplings (Table 1), and we invoke *ls2cd* macro. We analyze this design firstly as a simple Latin square and then as RBD to compare the results (Table 3).

Table 3. data set for Latin square design of ear length trait of maize

↓	C1-T	C2-T	C3	C4	C5	C6
	Genotype	Stress	Block	YLD	EL	SPE
1	BARAKAT-2	Control	1	10.13	20.67	34.67
2	BARAKAT-2	Control	2	7.44	18.33	35.67
3	BARAKAT-2	Control	3	5.52	20.00	31.67
4	BARAKAT-2	Drought-1	1	10.35	16.67	24.33
5	BARAKAT-2	Drought-1	2	6.37	19.67	26.00
6	BARAKAT-2	Drought-1	3	5.54	16.00	21.00
7	BARAKAT-2	Drought-2	1	11.57	19.33	28.33

The attributed ANOVA table will be constructed as follows:

Analysis of Variance for EL, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Block	2	2.395	2.395	***1.198***	0.26	0.791
Stress	2	6.395	6.395	***3.198***	0.71	0.586
Genotype	2	8.914	8.914	4.457	0.98	0.504
Error	2	9.062	9.062	***4.531***		
Total	8	26.765				

S = 2.12858   R-Sq = 66.14%   R-Sq(adj) = 0.00%

To run the program, you need to type in the command line as:

General Linear Model: Forage versus Treat

```
MTB > %ls2cd
Executing from file: D:\MINITAB14\MACROS\ls2cd.MAC
……
*       Example calling statement:
*    to run the macro (if executing file located
*    in macros folder of minitab) in session window type:
*    MTB> %ls2cd
```

```
**********************************************************
 Relative efficiency of LS to CR design:

 Please Enter: 'Block MS', 'Column MS', 'Error MS'
   and 'Treatment number', respectively
DATA> 1.198  3.198  4.531 3

Relative efficiency of RBD to CRD as a percent is:
 R.E.(%):    74.2551

Needed replications to perform the design as a CRD having the same efficiency
as LS:
Replications:    2.22765
```

According to the results above, the LS has a lower relative efficiency than CR design, and should we organize this 3 genotypes as a CR design, not an LS design. Analyzing this design as an RB design where rows considered as blocks, produces the results as:

MTB > %ls2rbr

```
Relative efficiency of LS to RB designs, where rows considered as blocks

 Please Enter 'Block MS', 'Error MS' & 'Treatments numbers', respectively
DATA> 1.198  4.531 3

Relative efficiency of LS to RB, where Rows considered as Blocks (%):
%R.E.    63.4032
```

If the relative efficiency of a given design, e.g., Latin square to CRD, is greater than 1, then Latin square design is more efficient than CRD and vice versa. If the relative efficiency of RBD is greater than that of CRD, the trial should be managed as an RBD, instead of CRD. However, the question arises: How many replications will be needed if one tries to run the current RCBD as a CRD to achieve the same accuracy? Or simply, how many more replications of an e.g. CRD would be necessary to gain the similar (same) precision as RBD for computing the treatment means? Our macro estimates the number of replication needed to run the trial as a CRD. For details of the equations used, see Gomez and Gomez (1984).

Regarding the comparison of an LS design against an RCB design through RE estimation, Gomez and Gomez (1984) pointed out: “the additional column-blocking, made possible by the use of an LS-design, is estimated to have increased the experimental precision over that of the RCB design with rows as blocks by 266%; whereas the additional row-blocking in the LS design did not increase precision over the RCB design with columns as blocks.

Orthogonal Contrasts

For the LS design comparing 20 genotypes of maize in this experiment, suppose we want to compare 9 of the 20 corn genotypes (from BARAKAT-2 to GARZA) against the remaining 11 genotypes (HYDRO to ZP-606) for the EL trait. The data on treatments are given, but the coefficients given here need not be defined or inserted manually when executing our macro “*contrast*”, because this macro estimates them automatically (Table 4), while other programs like SAS are not able to do this at all.

Table 4. Twenty corn genotypes tested in this experiment
for comparing 2 groups in an orthogonal contrast comparison.

C1-T	C2
Gen	EL
BARAKAT-2	157.67
BARAKAT-74	179.00
BOLSON	180.99
DKC-6589	193.33
DKC-5901	147.99
DKC-6630	164.01
DKC-7211	183.34
DRAGMA	187.35
GARZA	145.67
HYDRO	170.67
INOVE	180.33
KSC-704	180.00
KSC-705	172.00
MASYL	193.32
MAYMAY-70	177.33
MIMI	156.66
NS-6010	179.68
NS-770	183.99
ZP-600	179.66
ZP-606	177.31

In this table, 9 and 11 genotypes are categorized into 2 adjacent groups. Note that the user can define up to 4 groups in this macro, but it must be ensured that there is no interference between genotypes in the groups. That is, all observations of a given group are placed within that group. The ANOVA table for these 20 genotypes, which has been analyzed as a 2-factor design (ignoring of the interaction effect) is:

Analysis of Variance for EL, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Block	2	1.631	1.631	0.815	0.28	0.757 ns
ENV	2	95.192	95.192	47.596	16.25	0.000 ***
GEN	19	366.693	366.693	19.300	6.59	0.000 ***
Error	156	456.787	456.787	2.928		
Total	179	920.303				

As seen, there is a significant difference between 20 genotypes for EL trait, affirming t that orthogonal contrast to be done. If the file "*Contrast*" is located within the MACROS folder of MINITAB main root folder, as discussed earlier, in the enabled session window of MINITAB, type the following command and respond to the prompts and enter the information requested as follows:

MTB > %contrast c2

Contrast comparisons
Enter the size of the first, second, third and the last group:
DATA> 9 11
DATA> 0 0

Which groups do you need, e.g. 2 3:
DATA> 1 2

Enter: R, dfE, & MSE, respectively as Data prompt:
DATA> 3 156 2.928

Data Display

Row	EL	Coeffs	D.f.	SS=MS	F	P value
1	167.67	1.00000	1	43.2230	14.7620	0.0001771
2	179.00	1.00000				
3	180.99	1.00000				
4	193.33	1.00000				
5	147.99	1.00000				
6	164.01	1.00000				
7	183.34	1.00000				
8	187.35	1.00000				
9	145.67	1.00000				
10	170.67	-0.81818				
11	180.33	-0.81818				
12	180.00	-0.81818				
13	172.00	-0.81818				

```
14  193.32  -0.81818
15  177.33  -0.81818
16  155.66  -0.81818
17  179.68  -0.81818
18  183.99  -0.81818
19  179.66  -0.81818
20  177.31  -0.81818
```

According to results above, there is a significant difference between two groups of genotypes (9 against 11 genotypes). If this comparison is performed with the independent or 2-*t* test, the results will be as follows, in which there is no significant difference between the two groups, and this emphasizes that such comparison must be realized with the contrast method.

```
Two-Sample T-Test and CI: EL1, EL2

      N    Mean  StDev  SE Mean
EL9    9   172.2   17.0      5.7
EL11  11  177.27   9.32      2.8

Difference = mu (EL1) - mu (EL2)
Estimate for difference:  -5.11818
95% CI for difference:  (-19.01478, 8.77842)
T-Test of difference = 0 (vs not =): T-Value = -0.81  P-Value = 0.435  DF = 11
```

After entering the information necessary for the test, the main output of the macro is displayed, in which "*coeffs*" refers to the orthogonal coefficients for the test. It is important that these coefficients be estimated in the order in which their sum equals zero and for two groups (here 9 versus the rest 11 genotypes), the signs of the first group coefficients are opposite of the first group. However, the results of the test with one degree of freedom suggest that there was a significant difference between the two groups. In other words, the first 9 genotypes are different (P<0.001) from the remainder 11 maize genotypes.

Curve fitting

To run this macro, it is necessary that the traits *i.e.* the length of the ear (EL) and the number of grains per ear (SPE) to be located in 2 columns, such as c4 and c5, and type in the front of the command line prompt as *MTB>% curvfit c4 c5*, after activating the window session. In the following, the results will be as follows:

```
%"curvfit" C4 C5
….
Note *  Example calling statement:
Note *  Dependent and Independent variables must be raw data
Note *  e.g. in C4 & C5 columns of the worksheet, respectively.
Note *  Then type in session window as:
Note *  MTB> %curvfit c4 c5.
Note *----------------------
      …..
Data:
```

Row	EL	SPE
1	20.67	34.67
2	18.33	35.67
3	20.00	31.67
4	16.67	24.33
………………………………………………		

179	19.33	32.33
180	20.00	37.33

Trend 1:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	344.5	344.512	106.50	0.000
Error	178	575.8	3.235		
Total	179	920.3			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.79855	37.43%	37.08%	35.73%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	13.558	0.586	23.15	0.000	
X	0.1795	0.0174	10.32	0.000	1.00

Y = A + BX:
Regression Equation: $EL = 13.558 + 0.1975SPE$

Trend 1 shows a linear relationship between the two response variables (EL or ear length) and the predictor (SPE or number of seeds per ear), and in the analysis of variance, *F* index has a significant condition, which is also low (35.37%). And due to this low-level coefficient of determination, it is better to apply and fit other regression models to this data set.

Trend 2:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	352.253	176.127	54.88	0.000
X	1	0.156	0.156	0.05	0.826
X2	1	7.742	7.742	2.41	0.122
Error	177	568.050	3.209		
Total	179	920.303			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.79146	38.28%	37.58%	35.56%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	16.70	2.11	7.93	0.000	
X	-0.030	0.136	-0.22	0.826	61.59
X2	0.00329	0.00212	1.55	0.122	61.59

Y = A + BX + CX^2:
Regression Equation: $EL = 16.70 - 0.030SPE + 0.00329SPE^2$

The second trend is a quadratic trend between two maize traits. *i.e.* EL and SPE, with A nonsignificant slope but constant, and a low R^2 such as that for linear trend, and so, will not be mentioned as the best trend for the two data sets measured.

Trend 3:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value

Regression	3	376.90	125.633	40.69	0.000
X	1	24.23	24.233	7.85	0.006
X2	1	27.45	27.448	8.89	0.003
X3	1	24.64	24.644	7.98	0.005
Error	176	543.41	3.088		
Total	179	920.30			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.75714	40.95%	39.95%	32.93%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	31.95	5.78	5.53	0.000	
X	-1.631	0.582	-2.80	0.006	1173.61
X2	0.0565	0.0190	2.98	0.003	5138.14
X3	-0.000564	0.000200	-2.83	0.005	1476.37

Y = A + BX + CX^2 + CX^3
Regression Equation: $EL = 31.95 - 1.631SPE + 0.0565SPE^2 - 0.000564SPE^3$

The 3rd trend above is a cubic trend with significant slopes but low R^2 and so inappropriate.

Trend 4:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	333.3	333.298	101.07	0.000
Xsq	1	333.3	333.298	101.07	0.000
Error	178	587.0	3.298		
Total	179	920.3			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.81598	36.22%	35.86%	34.33%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	8.34	1.11	7.49	0.000	
Xsq	1.955	0.194	10.05	0.000	1.00

Y = SQRT(X) or $Y = \sqrt{X}$ **and simply linearized as:**
Regression Equation: $EL = 8.34 + 1.955\sqrt{SPE}$

The fourth equation above is square root of the predictor with significant slope and low R^2 and inappropriate.

Trend 5:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	316.3	316.264	93.20	0.000
logX	1	316.3	316.264	93.20	0.000
Error	178	604.0	3.393		
Total	179	920.3			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.84214	34.37%	34.00%	32.22%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.57	1.86	0.84	0.400	
logX	11.90	1.23	9.65	0.000	1.00

Y = log(X), linearized as: Y=A+B[log(X)]

Regression Equation: $EL = 1.57 + 11.90\log(SPE)$

Trend 6:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.9894	0.989363	105.58	0.000
X	1	0.9894	0.989363	105.58	0.000
Total	179	2.6573			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0968023	37.23%	36.88%	35.53%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2.6449	0.0315	83.92	0.000	
X	0.009621	0.000936	10.28	0.000	1.00

Y = e^X, linearized as; LnY=A+BX

Regression Equation: $Ln(EL) = 2.6449 + 0.009621SPE$

Trend 7:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	263.0	263.041	71.24	0.000
X´	1	263.0	263.041	71.24	0.000
Error	178	657.3	3.692		
Total	179	920.3			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.92158	28.58%	28.18%	25.40%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	23.653	0.519	45.55	0.000	
X´	-129.2	15.3	-8.44	0.000	1.00

Y = 1/X, linearized as: Y=A+B/(X)

Regression Equation: $EL = 23.653 - 129.2SPE'$

Trend 8:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.9894	0.989363	105.58	0.000
X	1	0.9894	0.989363	105.58	0.000
Error	178	1.6680	0.009371		
Total	179	2.6573			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0968023	37.23%	36.88%	35.53%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2.6449	0.0315	83.92	0.000	
X	0.009621	0.000936	10.28	0.000	1.00

Y = A(B^X), linearized as: LnY=LnA+[Ln(B)](X)
Real INTERCEPT and SLOPE are

Regression Equation: $Ln(EL) = 2.6449 + 0.009621SPE$

A 14.0826
B 1.00967

Trend 9:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.9153	0.915294	93.52	0.000
LnX	1	0.9153	0.915294	93.52	0.000
Error	178	1.7420	0.009787		
Total	179	2.6573			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0989282	34.44%	34.08%	32.33%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.9986	0.0997	20.05	0.000	
LnX	0.2780	0.0287	9.67	0.000	1.00

Y = A[(X)^B], linearized as: Ln(Y)=Ln(A)+B[Ln(X)]
Real INTERCEPT is

Regression Equation: $Ln(EL) = 1.9986 + 0.2780Ln(SPE)$

A 7.37906

Trend 10:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.002900	0.002900	102.42	0.000
X	1	0.002900	0.002900	102.42	0.000
Error	178	0.005040	0.000028		
Total	179	0.007939			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0053209	36.52%	36.17%	34.82%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.06927	0.00173	39.98	0.000	
X	-0.000521	0.000051	-10.12	0.000	1.00

Y = 1/(A+BX)

Regression Equation, **linearized as:** $EL' = 0.06927\text{-}0.000521SPE$

Trend 11:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.9894	0.989363	105.58	0.000
X	1	0.9894	0.989363	105.58	0.000
Error	178	1.6680	0.009371		
Total	179	2.6573			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0968023	37.23%	36.88%	35.53%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2.6449	0.0315	83.92	0.000	
X	0.009621	0.000936	10.28	0.000	1.00

Y = A[E^(BX)], linearized as: Ln(Y)=Ln(A)+BX
Real intercept is:

Regession Equation: $Ln(EL) = 2.6449{+}0.009621SPE$

A 14.0826

Trend 12:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.7675	0.767524	72.29	0.000
X´	1	0.7675	0.767524	72.29	0.000
Error	178	1.8898	0.010617		
Total	179	2.6573			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.103039	28.88%	28.48%	25.73%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.1878	0.0278	114.49	0.000	
X´	-6.981	0.821	-8.50	0.000	1.00

******** Y =A*E^(B/X), linearized as: Ln(Y)=Ln(A)+B(1/X) ********
Real INTERCEPT is

Regression Equation: $Ln(EL) = 3.1878 - 6.98SPE'$

A 24.2339

Trend 13:
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	33.7556	33.7556	1938.63	0.000
X	1	33.7556	33.7556	1938.63	0.000
Error	178	3.0994	0.0174		
Total	179	36.8550			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.131955	91.59%	91.54%	91.23%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-2.1155	0.0430	-49.24	0.000	
X	-0.05620	0.00128	-44.03	0.000	1.00

Y = A(X^2)E^(BX)

Regression Equation: $EL = -2.1155 - 0.05620\text{SPE}$

Real intercept is:

A 0.120572

Trend 14:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	0.002283	0.002283	71.84	0.000
X´	1	0.002283	0.002283	71.84	0.000
Error	178	0.005656	0.000032		
Total	179	0.007939			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0056371	28.76%	28.36%	25.63%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.03979	0.00152	26.12	0.000	
X´	0.3807	0.0449	8.48	0.000	1.00

********** **Y = 1/(A + B/X), as: 1/Y=A+B/X** **********

Regression Equation: $\text{EL}' = 0.03979 + 0.3807\text{SPE}'$

According to the above results, almost all (except for one) of the fourteen trends had a low coefficient of variations although they having significant regression coefficients, they were not proportional to the pair of variable EL and SPE. But one of them, a sophisticated trend, that had largely a high coefficient of determination (R^2=91.54%) has been considered. The greater R^2 suggests a stronger relationship. Of course, according to MINITAB Company, a high R-squared does not necessarily show that the model fits well. It may be a surprise, but one can look at the fitted line plot and also the residual plot. Our macro was able, to realize the almost real relationship between the trait of the EL response and the predictive trait of the SPE. Thus, this macro searches and finds the best relationship equation among 14 equations defined, for a dependent variable and relevant independent variable, and could be easily applied by researchers in various fields of study, especially genetics.

Path coefficient analysis

In this real experiment, the grain yield per hectare (YLD) was considered as the ultimate dependent or response variable and considered the two traits of EL(ear length) and SPE(the number of seeds per ear) as predictive or affective (causative) variables. Of course, it should be noted that there was no linear relationship between YLD and the two traits in this experiment, and instead, there were nonlinear relations. Thus, we note that in these calculations do not look for the direct, strong and significant direct effects of the predictors on the dependent. However, put the data of the three traits listed in Table 2 into 3 adjacent columns in the MINITAB worksheet, so that the dependent YLD is placed in the first column of the left, then other independent variables were put in the second to the last column (as shown in Table 2) and then in the front of the command line in session window, do by typing the following command and reply to the questions while popping up. An interesting innovative and exquisite problem about MINITAB 18 is that, unlike other previous versions, it separates the executable unnecessary commands from the numerical results of the session window, and the numerical results, such as the JMP software, are divided into segments, which also have the storage functionality, both text, and image.

Command line (unnecessary commands):

```
MTB > % multpath c1-c3
Executing from file: D:/multpath.MAC
 …..

 How many Dependent & Independent variables
 do you have, respectively?
DATA> 1 2
 All DATA: Y1 Y2 ...Yk.1  vs  X1 X2 ... X.p  for raw data,
 &  rXiY1 ... rXiYk.1 vs  1 ... rX1Xp for correlation coefficients.

 For Raw data enter 1,
 For correlation coeffs Enter 0, respectively:
DATA> 1
 Descriptive statistics of data
 Correlation coefficients, and their significance:
 **************************************************
 *  Multiple regression outputs for various sets of
 *  Dependent versus Independent variables, respectively:
 ********************************************************
 Y1  vs   X1, X2, ...
 Y2  vs   X1, X2, ...
 ...................
  *****    ***   *****
 *  *  *   * *     *
 *  *  *   ***     *
 *  *  *   * *     *

 Matrix outputs for first Dependent vs. Independent variables,
  like Y1 vs. X1 X2 X3, ...

Answer = 0.0103969

 U is RESIDUAL of PATH model.
```

The main numerical(2nd) section of the session outputs i.e. descriptive statistics:

Data

Row	YLD	EL	SPE
1	10.13	20.67	34.67
2	7.44	18.33	35.67
3	5.52	20.00	31.67
4	10.35	16.67	24.33
5	6.37	19.67	26.00

Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
YLD	180	0	8.596	0.190	2.543	0.000	6.900	8.720	10.495	15.050
EL	180	0	19.441	0.169	2.267	13.670	18.000	20.000	21.000	23.670
SPE	180	0	32.767	0.576	7.727	11.670	26.083	33.670	38.670	48.000

Correlations

	YLD	EL
EL	0.102	
	0.174	
SPE	0.057	0.612
	0.445	0.000

Cell Contents
Pearson correlation
P-Value

Regression Analysis: YLD versus EL, SPE

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	12.04	6.01842	0.93	0.397
EL	1	8.24	8.24101	1.27	0.261
SPE	1	0.05	0.04638	0.01	0.933
Error	177	1145.69	6.47285		
Lack-of-Fit	166	1110.98	6.69264	2.12	0.081
Pure Error	11	34.72	3.15601		
Total	179	1157.73			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.54418	1.04%	0.00%	0.00%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	6.36	1.66	3.83	0.000	
EL	0.120	0.106	1.13	0.261	1.60
SPE	-0.0026	0.0311	-0.08	0.933	1.60

Regression Equation

YLD = 6.36 + 0.120 EL - 0.0026 SPE

Fits and Diagnostics for Unusual Observations

Obs	YLD	Fit	Resid	Std Resid	
31	11.660	8.824	2.836	1.15	X
32	6.300	8.864	-2.564	-1.04	X
33	7.610	9.144	-1.534	-0.64	X
40	0.000	8.052	-8.052	-3.21	R
41	0.000	8.373	-8.373	-3.32	R
42	0.000	8.132	-8.132	-3.24	R
69	15.050	8.969	6.081	2.41	R
127	14.100	8.576	5.524	2.18	R
133	14.960	8.679	6.281	2.49	R
140	6.780	7.982	-1.202	-0.49	X

R Large residual
X Unusual X

Matrix m.1

0.101769
0.057260

Cumulative Distribution Function

Data

Row	Direct effects	T-values	P-value
1	0.106664	0.631127	0.264381
2	-0.008002	0.013893	0.494465

Matrix m.3

0.106664	-0.0048957
0.065261	-0.0080016

Residual effect:

U 0.994788

In outputs, the dependent variable (grain yield or YLD) is associated with independent variables (EL and SPE) in the linear equation, e.g., YLD = 6.36 + 0.120*EL - 0.0026*SPE. In this equation, 6.36 is the constant (intercept) and 0.120 and 0.0026 are the slopes (β_1 *and* β_2) of the linear equation. The “VIF” or variation inflation factors of two independent traits were lower 10, and *F* test of ANOVA table was nonsignificant too. Nevertheless, the multiple linear equation indicated the nonlinearity of the model, and the determination coefficient of the model (1%) proves that linear model does not fit the data and so other models should be evaluated, which has been discussed earlier.

In the last part of the macro output, the matrix results are presented. The “*m.1*” matrix represents the correlation coefficients of all independent variables (EL and SPE) with the relevant dependent variable (YLD). Then, direct effects of independent variables on the dependent variable, along with *t*-test and P-values, are presented. The direct effects of grain yield (YLD) variable on two independents EL and SPE (0.106664 and -0.008002, respectively) were not significant (P>0.05). Finally, “*m.3*” matrix shows direct (diagonal) and indirect effects (off-diagonal elements). The U or residuals of path analysis were 0.994788, indicating the proportional effects of unknown factors on the dependent variable (Figure 2). We suggest drawing the path diagram by means of specific graph builders like *Graphviz*. Analyses mentioned above showed that the "*MULTPATH*" macro may be practically used by many researchers in various fields of study, especially agriculture and biological sciences. For more information refer to our MULTPATH paper published (Arminian *et al.,* 2008).

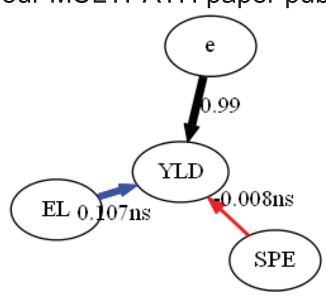


Fig. 2. Path analysis with dependent (YLD or grain yield) and 2 independent variables (EL or ear length and SPE or number of seeds per ear) in maize. One- and two-headed arrows for EL and SPE denote direct effects (coefficients) and simple Pearson correlation coefficients, respectively. U: residual. *(P<0.05); **(P<0.01).

Univariate parametric stability

The "*MINISTAB*" exec macro operates as follows: as mentioned, for stability analyses, our maize data set (grain yield) was used. The dataset tests consisted of 20 genotypes in 3 environments (drought stress) with 3 replications (Table 5).

Table 5. The effect of 20 genotypes in 3 environments at 3 blocks(replicates) on grain yield of maize

↓	C1-T	C2-T	C3	C4
	Gen	ENV	Block	YLD
1	BARAKAT-2	Control	1	10.13
2	BARAKAT-2	Control	2	7.44
3	BARAKAT-2	Control	3	5.52
4	BARAKAT-2	Drought-1	1	10.35

Before executing the *MINISTAB* macro, it is essential that data be first subjected to ANOVA, whose results are as follows. The ANOVA table as a 2 factorial experiment in randomized completely design via GLM schedule in MINITAB 14, is performed through Stat>General linear model> and the result is as follows:

Analysis of Variance for YLD, using Adjusted SS for Tests

Source DF Seq SS Adj SS Adj MS F P

Block 2 145.874 145.874 72.937 26.38 0.000 **
Gen 19 494.437 494.437 26.023 9.41 0.000 **
ENV 2 10.244 10.244 5.122 1.85 0.161 ns
Gen*ENV 38 180.862 180.862 4.760 1.72 0.014 *
Error 118 326.313 326.313 2.765
Total 179 1157.731

S = 1.66294 R-Sq = 71.81% R-Sq(adj) = 57.24%

According to the ANOVA results, there was a significant difference for blocks, genotypes and their effects. So the stability analyses like AMMI could be logically performed. To perform the stability macros, suppose the files and their subroutines are located in a local drive (here D:\). For executing the univariate stability "*MINISTAB*", follow these steps: *File > Other files > Run an exec > Select files > "D:\MINISTAB folder" > Ministab.* To choose the main macro file (*MINISTAB*) and execute it, the following table is used for the dataset as arranged in a matrix (G×E) form where rows denote the genotypes and columns denote the environments or locations (Table 6):

Table 6. Data of grain yield sorted as a two-way table (G×E) suitable for the stability analysis by MINISTAB macro:

	Env1	Env2	Env3	Genotypes means	Gen
1	7.7000	7.4200	8.2600	7.7933	BARAKAT-2
2	8.4400	8.4300	8.4900	8.4533	BARAKAT-74
3	5.3400	5.0500	5.2700	5.2200	BOLSON
...	...	...	...	...	...
20	11.4600	12.0100	11.8600	11.7767	ZP-606
21	8.8355	8.2710	8.6830	8.5965	
22	0.2390	-0.3255	0.0865		

The order of the data matrix must be, as shown above, i.e., the data must be arranged for example from the first column (e.g. C1) to the next columns (here e.g. C4), left to right. Of course, local macros do not have such limitations as in our *AMMI* macro. Of course, we adjusted the MINISTAB macro to construct the two-way table (Table 6) automatically from a design table (Table 5), which will be employed later. As genotype-by-environment interaction effect(GE) was significant (prerequisite of stability analysis), so the stability analysis was performed. For this, the *MINISTAB* executive macro produced the following output, in which some parameters are taken from ANOVA table, which must be inserted as requested by macro, including the number of environments and replications, error MS and confidence level ($1-\alpha$):
MTB > Execute "D:\MINISTAB_parametric stability\MINISTAB.MTB" 1.

Executing from file: D:\MINISTAB_parametric stability\MINISTAB.MTB

How many Environments(Locations)do you have? Please choose At must 3500).
DATA> 3

Enter EMS(Error Mean Square); R(Replications number); and Confidence level in C3501 as DATA>prompt, respectively.
DATA> 2.765 3 .95

Data Display

Row	Gen	Env1	Env2	Env3	means	Si^2-Roemer
1	BARAKAT-2	7.7000	7.4200	8.2600	7.7933	0.1829
2	BARAKAT-74	8.4400	8.4300	8.4900	8.4533	0.0010
3	BOLSON	5.3400	5.0500	5.2700	5.2200	0.0229
4	DKC-5689	9.3200	8.5200	8.0900	8.6433	0.3896
5	DKC-6101	9.3800	0.0000	8.3000	5.8933	26.3401

6	DKC-6630	9.4700	8.0000	8.5200	8.6633	0.5556
7	DKC-7211	7.8300	9.0200	9.4900	8.7800	0.7321
8	DRAGMA	11.0600	12.1900	11.5600	11.6033	0.3206
9	GARZA	10.1900	9.6800	10.0200	9.9633	0.0674
10	HYDRO	9.5300	9.3600	9.5800	9.4900	0.0133
11	INOVE	8.4200	9.1800	7.8000	8.4667	0.4777
12	KSC-704	8.1500	7.3300	8.1200	7.8667	0.2162
13	KSC-705	8.2900	9.2600	8.4500	8.6667	0.2704
14	MASYL	8.8900	6.5500	7.1600	7.5333	1.4734
15	MAYMAY-70	9.6500	9.1800	10.2100	9.6800	0.2659
16	MIAMI	6.8100	6.9700	6.5300	6.7700	0.0496
17	NS-6010	10.8200	10.6600	9.8700	10.4500	0.2587
18	NS-770	6.8100	6.9700	6.5300	6.7700	0.0496
19	ZP-600	9.1500	9.6400	9.5500	9.4467	0.0680
20	ZP-606	11.4600	12.0100	11.8600	11.7767	0.0808
21		8.8355	8.2710	8.6830	8.5965	
22		0.2390	-0.3255	0.0865		

Row	CVi-Francis/Kannenberg	Wi^2-Wricke	Si^2-Shukla	Plaisted/Peterson
1	5.4881	0.2573	0.0548	0.8606
2	0.3803	0.1575	-0.0006	0.8344
3	2.8990	0.0397	-0.0661	0.8034
4	7.2218	0.6418	0.2685	0.9618
5	87.0859	46.9313	25.9848	13.1432
6	8.6042	0.4892	0.1837	0.9217
7	9.7452	2.1223	1.0909	1.3514
8	4.8800	1.4610	0.7235	1.1774
9	2.6063	0.0028	-0.0865	0.7937
10	1.2152	0.0778	-0.0449	0.8134
11	8.1636	1.7280	0.8719	1.2477
12	5.9111	0.0744	-0.0468	0.8125
13	6.0004	1.3152	0.6426	1.1390
14	16.1131	1.8934	0.9638	1.2912
15	5.3270	0.2995	0.0783	0.8717
16	3.2897	0.4224	0.1465	0.9041
17	4.8672	0.7481	0.3275	0.9898
18	3.2897	0.4224	0.1465	0.9041
19	2.7611	0.5564	0.2210	0.9393
20	2.4142	0.6211	0.2569	0.9564

Row	Plaisted	bi: Regression	t-test for bi	R^2-Pintus	Bi-Perkins/Jinks
1	1.66641	0.8184	-0.1495	0.312222	-0.1816
2	1.66932	0.0444	-9.4883	0.163007	-0.9556
3	1.67277	0.5180	-30.2331	0.999053	-0.4820
4	1.65516	0.9030	-0.0501	0.178453	-0.0970
5	0.30167	17.3539	5.8859	0.975006	16.3539
6	1.65962	2.3237	1.2529	0.828688	1.3237
7	1.61187	-1.4292	-0.9497	0.237928	-2.4292
8	1.63121	-1.9030	-7.8014	0.963172	-2.9030
9	1.67385	0.8871	-1.8333	0.995205	-0.1129
10	1.67165	0.3498	-3.5471	0.784557	-0.6502
11	1.62340	-1.7649	-1.7531	0.556014	-2.7649
12	1.67175	1.5497	1.5021	0.947175	0.5497
13	1.63547	-1.7701	-14.2054	0.988009	-2.7701
14	1.61857	3.5885	1.2338	0.745276	2.5885
15	1.66517	1.1810	0.1379	0.447319	0.1810

16	1.66158	-0.4474	-2.3433	0.344094	-1.4474
17	1.65205	-0.1765	-0.6789	0.010264	-1.1765
18	1.66158	-0.4474	-2.3433	0.344094	-1.4474
19	1.65766	-0.7323	-3.3873	0.672168	-1.7323
20	1.65577	-0.8468	-3.8440	0.756491	-1.8468

Row	Di^2-Perkins/Jinks	D^2-Eberhart/Russel	Pi-Lin/Binns	PGEi-Lin/Binns
1	0.14303	0.25164	0.1699	0.1006
2	0.15712	0.00173	0.0010	0.2196
3	-0.00608	0.00004	0.0148	0.1682
4	0.50275	0.64020	0.3588	0.2023
5	-4.43230	1.31669	14.8585	19.5193
6	-0.43169	0.19037	0.5106	0.3173
7	1.77389	1.11583	0.4961	0.3793
8	0.84330	0.02362	0.2790	0.1604
9	-0.13140	0.00065	0.0482	0.1258
10	0.05696	0.00573	0.0085	0.1846
11	1.19679	0.42421	0.4137	0.2528
12	-0.33523	0.02284	0.1122	0.1728
13	0.78083	0.00649	0.2662	0.1354
14	-0.30286	0.75064	1.4114	1.1925
15	0.06162	0.29392	0.2291	0.1346
16	0.38822	0.06507	0.0365	0.1315
17	0.74283	0.51209	0.1547	0.1615
18	0.38822	0.06507	0.0365	0.1315
19	0.46495	0.04461	0.0414	0.1445
20	0.49877	0.03937	0.0542	0.1290

Row	Di-Lin/Binns	Di^2-Hanson	dsi(B)^2-Hanson/Mongomery
1	7.7579	51.352	0.3138
2	8.4514	50.486	0.0395
3	5.1976	50.837	0.6761
4	8.6043	51.820	49.9506
5	5.1428	114.400	41.0378
6	8.5628	53.074	4.6301
7	8.8418	50.991	0.8531
8	11.6856	49.861	1.3598
9	9.9250	51.166	0.0595
10	9.4749	50.709	1.2913
11	8.5430	50.265	2.5178
12	7.7996	51.893	1.9334
13	8.7432	49.847	5.5149
14	7.3781	55.731	2.9725
15	9.6289	51.753	1.0878
16	6.7893	50.264	0.2246
17	10.4576	50.858	0.2246
18	6.7893	50.264	0.2313
19	9.4783	50.116	0.0024
20	11.8133	50.067	

Row	dst(A)^2-Fox/Rosielle	F-test for dst(B)^2-Abou-El-Fitouh...	Shukla
1	0.00246	0.76582	0.0595
2	0.65648	0.00333	-0.0007
3	1.48077	0.37873	-0.0717
4	1.83708	1.29355	0.2913

5	0.43721	0.44526	28.1933
6	1.04466	0.57244	0.1993
7	0.82143	0.28666	1.1837
8	0.08687	0.20644	0.7850
9	0.31605	0.32985	-0.0939
10	3.04478	0.34280	-0.0487
11	2.27806	1.40334	0.9460
12	0.92435	0.76744	-0.0507
13	0.01190	0.00488	0.6972
14	1.78324	1.31004	1.0457
15	4.43292	1.72340	0.0849
16	0.00393	0.07191	0.1590
17	0.00393	0.07191	0.3554
18	2.17209	0.27558	0.1590
19	0.01062	0.00071	0.2398
20		0.2788	

Inverse Cumulative Distribution Function
Student's t distribution with 2 DF

P(X <= x) x
0.95 2.91999

Inverse Cumulative Distribution Function

F distribution with 2 DF in the numerator and 120 DF in the denominator

P(X <= x) x
0.95 3.07178

Using the above outputs for the MINISTAB macro, stable varieties via univariate stability parameters were distinguished. For example, BARAKAT-74 was considered to be the most stable maize genotype on the basis of the Si^2 criterion of Roemer (1917), CVi index of Francis and Kannenberg and Pi of Lin and Binns stability indices, and also genotype GARZA was the most stable genotype based on the stability indices of Wi^2 of Wricke, Si^2 of Shukla, Plaisted and Peterson and F-test of Shukla test. Genotype DKC-6101 was the most stable genotype on the basis of stability tests of Plaisted, Di^2 of Perkins and Jinks, Di of Lin and Binns. Genotype BOLSON was the most stable according to the deviation of Eberhart and Russel. However, it is obvious that some genotypes as mentioned above could be selected as stable genotypes against various fluctuating environments (here drought stress). Another point of view which is important is that the selection approach could be not only on the basis of stability parameters but should be remembered on the basis of yield. In this subject, GARZA was a high yielding genotype, and BARAKAT-74 was a good-yielding too. Thus, one can select GARZA as a stable and high-yielding genotype. To achieve the most stable genotype, one may include a known genotype as the control in the experiment. This macro is useful for plant breeders and geneticists working on yield stability.

Multivariate parametric stability analysis of genotypes

According to the data set (Table 5), we organize such data, so that genotypes and environments (or locations) must be numerical but not characters (as Table 7).

Table 7. Four column data set organized for most common form of AMMI analysis.

↓	C1	C2	C3	C4
	GEN	ENV	Block	Response
1	1	1	1	10.13
2	1	1	2	7.44
3	1	1	3	5.52

To execute the macro, it is best to place the program file inside the Macros folder of MINITAB and then type in the front of the command line: %AMMI C1-C4. The program then runs and gives the following results.

```
****************************************************************************
*                   AMMI.MAC
*   Local macro performing Additive Main Effects and Multiplicative
*   Interaction (AMMI) analysis on an ANOVA data or an ExG table of means,
…….
****************************************************************************
*   Example calling statements:
*   1: Two-way data serial (data in 3 columns: Gen, Env, and Response),
*  e.g.,in C4-C6 columns. Then at command prompt type, MTB> %AMMI C4-C6.
….

Additive Main effects & Multiplicative Interaction (AMMI) analysis

 *********   ************   ************   ******
 *********   ************   ************   ******
 **    **   **  **  **   **  **  **     **
 *********   **  **  **   **  **  **     **
 *********   **  **  **   **  **  **     **
 **    **   **  **  **   **  **  **     **
 **    **   **  **  **   **  **  **   ******
 **    **   **  **  **   **  **  **   ******

Macro is running,
please wait...

Do you work with Locations instead of Environments(Y/S)?
n

Do you work with a two-way (ExG or LG) table (genotypes in columns) (Y/S)?
n
```

In this part of the program, MINITAB asks the user to specify whether it is dealing with columns of data or with a matrix or two-way table of genotypes and environments (locations)?. The key merit of our program over other stability analysis programs is that the user can enter both columns of data or a matrix or two-sided table of info.

```
 ************ All raw data ************
Data Display

 Row   GEN  ENV  Block    YLD

   1     1    1     1   10.13
   2     1    1     2    7.44
   3     1    1     3    5.52
   ………………………………………………………………………………
 180    20    3     3   10.58

 ************ Two-way Analysis of Variance ************

General Linear Model: YLD versus ENV, GEN, Block

Factor      Type Levels Values
ENV        fixed      3 1 2 3
```

```
Block(ENV) random      9 1 2 3 1 2 3 1 2 3
GEN        fixed      20  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18
                       19 20
Analysis of Variance for YLD, using Adjusted SS for Tests
Source       DF     Seq SS     Adj SS     Adj MS       F      P
ENV          2     10.244     10.244      5.122    0.20  0.826
Block(ENV)   6    155.733    155.733     25.956    9.35  0.000
GEN         19    494.437    494.437     26.023    9.37  0.000
ENV*GEN     38    180.862    180.862      4.760    1.71  0.016
Error      114    316.454    316.454      2.776
Total      179   1157.731
```

In this section, the variance analysis table is displayed as above. The significant interaction effect (ENV*GEN) implies that the multivariate stability analyses like AMMI could be applied.

******* Genotype means and standard errors ******

Data Display

Row	Genotype	Genmean	StEr
1	1	7.79	1.11
2	2	8.45	0.68
3	3	5.22	0.38
4	4	8.65	1.01
5	5	5.89	2.43
6	6	8.66	0.94
7	7	8.78	0.53
8	8	11.60	0.68
9	9	9.97	0.80
10	10	9.49	0.62
11	11	8.47	0.90
12	12	7.87	0.33
13	13	8.66	0.96
14	14	7.53	0.64
15	15	9.68	1.94
16	16	6.77	0.40
17	17	10.45	0.40
18	18	6.77	0.40
19	19	9.45	0.61
20	20	11.78	0.40

****** Coefficent of Variation (CV) ******

Data Display

CV   0.295839

Below user must enter the EMS from the ANOVA table, manually:

Enter acquired 'Error mean square' from ANOVA table above, as data>prompt
DATA> 2.776

******** All data averaged over reps ********
Res=residuals, from the additive model, i.e. interactions.

Data Display (residuals)

Row	GEN	ENV	YLD	Res
1	1	1	7.6967	-0.33500
2	1	2	7.4200	-0.04667
3	1	3	8.2600	0.38167
………				
60	20	3	11.8633	-0.00056

********** ExG table means, Averaged over replications.***********
* 'Emeans' and 'Enveff' are env means & effects in ExG table.
* The genot means and effects are two last rows, respectively.

Data Display

Row	Envs	Genot_1	Genot_2	Genot_3	Genot_4	Genot_5	Genot_6
1	E1	7.69667	8.44000	5.34333	9.32333	9.38333	9.47333
2	E2	7.42000	8.43000	5.05000	8.52333	0.00000	7.99667
3	E3	8.26000	8.49000	5.27000	8.09333	8.29667	8.52000
4	Gmeans	7.79222	8.45333	5.22111	8.64667	5.89333	8.66333
5	Genoteff	-0.80417	-0.14306	-3.37528	0.05028	-2.70306	0.06694

Row	Genot_7	Genot_8	Genot_9	Genot_10	Genot_11	Genot_12	Genot_13
1	7.83333	11.0567	10.1900	9.53333	8.42000	8.15333	8.28667
2	9.01667	12.1933	9.6833	9.36333	9.17667	7.33000	9.25667
3	9.49000	11.5600	10.0233	9.57667	7.80000	8.12000	8.44667
4	8.78000	11.6033	9.9656	9.49111	8.46556	7.86778	8.66333
5	0.18361	3.0069	1.3692	0.89472	-0.13083	-0.72861	0.06694

Row	Genot_14	Genot_15	Genot_16	Genot_17	Genot_18	Genot_19	Genot_20
1	8.89333	9.6467	6.80667	10.8233	6.80667	9.15000	11.4567
2	6.54667	9.1833	6.97000	10.6567	6.97000	9.63667	12.0133
3	7.16333	10.2067	6.52667	9.8667	6.52667	9.55000	11.8633
4	7.53444	9.6789	6.76778	10.4489	6.76778	9.44556	11.7778
5	-1.06194	1.0825	-1.82861	1.8525	-1.82861	0.84917	3.1814

Row	Emeans	Enveff
1	8.83583	0.239444
2	8.27083	-0.325556
3	8.68250	0.086111
4	8.59639	

******** Adjusted scores(weights) for factor 1(e.g. genot.) *******

Data Display

Matrix Adscor.1

-0.00396	0.34259
-0.14149	0.06265
-0.07148	0.02075
-0.05284	-0.52915

```
 2.50965   0.17036
 0.17938  -0.33935
-0.31261   0.79074
-0.43004   0.20894
-0.01882  -0.01103
-0.09479   0.07095
-0.44518  -0.33330
 0.09029   0.07975
-0.41968   0.04690
 0.34906  -0.66991
 0.05103   0.35453
-0.23047  -0.11746
-0.20960  -0.43671
-0.23047  -0.11746
-0.24899   0.20397
-0.26899   0.20224
```

******* Adjusted scores(weights) for factor 2(e.g. env.) *******

Data Display

Matrix Adscor.2

```
 1.36053  -0.96212
-2.20757  -0.16174
 0.84703   1.12386
```

******** How many IPCA scores do you want? ********
DATA> 2

The AMMI analysis on the base of Gollob model is as follows:

****** Analysis of Variance for AMMIF (full model) *****

******** Gollob ANOVA table ********

Data Display

Row	Source	DfG	SSG	MSG	FGollob	PG	Percent	%Cumulative
1	IPCA.1	20	166.14	8.31	2.99	0.0001	91.86	91.86
2	IPCA.2	18	14.72	0.82	0.29	0.9977	8.14	100.00
3	Residual	0	0.00	0.00	0.00			

The table above shows that 2 first IPCA scores account for all the variation of G*E. So biplot of two first scores is best for studying the dispersión of genotypes against 2 IPCA scores. The AMMI analysis on the base of Cornelius model is as follows:

******** Cornelius ANOAVA table ********

Data Display

Row	Dfc	SSc	MSc	FCornel	Pc

```
 1    18   14.72   0.82     0.29  0.9977
```

******** Interaction effects matrix ********

Data Display

Matrix m.2

-0.00539	0.00874	-0.00335
-0.19251	0.31235	-0.11985
-0.09725	0.15780	-0.06055
-0.07188	0.11664	-0.04475
3.41446	-5.54022	2.12576
0.24406	-0.39600	0.15194
-0.42532	0.69012	-0.26479
-0.58509	0.94935	-0.36426
-0.02561	0.04155	-0.01594
-0.12896	0.20925	-0.08029
-0.60568	0.98276	-0.37708
0.12284	-0.19932	0.07648
-0.57099	0.92647	-0.35548
0.47491	-0.77057	0.29567
0.06943	-0.11266	0.04323
-0.31356	0.50878	-0.19522
-0.28516	0.46270	-0.17754
-0.31356	0.50878	-0.19522
-0.33876	0.54966	-0.21090
-0.36597	0.59382	-0.22785

********* AMMI-estimates per environments *********

Data Display

Row	Genots	Envi.1	Envi.2	Envi.3
1	Gen. 1	7.70	7.42	8.26
2	Gen. 2	8.44	8.43	8.49
3	Gen. 3	5.34	5.05	5.27
4	Gen. 4	9.32	8.52	8.09
5	Gen. 5	9.38	0.00	8.30
6	Gen. 6	9.47	8.00	8.52
7	Gen. 7	7.83	9.02	9.49
8	Gen. 8	11.06	12.19	11.56
9	Gen. 9	10.19	9.68	10.02
10	Gen.10	9.53	9.36	9.58
11	Gen.11	8.42	9.18	7.80
12	Gen.12	8.15	7.33	8.12
13	Gen.13	8.29	9.26	8.45
14	Gen.14	8.89	6.55	7.16
15	Gen.15	9.65	9.18	10.21
16	Gen.16	6.81	6.97	6.53
17	Gen.17	10.82	10.66	9.87
18	Gen.18	6.81	6.97	6.53
19	Gen.19	9.15	9.64	9.55
20	Gen.20	11.46	12.01	11.86

***** Ranked genotypes in environments via AMMI-estimates ******

Data Display

Row	Genots	Envi.1	Envi.2	Envi.3
1	Gen. 1	4.0	7.0	8.0
2	Gen. 2	9.0	9.0	11.0
3	Gen. 3	1.0	2.0	1.0
4	Gen. 4	12.0	10.0	6.0
5	Gen. 5	13.0	1.0	9.0
6	Gen. 6	14.0	8.0	12.0
7	Gen. 7	5.0	11.0	13.0
8	Gen. 8	19.0	20.0	19.0
9	Gen. 9	17.0	17.0	17.0
10	Gen.10	15.0	15.0	15.0
11	Gen.11	8.0	12.5	5.0
12	Gen.12	6.0	6.0	7.0
13	Gen.13	7.0	14.0	10.0
14	Gen.14	10.0	3.0	4.0
15	Gen.15	16.0	12.5	18.0
16	Gen.16	2.5	4.5	2.5
17	Gen.17	18.0	18.0	16.0
18	Gen.18	2.5	4.5	2.5
19	Gen.19	11.0	16.0	14.0
20	Gen.20	20.0	19.0	20.0

The table of ranked genotypes showed that e.g. in the first environment (or control condition), genotype 3 was the best. But in first drought condition, genotype 5 was the best.

******* Table of AMMI selections per environment *******
up to first 6 AMMI Selections

Data Display

Row	Envno	Emeans	Score	Sel.1	Sel.2	Sel.3	Sel.4	Sel.5	Sel.6
1	E1	8.83583	1.36053	g20	g 8	g17	g9	g15	g10
2	E3	8.68250	0.84703	g20	g 8	g15	g9	g17	g10
3	E2	8.27083	-2.20757	g 8	g20	g17	g9	g19	g10

In the table above, 6 best genotypes of 20 genotypes were selected as the best 6 ones in each environment. For example under control condition, genotypes NO 20, 8, 17, 9, 15 and 10 were the best ones with the highest grain yield and positive AMMI score.

****** Supplementary Genotype Analysis (Pacheco *et al*., 2005) *****
Do you want to perform the supplementary genotype(s) analysis(y/n)?
n
Note suppl genotype must be the last genotype in order.

Do you apply MINITAB release 15 (Y/N)?
n

***** Pairwise simple Pearson Correlation Coefficients *****
Between Environments

Correlations: Envi.1, Envi.2, Envi.3

	Envi.1	Envi.2
Envi.2	0.516	
	0.020	

Envi.3   0.881   0.663
          0.000   0.001

Cell Contents: Pearson correlation
               P-Value

Cluster Analysis of Observations outputs:
...
...
...
Principal component analysis outputs:
...
...
...

Results showed that variation attributable to genotypes (G), and their interaction (G×E), each was significant. Thus, decomposition of the GE effect into IPCA scores via AMMI method should be examined. The AMMI analysis showed that 2 IPCA scores accounted for almost all the variation relative to G × E as 91.86 and 8.14%, respectively. So the AMMI2 model is the best model for this data set. For this, the biplot of first and second scores against genotype means and environment means, and each other was drawn, which represented the separation of 20 genotypes and 3 environments in 2D space or other complicated ones (Fig. 3).

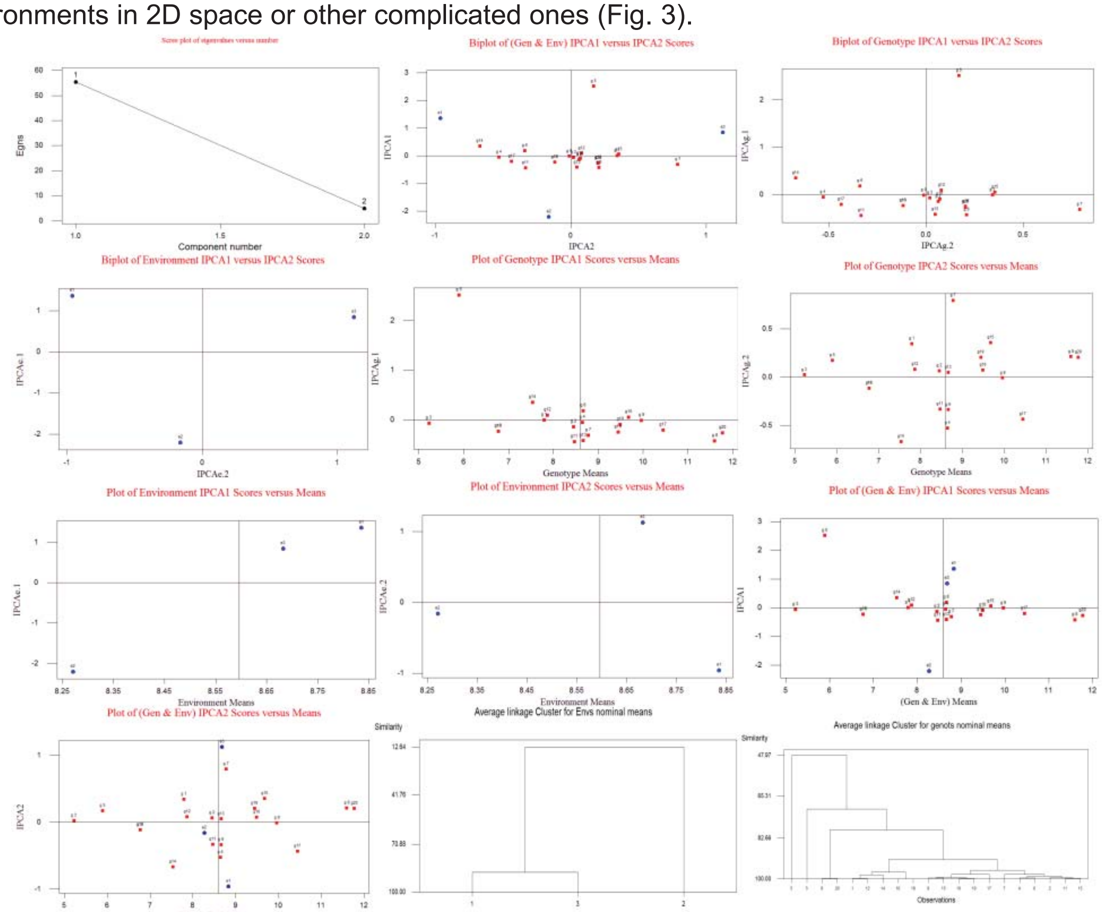


Fig. 3. Most of the graphs obtained by the AMMI macro in this experiment including two first components, biplots of IPCA scores, and IPCA scores against genotypes and environments mean.

As can be seen from the figure above, two of the IPCA scores were able to account for all (100%) of the variation in the GE interaction, based on which required biplots were plotted.

Accordingly, two IPCAs were drawn against each other and the distribution of genotypes and environments (under stress and control) was investigated. Accordingly, genotype 5 has the highest IPCA1 score, and genotypes such as 3, 9, 12, 13 and 16, which were located in the center of the biplot, possessed nearly zero interaction effect and had more general stability. Likewise, genotypes 7 and 14 had the highest and lowest IPCA2 scores, respectively. Genotypes 8 and 20 had the highest grain yield and showed good stability and can be considered as the most stable and high yielding corn genotypes in this research. Genotypes 3 and 5 had the least yield and were sensitive. The results of PCA and cluster analysis were consistent, so that the genotype 5 was located in a separate cluster (low grain yield and the highest IPCA1), and then the genotype 3 was located in another separate cluster (the lowest grain yield and IPCA was about zero, and then genotypes 8 and 20 were in the same group with highest grain yield and high IPCA1), and the remaining genotypes were in the same category, which had a moderate performance and stability. In the control conditions, the highest grain yield and then, respectively, in stress in the dough and milk stage, were the highest grain yield. When there is more than 2 IPCA scores or more than 2 environments, the AMMI macro produces IPCA1 vs IPCA2, IPCA1 vs IPCA3, IPCA2 vs IPCA3, Gen & Env means vs IPCA3 and IPCA4, a 3D surface plot, a contour plot of genotype IPCA1, IP CA2 & IPCA3 scores and moren (Fig. 4).

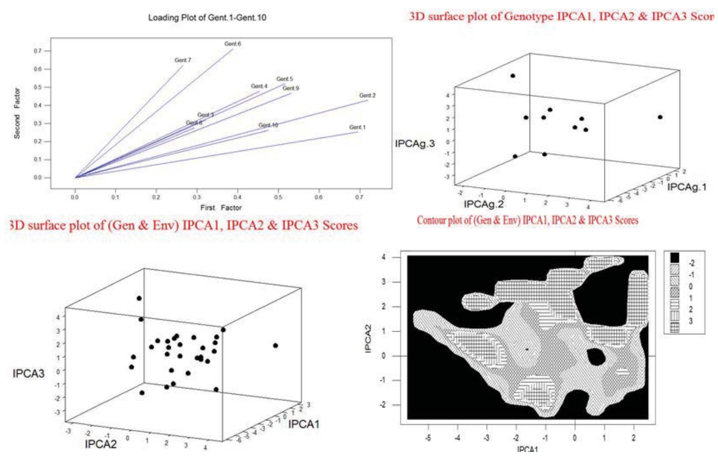


Fig. 4. Loading plot, 3-dimensional plot of first 3 IPCA scores of AMMI,

In this case, Oyekunle *et al.* (2017) studied the stability of 18 maize cultivars adapted to tropical environments using AMMI method, during 2 years, and have been able to evaluate the interaction of GE and introduce and select superior and stable cultivars in different regions, and also explore the relationship between regions. As in the present study, they showed that the main effect as well as the interaction effect of GE of grain yield were significant. In their research, Bipolt AMMI accounted for only 90.5% of the diversity of interactions and based on this, the distribution and separation of maize genotypes and environments were performed and the most stable and the high-yielding genotypes were identified. Likewise, Sserumaga *et al.* (2018) studied grain-yield stability of maize tropical hybrids under drought and optimum conditions in Africa and reported that locations, genotypes and their interaction were significant, and accordingly introduced superior drought stable and high-yielding cultivars. Such researches demonstrate the potential and value of AMMI strategy in surveying stability analysis in crops such as corn.

Another capability of our AMMI macro is regression fitting of genotypes via the Finlay-Wilkinson method versus environment means, whereby more stable genotypes in the experiment are identified. For example, the results of this approach for fitting Finlay-Wilkinson regression for genotype 1 versus environment means are as follows:

****** Pairwise correlation between gen & env means *****

***** Along with a Finlay_Wilkinson linear regression fit ****

Correlations: Gent.1, Emeans

Pearson correlation of Gent.1 and Emeans = 0.554
P-Value = 0.626

In the results below, the regression of each genotype with the environments mean is estimated as a stability schedule. For example for genotype 1, there was not a significant linear relationship, and a low *r-squared* and may other trends exist but was a high positive linear trend for genotypes 3 and 9.

Regression Analysis: Gent.1 versus Emeans
The regression equation is
Gent.1 = 0.8 + 0.81 Emeans

Predictor	Coef	SE Coef	T	P
Constant	0.81	10.49	0.08	0.951
Emeans	0.812	1.220	0.67	0.626

S = 0.5040 R-Sq = 30.7% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.1125	0.1125	0.44	0.626
Residual Error	1	0.2540	0.2540		
Total	2	0.3665			

Correlations: Gent.2, Emeans

Pearson correlation of Gent.2 and Emeans = 0.403
P-Value = 0.736

Regression Analysis: Gent.2 versus Emeans
The regression equation is
Gent.2 = 8.07 + 0.044 Emeans

Predictor	Coef	SE Coef	T	P
Constant	8.0726	0.8661	9.32	0.068
Emeans	0.0443	0.1007	0.44	0.736

S = 0.04162 R-Sq = 16.2% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.000335	0.000335	0.19	0.736
Residual Error	1	0.001732	0.001732		
Total	2	0.002067			

Correlations: Gent.3, Emeans

Pearson correlation of Gent.3 and Emeans = 1.000
P-Value = 0.015

Regression Analysis: Gent.3 versus Emeans

The regression equation is
Gent.3 = 0.731 + 0.522 Emeans

Predictor	Coef	SE Coef	T	P
Constant	0.7309	0.1031	7.09	0.089
Emeans	0.52234	0.01199	43.58	0.015

S = 0.004953 R-Sq = 99.9% R-Sq(adj) = 99.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.046583	0.046583	1899.22	0.015
Residual Error	1	0.000025	0.000025		
Total	2	0.046607			

Correlations: Gent.4, Emeans

Pearson correlation of Gent.4 and Emeans = 0.424
P-Value = 0.722

Regression Analysis: Gent.4 versus Emeans

The regression equation is
Gent.4 = 0.9 + 0.91 Emeans

Predictor	Coef	SE Coef	T	P
Constant	0.87	16.64	0.05	0.967
Emeans	0.905	1.935	0.47	0.722

S = 0.7996 R-Sq = 17.9% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.1399	0.1399	0.22	0.722
Residual Error	1	0.6394	0.6394		
Total	2	0.7793			

Correlations: Gent.5, Emeans

Pearson correlation of Gent.5 and Emeans = 0.987
P-Value = 0.102

Regression Analysis: Gent.5 versus Emeans
The regression equation is
Gent.5 = - 143 + 17.3 Emeans

Predictor	Coef	SE Coef	T	P
Constant	-143.20	23.99	-5.97	0.106
Emeans	17.344	2.789	6.22	0.102

S = 1.152 R-Sq = 97.5% R-Sq(adj) = 95.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	51.359	51.359	38.67	0.102
Residual Error	1	1.328	1.328		

Total	2	52.687

Correlations: Gent.6, Emeans

Pearson correlation of Gent.6 and Emeans = 0.911
P-Value = 0.270

Regression Analysis: Gent.6 versus Emeans
The regression equation is
Gent.6 = - 11.4 + 2.33 Emeans

Predictor	Coef	SE Coef	T	P
Constant	-11.408	9.079	-1.26	0.428
Emeans	2.335	1.056	2.21	0.270

S = 0.4362 R-Sq = 83.0% R-Sq(adj) = 66.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.9308	0.9308	4.89	0.270
Residual Error	1	0.1903	0.1903		
Total	2	1.1211			

Correlations: Gent.7, Emeans

Pearson correlation of Gent.7 and Emeans = -0.486
P-Value = 0.677

Regression Analysis: Gent.7 versus Emeans
The regression equation is
Gent.7 = 21.0 - 1.42 Emeans

Predictor	Coef	SE Coef	T	P
Constant	20.99	21.94	0.96	0.514
Emeans	-1.421	2.552	-0.56	0.677

S = 1.054 R-Sq = 23.7% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.345	0.345	0.31	0.677
Residual Error	1	1.112	1.112		
Total	2	1.456			

Correlations: Gent.8, Emeans

Pearson correlation of Gent.8 and Emeans = -0.982
P-Value = 0.122

Regression Analysis: Gent.8 versus Emeans
The regression equation is
Gent.8 = 28.1 - 1.91 Emeans

Predictor	Coef	SE Coef	T	P

Constant	28.053	3.202	8.76	0.072
Emeans	-1.9135	0.3723	-5.14	0.122

S = 0.1538 R-Sq = 96.4% R-Sq(adj) = 92.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.62516	0.62516	26.42	0.122
Residual Error	1	0.02366	0.02366		
Total	2	0.64882			

Correlations: Gent.9, Emeans

Pearson correlation of Gent.9 and Emeans = 0.998
P-Value = 0.040

Regression Analysis: Gent.9 versus Emeans

The regression equation is
Gent.9 = 2.38 + 0.882 Emeans

Predictor	Coef	SE Coef	T	P
Constant	2.3831	0.4792	4.97	0.126
Emeans	0.88205	0.05572	15.83	0.040

S = 0.02302 R-Sq = 99.6% R-Sq(adj) = 99.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.13283	0.13283	250.58	0.040
Residual Error	1	0.00053	0.00053		
Total	2	0.13336			

Correlations: Gent.10, Emeans

Pearson correlation of Gent.10 and Emeans = 0.897
P-Value = 0.292

Regression Analysis: Gent.10 versus Emeans
The regression equation is
Gent.10 = 6.52 + 0.346 Emeans

Predictor	Coef	SE Coef	T	P
Constant	6.517	1.470	4.43	0.141
Emeans	0.3460	0.1709	2.02	0.292

S = 0.07063 R-Sq = 80.4% R-Sq(adj) = 60.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.020441	0.020441	4.10	0.292
Residual Error	1	0.004989	0.004989		
Total	2	0.025430			

Correlations: Gent.11, Emeans

Pearson correlation of Gent.11 and Emeans = -0.744

P-Value = 0.466

Regression Analysis: Gent.11 versus Emeans
The regression equation is
Gent.11 = 23.6 - 1.76 Emeans

Predictor	Coef	SE Coef	T	P
Constant	23.56	13.56	1.74	0.333
Emeans	-1.756	1.577	-1.11	0.466

S = 0.6516   R-Sq = 55.3%   R-Sq(adj) = 10.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.5262	0.5262	1.24	0.466
Residual Error	1	0.4245	0.4245		
Total	2	0.9507			

Correlations: Gent.12, Emeans

Pearson correlation of Gent.12 and Emeans = 0.974
P-Value = 0.146

Regression Analysis: Gent.12 versus Emeans
The regression equation is
Gent.12 = - 5.48 + 1.55 Emeans

Predictor	Coef	SE Coef	T	P
Constant	-5.483	3.124	-1.76	0.330
Emeans	1.5531	0.3632	4.28	0.146

S = 0.1501   R-Sq = 94.8%   R-Sq(adj) = 89.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.41184	0.41184	18.28	0.146
Residual Error	1	0.02252	0.02252		
Total	2	0.43436			

Correlations: Gent.13, Emeans

Pearson correlation of Gent.13 and Emeans = -0.994
P-Value = 0.071

Regression Analysis: Gent.13 versus Emeans
The regression equation is
Gent.13 = 23.9 - 1.77 Emeans

Predictor	Coef	SE Coef	T	P
Constant	23.869	1.696	14.07	0.045
Emeans	-1.7689	0.1973	-8.97	0.071

S = 0.08151   R-Sq = 98.8%   R-Sq(adj) = 97.5%

Analysis of Variance

Source	DF	SS	MS	F	P

Regression	1	0.53422	0.53422	80.41	0.071
Residual Error	1	0.00664	0.00664		
Total	2	0.54087			

Correlations: Gent.14, Emeans

Pearson correlation of Gent.14 and Emeans = 0.865
P-Value = 0.335

Regression Analysis: Gent.14 versus Emeans
The regression equation is
Gent.14 = - 23.4 + 3.60 Emeans

Constant	-23.43	17.96	-1.30	0.416
Emeans	3.602	2.089	1.72	0.335

S = 0.8630 R-Sq = 74.8% R-Sq(adj) = 49.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2.2153	2.2153	2.97	0.335
Residual Error	1	0.7447	0.7447		
Total	2	2.9600			

Correlations: Gent.15, Emeans

Pearson correlation of Gent.15 and Emeans = 0.665
P-Value = 0.537

Regression Analysis: Gent.15 versus Emeans
The regression equation is
Gent.15 = - 0.3 + 1.17 Emeans

Predictor	Coef	SE Coef	T	P
Constant	-0.34	11.27	-0.03	0.981
Emeans	1.166	1.310	0.89	0.537

S = 0.5414 R-Sq = 44.2% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	**1**	**0.2321**	**0.2321**	**0.79**	**0.537**
Residual Error	1	0.2931	0.2931		
Total	2	0.5252			

Correlations: Gent.16, Emeans

Pearson correlation of Gent.16 and Emeans = -0.590
P-Value = 0.598

Regression Analysis: Gent.16 versus Emeans
The regression equation is
Gent.16 = 10.7 - 0.453 Emeans

Predictor	Coef	SE Coef	T	P
Constant	10.659	5.329	2.00	0.295
Emeans	-0.4527	0.6197	-0.73	0.598

S = 0.2560 R-Sq = 34.8% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.03498	0.03498	0.53	0.598
Residual Error	1	0.06556	0.06556		
Total	2	0.10054			

Correlations: Gent.17, Emeans

Pearson correlation of Gent.17 and Emeans = -0.094
P-Value = 0.940

Regression Analysis: Gent.17 versus Emeans
The regression equation is
Gent.17 = 11.9 - 0.16 Emeans

Predictor	Coef	SE Coef	T	P
Constant	11.86	14.98	0.79	0.573
Emeans	-0.165	1.741	-0.09	0.940

S = 0.7195 R-Sq = 0.9% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.0046	0.0046	0.01	0.940
Residual Error	1	0.5177	0.5177		
Total	2	0.5224			

Correlations: Gent.18, Emeans

Pearson correlation of Gent.18 and Emeans = -0.590
P-Value = 0.598

Regression Analysis: Gent.18 versus Emeans
The regression equation is
Gent.18 = 10.7 - 0.453 Emeans

Predictor	Coef	SE Coef	T	P
Constant	10.659	5.329	2.00	0.295
Emeans	-0.4527	0.6197	-0.73	0.598

S = 0.2560 R-Sq = 34.8% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.03498	0.03498	0.53	0.598
Residual Error	1	0.06556	0.06556		
Total	2	0.10054			

Correlations: Gent.19, Emeans

Pearson correlation of Gent.19 and Emeans = -0.817
P-Value = 0.391

Regression Analysis: Gent.19 versus Emeans
The regression equation is
Gent.19 = 15.7 - 0.726 Emeans

Predictor	Coef	SE Coef	T	P
Constant	15.688	4.402	3.56	0.174
Emeans	-0.7262	0.5119	-1.42	0.391

S = 0.2115 R-Sq = 66.8% R-Sq(adj) = 33.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.09005	0.09005	2.01	0.391
Residual Error	1	0.04474	0.04474		
Total	2	0.13479			

Correlations: Gent.20, Emeans

Pearson correlation of Gent.20 and Emeans = -0.869
P-Value = 0.330

Regression Analysis: Gent.20 versus Emeans
The regression equation is
Gent.20 = 19.1 - 0.856 Emeans

Predictor	Coef	SE Coef	T	P
Constant	19.139	4.200	4.56	0.138
Emeans	-0.8563	0.4883	-1.75	0.330

S = 0.2018 R-Sq = 75.5% R-Sq(adj) = 50.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.12520	0.12520	3.08	0.330
Residual Error	1	0.04072	0.04072		
Total	2	0.16592			

Conclusions

We present in this paper some macros for conducting statistical analyses for different experimental designs using MINITAB, which is a user-friendly platform. The macros introduced here allow users to do many analyses that may not be possible with other platforms. We hope all researchers, especially agriculturists and biologists, would benefit from these programs. These macros are listed as follows: (1) a macro for analyzing the Latin square designs with subsampling: This applies when there is more than one sampling in a replication of a Latin square design and is rarely done due to the difficult analysis process. (2) a macro for analyzing the relative efficiency of an experimental design, which is suitable to compare different designs to increase the efficiency of the experiments. (3) a macro performing orthogonal contrasts that are used when contrasts or group comparisons are done following an ANOVA. Our macro does this task better than any other macro where it estimates the contrasts coefficients automatically. (4) macro for curve fitting when nonlinear trends are used other than linear trends when a cause and effect is tested. (5) a macro called MULTPATH to do path coefficient analysis, along with estimating and testing the significance of direct effects and also related standard errors. (6) some macros to analyze univariate or classic parametric stability analysis of genotypes. (7) a macro called AMMI to do multivariate parametric stability analysis (AMMI) which is a useful means to study the stability analysis in plant breeding programs. All the achievements in this paper on the stability analysis of the corn plant have led to the identification of the best and high-yielding relatively new corn genotypes that can be used in areas with problems such as drought stress. It should be noted that all of these results are easy to accomplish in Windows XP and 7 with MINITAB 13. The macros presented in this paper could be easily obtained by emailing the first author of this paper.

Acknowledgements

The first author expresses his deepest thanks to professor Manjit S. Kang, Department of Plant Pathology, Kansas State University, Manhattan, Kansas, USA for reviewing this manuscript and also to professor Hugh G. Gauch, Jr., Cornell University, Ithaca, the USA for his nice help and comments while preparing the AMMI macro.

References

Arminian A., Kang M.S., Kozak M., Houshmand S., and Mathews P., 2008. MULTPATH: A Comprehensive MINITAB program for computing path coefficients and multiple regression for multivariate analyses. Journal of Crop Improvement 22, 82-120.

Becker HC, Léon J, 1988. Stability analysis in plant breeding. Plant Breeding 101: 1-23. Dorst JC, 1957: Adaptation. Euphytica 6: 247-254.

Board, J.E., M.S. Kang, and B.G. Harville., 1997. Path analysis identifies indirect selection criteria for yield of late-planted soybean. Crop Science 37, 879–884.

Board, J.E., M.S. Kang, and M.L. Bodrero., 2003. Yield components as indirect criteria for late-planted soybean cultivars. Agronomy Journal 95, 420–429.

Carey, G. 1998a. Path analysis using Proc CALIS. Accessed August 2 2014 at http://psych.colorado.edu/∼carey/courses/psyc7291/handouts/pathcalis.pdf

Carey, G. 1998b. Multiple regression and path analysis. Colorado State University. Available on: http://ibgwww.colorado.edu/~carey/p7291dir/handouts/pathanal2.pdf

Cochran, W.G. and Cox, GM. 1957. Experimental Designs. John Wiley and Sons, Inc. 460 p.

Cornelius, P.L., 1993. Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. Crop Science 33, 1186-1193.

Cramer, C.S., Wehner, T.C., and Donaghy, S.B. 1999. PATHSAS: A SAS computer program for path coefficient analysis of quantitative data.Journal of Heredity, 90(1): 260-262.

Cristina, M. and Nogueira, S. 2004. Orthogonal contrasts: definitions and concepts. Sci. Agric. (Piracicaba, Braz.), v.61, n.1, p.118-124.

Dewey, D.R. and K.H. Lu, 1959. A correlation and path coefficient analysis of components of crested wheatgrass and seed production. Agron. J. 51: 515-518.

Eberhart, S.A. and W.A. Russell. 1966. Stability parameters for comparing varieties, Crop Science 6, pp. 36–40.

Finlay, K.W., Wilkinson, G.N., 1963. The analysis of adaptation in a plant breeding programme. Australian Journal of Agricultural Research 14, 742–754.

Francis, T.R. and L.W. Kannenberg, 1978. Yield stability studies in short-season maize. I. A descriptive method for grouping genotypes. Canadian Journal of Plant Sciences, 58: 1029-1034.

Freeman, G.H. 1973. Statistical methods for the analysis of Genotype-environment interactions. Heredity, 31(3): 339-354.

Gauch, H.G. Jr., 1992. Statistical Analysis of Regional Trials: AMMI Analysis of Factorial Design. Elsevier, Amsterdam.

Gollob, H.F., 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33, 73-155.

Gomez, K.A. and A.A. Gomez. 1984. Statistical procedures for agricultural research. 2nd Edn. John Wiley and Sons. 704 p.

Hanson WD (1970).Genotypic stability. Theor. Appl. Genet. 40:226-231.

Hinkelmann, K. and Kempthorne, O. 2008. Design and analysis of Experiments. Volume 1. 2nd Edn. John Wiley and Sons, Inc. 666 p.

Hussein MA, Bjornstad A, Aastveit AH (2000). SASG × ES-TAB, A SAS program for computing genotype x environment stability statistics. Agron J 92:454-459.

Johnson, R.A. and Bhattacharyya, G.K. 2010. Statistics: Principles and Methods (6th Edn). John Wiley and Sons Inc. p 686.

Kang M.S. (1998). Using genotype-by-environment interaction for crop cultivar development. Adv. Agron. 62:199-252.

Kang M.S. (2015). Efficient SAS programs for computing path coefficients and index weights for selection indices. J. Crop Improv. 29:6-22.

Kang MS (1993). Simultaneous selection for yield and stability in crop performance trials: consequences for growers. Agron J. 85:754-757.

Kang, 2015. Efficient SAS programs for computing path coefficients and index weights for selection indices. Journal of Crop Improvement, 29:6–22.

Kang, M.S., and Gauch, H.G. Jr. 1996. Genotype-by-environment interaction. CRC Press. 409 p.

Kang, M.S., J.D. Miller, and P.Y.P. Tai., 1983. Genetic and phenotypic path analyses and heritability in sugarcane. Crop Science 23, 643–647.

Karafilowska, E. 2004. MINITAB: the Information Division, Teaching, Learning and Research Support, University of Melbourne, Australia.

Kenny, D.A. 2004. Correlation and Causality (Rev. ed.). Retrieved August 9, 2005, from http://davidakenny.net/doc/cc_v1.pdf

Lawal, B. 2014. Applied Statistical Methods in Agriculture, Health and Life Sciences. Springer, 799 p.

Li, C.C. 1975. Path analysis-A primer. Boxwood Press, Pacific Grove, Calif.

Lin, C.S. and M.R. Binns, 1988. A superiority measure of cultivar performance for cultivar x location data. Canadian Journal of Plant Science, 68: 193-198.

Lin, C.S., M.R. Binns and L.P. Lefkovitch, 1986. Stability Analysis: Where do we stand? Crop Science, 26: 894-899.

Lohmor, N., Khan, M., Kapoor, K., Kumar Tripathi, R. 2017. Studies on the relative efficiency of different experimental designs for sunflower (Helianthus annuus L.). Annual Research & Review in Biology, 12(4): 1-7.

Mathews, P.G., 2005. Design of experiments with MINITAB. ASQ (American Society for Quality) Press, Milwaukee, WI.

Milligan, S.B., K.A. Gravois, K.P. Bischoff and F.A. Martin, 1990. Crop effects on genetic relationships among sugarcane traits. Crop Sci.30: 927-931.

MINITAB., 2018. MINITAB statistical software Inc.

Montgomery, D.C., Peck, E.A. and Vining, G.G. 2012. Introduction to Linear Regression Analysis. 5th Edition,

Neale M.C. and Cardon L.R. (1992) Methodology for Genetic Studies of Twins and Families (Kluver, Dordrecht, The Netherlands). New York, John Wiley & Sons.

Nielsen, R.L. 2018. Grain fill stages in corn. Agronomy Dept., Purdue Univ. West Lafayette, IN 47907-2054. URL: http://www.king.corn/news/timeless/GrainFill.html

Oladugba, A.V., Onuoha, D.O., Opara P.N.2013. Relative Efficiency of Split-plot Design (SPD) to Randomized Complete Block Design (RCBD). West African Journal of Industrial and Academic Research Vol.7 No. 1: 113-120.

Olive, D.J. 2017. Linear Regression. Springer International Publishing AG, USA. 497 p.

Oyekunle, M., Menkir, A., Mani, H., Olaoye, G., Usman, I.S., Ado, S.G., Abdullahi, U.S., Ahmed, H.O., Hassan, L.B., Abdulmalik, R.O. and Abubakar, H. 2017. Stability Analysis of Maize Cultivars Adapted to Tropical Environments Using AMMI Analysis. Cereal Research Communications, 45(2): 336–345.

Pacheco, R.M., Duarte,J.B., Vencovsky,R., Pinheiro,J.B., Oliveira, A.B., 2005. Use of supplementary genotypes in AMMI analysis. Theoretical Applied Genetics 110, 812–818.

Pedhazur, E.J. 1997. Multiple regression in behavioral research. Explanation and prediction. 3rd edn. Thomson Learning, Inc. Thomson Learning1M. 1058 p.

Pinthus JM (1973). Estimate of genotype value: a proposed method. Euphytica 22:121-123.

Plaisted, R.L. and L.C., Peterson. 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. American Potato Journal, 36: 381-385.

Ransom, J. 2013. Corn, growth and management, quick guide. A1173. Cereal Crops NDSU Extension Service. NDSU extension service, North Dakota State University Fargo, North Dakota, USA.

Richardson, J.T.E. 2018a. The use of Latin-square designs in educational and psychological research. Educational Research Review, 24 pp. 84–97.

Richardson, J.T.E. 2018b. Latin Square Design. In: Frey, Bruce B. ed. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. SAGE Publications, p. 949.

Rodrigues, P.C., Monteiro, A., and V.M. Lourenço. 2015. A robust AMMI model for the analysis of genotype-by-environment data. Bioinformatics, 32(1): 58–66.

Roemer T. 1917: Sind die ertragreichen Sorten ertragssicherer? Mitteilungen der Deutschen Landwirtschaftlichen Gesellschaft 32: 87-89.

Seber, G.A.F. and Wild, C.J. 2003. Nonlinear regression. John Wiley and Sons Inc., Publication. 800 p.

Shukla, G.K., 1972. Some statistical aspects of partitioning genotype environmental components of variability. Heredity, 29: 237-245.

Sidwell, R.J., E.L. Smith, and R.W. McNew. 1976. Inheritance and interrelationships of grain yield and selected yield-related traits in a hard red winter wheat crops. Crop Science 16, 650–654.

SserumagaA, J.P., Yoseph, B., Pillay, K., Kullaya, A., Oikeh, S.O., Mugo, S., Machida, L., Ngolinda, I., Asea, G., Ringo, J., Otim, M., Abalo, G. and Kiula, B. 2018. Grain-yield stability among tropical maize hybrids derived from doubled-haploid inbred lines under random drought stress and optimum moisture conditions. Crop and Pasture Science, 69: 691-702.

Sokal, R.R., and Rohlf, F.J. 1995. Biometry; the principles and practice of statistics in biological research. W.H. Freeman and Company, 889 p.

WRICKE, G. 1962.Uber eine Methode zur Erfassung der okologischen Streubreite in Feldversuchen.-2.Pflunrenzzichr.47:92-96.

WRICKE, G. 1965. Die Erfassung der Wechselwirkung zwischcn Genotyp und Umwelt bei quantitativen Eigenschaften. - 2. Pflanzenziicht. 53: 266-343.

Wricke, G. 1962. Uber eine method zur Erfassung der Okologischen streuberite in Feldversuchen. Z. Pflanzerzuentz, 47: 92-96. (in German).

WRICKE, G. 1964. Zur Berechnung der Okovalenz bei Somrner-weizen und Hafer.2.Pflanzenziicht.52:127-138.

Wright, S. 1934. The method of path coefficients. Annals of Mathematical Statistics 5(3), 161–215.

Wuensch, K.L. 2016. An introduction to path analysis. Retrieved from http://core.ecu.edu/psyc/wuenschk/MV/SEM/Path.pdf

Zar, J.H. 2014. Biostatistical analysis (5th edn.). Pearson Education Limited. 756 p.

DECISION SCIENCES INSTITUTE

Predicting Backorders using Machine Learning Techniques

Saban Adana
University of North Texas
E-mail: saban.adana@unt.edu

Sedat Cevikparmak
University of North Texas
E-mail: sedat.cevikparmak@unt.edu

Hasan Celik
University of North Texas
E-mail: hasan.celik@unt.edu

Dr. Hasan Uvet
Georgia Gwinnett College
E-mail: huvet@ggc.edu

ABSTRACT

This paper attempts to demonstrate the utility of machine learning in an inventory case with the aim of demonstrating empirically that machine learning can be an important tool to optimally tackle supply chain problems. More specifically, we examine the usefulness of advanced machine learning techniques, including neural networks, auto neural networks, and decision trees to forecasting backorders happening in a supply chain. Following, we compare these methods with a more traditional one, logistic regression and parse out the differences. We conclude that the use of machine learning techniques for forecasting backorders provide more precise forecasts than traditional forecasting techniques such as logistic regression. By implementing our auto neural model, businesses can more accurately predict the backorder status of a product, in turn, minimize cost by more efficient inventory control, and increase profits by higher customer satisfaction. Future research could be directed to investigating other supply chain phenomena such as transportation management, demand forecasting, inventory levels, cash-to-cash cycles, supplier quality and production planning using machine-learning techniques to increase the robustness and generalizability of this approach

KEYWORDS: backorder, machine learning, inventory management, forecasting

INTRODUCTION

Customer satisfaction has always been at the center of marketing and operations management to grow market share and gain competitive advantages (Anderson et al. 1994). Therefore, supply chain managers have focused their efforts in better managing product inventory and preventing stock outs to improve customer satisfaction and retention. However, stock outs still occur causing backorders. Backorder can be defined as “a customer order that cannot be fulfilled, and for which the customer is prepared to wait for some time ” (Business dictionary).

The level of products that go into backorders and the days they remain are significant indicators of the firm`s customer service and efficiency in managing its inventory management. If backorders are not dealt with properly, it will have a significant impact on the firms's revenue, market share, credibility and customers' trusts and may end up with a lost sale. Disappointing customers will indirectly drive the customer into the hands of a competitor.

On the other hand, the appropriate actions to avoid backorders would put extra pressure on different stages of the supply chain, which will increase the overall costs as well as the associated transportation costs (Carter and Rogers, 2008). Moreover, the globalization has created an environment where customers' demands and expectations are highly probabilistic, which made traditional supply chain management systems less effective (Simchi-Levi et al., 2008). This results in errors in demand forecasting and backorders. Accordingly, many companies started to utilize algorithms to predict the backorder status of products to avoid the tangible and intangible costs. Reliable inventory systems are also key to avoid and handle backorders. Optimizing supply chain means providing customer what they want and when they want. If that does not happen, this will leave a `bad taste` in the customer's mouth, which will hurt the company's revenue.

With the easiness of accessing low cost, high computing power machines, machine learning and big data has started to alter the analytical processes of business decision making. A highly studied topic in supply chain management is the issue of inventory management, which is a crucial activity for any business as they try to determine the decision of when, and how much to order, in a highly dynamic market. Many of the approaches so far viewed the issue primarily as an optimization problem. Identifying the supplies at risk of backorder before the event would be another way to address the problem allowing the firms to respond to the issue (Vijayan,2008).
In the literature, the main antecedents discussed of backorders are current inventory level of products, lead time, in transit quantity, forecasted sales, actual sales (demand), minimum recommended amount in stock and identified issue of products (e.g. damage due to packaging or deck problems) (Koh, 2004, Chandra&Grabis, 2008, Koh, Saad, & Jones, 2002). Through this analysis, we may be able to answer which variables play a bigger role in backorder status of a product. Accordingly, companies can predict the backorder status of a product and align their supply chains, which in turn, will minimize the inventory costs and improve customer satisfaction.

This paper attempts to demonstrate the utility of machine learning in an inventory case with the aim of providing insight on the predictive powers of various techniques and demonstrate empirically that machine learning can be an important tool to optimally tackle supply chain problems.

In this study, we examine the usefulness of advanced machine learning techniques, including neural networks, auto neural networks, and decision trees to forecasting backorders happening in a supply chain. Following, we compare these methods with a more traditional one, logistic regression and parse out the differences.

LITERATURE REVIEW

In a general sense, the demand during a stock out is either lost or completely backordered. Although, there might be customers who are willing to wait, the majority would place more emphasis on time utility resulting a lost sale case. There are several models in the literature combine backorders and lost sales. The first attempt to this issue was made by Montgomery et al. (1973). Quyang and Wu (1999) followed with a model that incorporated varying lead-time with a fixed reorder point. Hariga and Ben Daya (1999) combined backorders and lost sales in scenarios of partial and full demand information. Abad (2000) increased the complexity of the model by including exponential decay in a perishable good context.

All the aforementioned papers brought forward mathematical models to predict backorder or lost status and through that what should be the inventory level. Our paper purely approaches the issue from a machine learning perspective. The major contribution of the paper is assessment of different learning classifiers algorithms such as neural networks and decision trees, based on specific methods to deal with the class imbalanced problem in a real supply chain scenario.

THEORETICAL DEVELOPMENT/MODEL

Data Description

The original data includes 22 columns and 1.048.576 rows primarily interval and some binary data. Explanations of the variables are presented in the Table 1.

Variable Name	Explanation
Went_on_backorder_Yes	If the product went to backorder or not
Deck_Risk_yes	The products that might remain in the deck/shop/stock
Forecast 3 month	The demand estimated for 3 months
Forecast 6 month	The demand estimated for 6 months
Forecast 9 month	The demand estimated for 9 months
In_transit_quantity	The total number of products that are currently being shipped from one location to another
Lead_time	The time taken from release of an order to production and shipment
Local_bo_quantity	Amount of overdue stock orders
Min_bank	Minimum recommended amount in stock.
National_inv	Current inventory level of different products
Oe_constraint	Products that are facing operational limiting factors such as bottleneck

Sales_1_month	Actual sales for 1 month
Sales_3_month	Actual sales for 3 month
Sales_6_month	Actual sales for 6 month
Sales_9_month	Actual sales for 9 month
Potential_Issue_Yes	Potential damage
Pieces_past_due	Products overdue from source
Perf_6_month	Source average performance in last 6 months
Perf_12_month	Source average performance in last 12 months
Ppap_risk_yes	Risks associated with packaging and production
Rev_stop_yes	Revenue status for product
Stop_autobuy_yes	Whether automatic selling process has been stopped or not

Table 1: Variable Descriptions

Data Preparation

In the dataset, since our target variable (Went_on_backorder_Yes) has only 9235 out of 1.048.576 observations, which consists of less than %1 of the data, we used stratified hold-out method and randomly undersampled the frequent event (Went_on_backorder_No) to have a balanced sample. That is, we created a new sample with 18470 observations, of which our target variable consisted 50%.

We have used the impute data node in SAS Enterprise Miner to impute missing values which consisted of 5.98% of the sample. The method it has used was taking the mean of the columns and replacing them with a value. We used this technique not to distort the current distribution of data. Moreover, since the skewness and Kurtosis values were more than 2 and 7, respectively, which is a sign of violation of normality assumption, we transformed the data with log10 method to meet this assumption. The descriptive statistics of the data is presented in Table 2.

Variable	Role	Mean	Standard Deviation	Non Missing	Missing	Minimum	Median	Maximum	Skewness	Kurtosis
LG10_forecast_3_month	INPUT	0.839612	0.914994	18470	0	0	0.69897	5.079185	0.908258	0.072592
LG10_forecast_6_month	INPUT	1.000099	0.988346	18470	0	0	0.90309	5.232694	0.719043	-0.2552
LG10_forecast_9_month	INPUT	1.095721	1.034089	18470	0	0	1	5.316308	0.627752	-0.3973
LG10_in_transit_qty	INPUT	0.193565	0.52305	18470	0	0	0	4.438242	3.129162	10.14996
LG10_lead_time	INPUT	0.830488	0.263552	18470	0	0	0.954243	1.724276	-0.31288	0.448377
LG10_local_bo_qty	INPUT	0.058441	0.270611	18470	0	0	0	3.477121	5.891591	39.94688
LG10_min_bank	INPUT	0.478738	0.70688	18470	0	0	0	4.08934	1.610285	1.904321
LG10_national_inv	INPUT	3.489749	0.073725	18470	0	0	3.477844	5.121494	-4.78442	621.4353
LG10_perf_12_month_avg	INPUT	1.979991	0.21438	18470	0	0	2.003461	2.004321	-9.12835	81.33788
LG10_perf_6_month_avg	INPUT	1.972962	0.244348	18470	0	0	2.003547	2.004321	-7.95122	61.23004
LG10_pieces_past_due	INPUT	0.069206	0.307748	18470	0	0	0	3.699057	5.134822	28.55107
LG10_sales_1_month	INPUT	0.568956	0.685353	18470	0	0	0.30103	3.982678	1.360584	1.573597
LG10_sales_3_month	INPUT	0.848885	0.832828	18470	0	0	0.69897	4.467282	0.919936	0.368058
LG10_sales_6_month	INPUT	1.027059	0.914921	18470	0	0	0.954243	4.819511	0.716242	-0.03745
LG10_sales_9_month	INPUT	1.14081	0.96619	18470	0	0	1.079181	4.994163	0.614755	-0.19504

Table 2: Descriptive Statistics of the Data

Lastly; we divided our data set for training and validation models as 60% and 40% respectively (See partition result summary in the Table 3).

Type	Data Set	Number of Observations
DATA	EMWS1.Trans_TRAIN	18470
TRAIN	EMWS1.Part_TRAIN	11082
VALIDATE	EMWS1.Part_VALIDATE	7388

Table 3: Partition Summary

Data Mining Techniques

Since the target variable that we are interested is a binary variable (backorder status: Yes or No), we decided that it would be first appropriate to conduct a logistic regression analysis, in which we used stepwise approach to find the best model, with all the input variables. Afterwards, we used decision tree, neural network and autoneural methods. Moreover, we conducted a principal component analysis (PCA) and then repeated aforementioned methods to find the best model to predict our target variable. Finally, we compared the models to find the best one according to “misclassification” and “average square error (ASE)” assessment. We used SAS Enterprise Miner for all the models and model comparison. The diagram created for aforementioned models and data preparation is presented in the Figure 1 below.

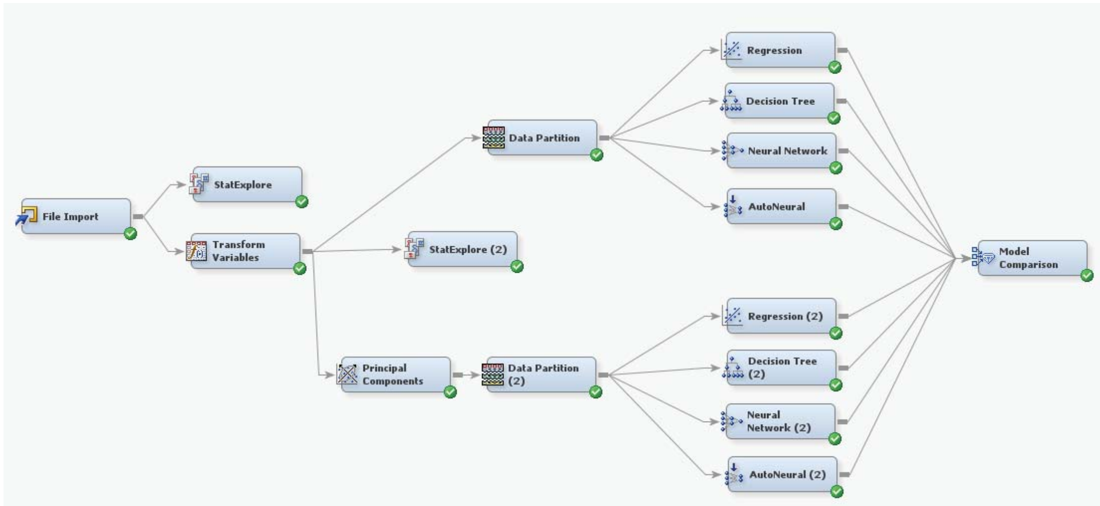


Figure 1: SAS Enterprise Miner Diagram.

Findings

a. Logistic Regression_1:

According to the logistic regression outputs provided in Table 4 below, we observed that the overall model including 21 predictor variables is statistically significant (p <.0001).

Likelihood Ratio Test for Global Null Hypothesis: BETA=0

-2 Log Likelihood		Likelihood		
Intercept Only	Intercept & Covariates	Ratio Chi-Square	DF	Pr > ChiSq
15362.914	9966.767	5396.1474	12	<.0001

Table 4: Overall Model Evaluation Results.

Of the 21 predictor variables 12 were significant: forecast_3 month, forecast_6 month, in_transit_qty, lead-time, local_bo_qty, min_bank, national_inv, pieces_past_due, sales_3_month, deck_risk_yes, ppap_risk_yes, and stop_auto_buy_yes. Furthermore, exp (b) values (odd ratios) of significant variables show that with the increase in those variables that are over 1.0 (forecast_3 month, forecast_6 month, local_bo_qty, pieces_past_due, sales_3_month, deck_risk_yes, stop_auto_buyYes), it would be more likely to face a backorder status (See the results in Table 5).

Analysis of Maximum Likelihood Estimates

Parameter		went_on_ backorder Yes	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept		1	1	164.2	11.5899	200.79	<.0001		999.000
LG10_forecast_3_month		1	1	1.1247	0.0978	132.27	<.0001	0.5599	3.079
LG10_forecast_6_month		1	1	0.6307	0.0932	45.83	<.0001	0.3397	1.879
LG10_in_transit_qty		1	1	-1.7850	0.0824	469.47	<.0001	-0.5038	0.168
LG10_lead_time		1	1	-1.0836	0.1025	111.86	<.0001	-0.1572	0.338
LG10_local_bo_qty		1	1	2.0909	0.2585	65.41	<.0001	0.3031	8.092
LG10_min_bank		1	1	-1.3834	0.0742	347.80	<.0001	-0.5312	0.251
LG10_national_inv		1	1	-47.2028	3.3324	200.65	<.0001	-1.7032	0.000
LG10_pieces_past_due		1	1	1.1541	0.1605	51.72	<.0001	0.1971	3.171
LG10_sales_3_month		1	1	0.8509	0.0612	193.23	<.0001	0.3873	2.342
deck_riskYes	0	1	1	0.0977	0.0333	8.64	0.0033		1.103
ppap_riskYes	0	1	1	-0.1499	0.0367	16.72	<.0001		0.861
stop_auto_buyYes	0	1	1	0.2379	0.0921	6.67	0.0098		1.269

Table 5: Results of Logistic Regression_1 Analysis.

The classification table (See Table 6) provides the required information to calculate the model accuracy (proportion of correct predictions), sensitivity (proportion of positive cases correctly classified as positive) and specificity (proportion of negative cases correctly classified as negative) and precision (proportion of predicted positive cases that were correctly classified as positive). This model has 81.71% accuracy, 80.5% sensitivity, 82.91% specificity and 82.49% precision based on validation model.

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	80.7665	83.2882	4615	41.6441
1	0	19.2335	19.8340	1099	9.9170
0	1	17.2504	16.7118	926	8.3559
1	1	82.7496	80.1660	4442	40.0830

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	80.9675	82.9182	3063	41.4591
1	0	19.0325	19.4911	720	9.7455
0	1	17.5035	17.0818	631	8.5409
1	1	82.4965	80.5089	2974	40.2545

Table 6: Classification Results of Logistic Regression_1 Model.

Moreover, the cumulative lift of the model is presented in the Figure 2.

Figure 2: Cumulative Lift of the Logistic Regression_1 Model.

b. Decision Tree_1:

In the Decision Tree_1 model, we specified the assessment measure as misclassification with 2 max branch and 6 max depth. The decision tree elicited is presented in Figure 3 below.

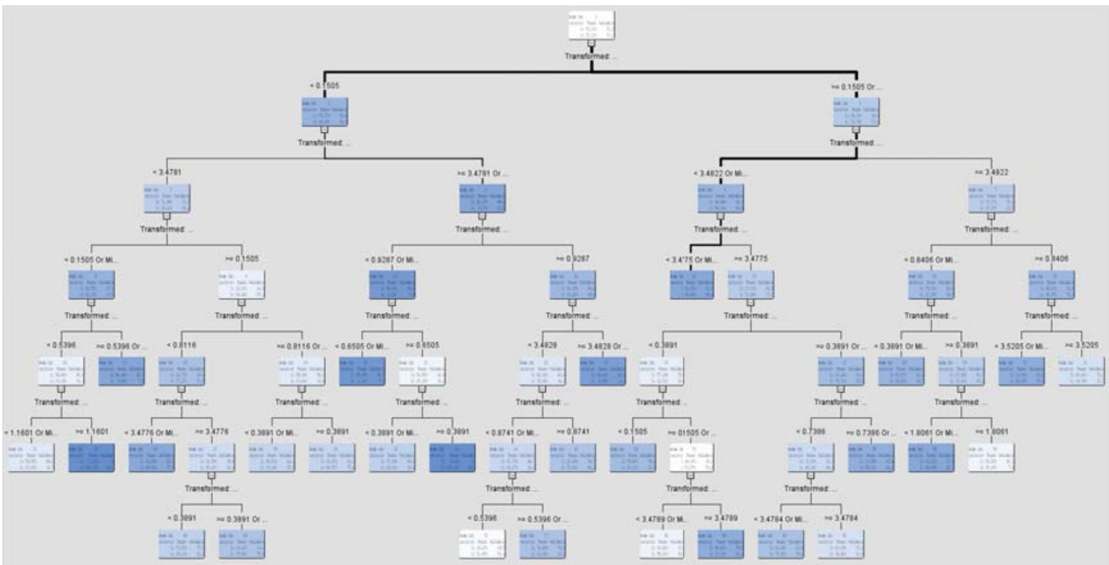


Figure 3: Decision Tree_1

The variables are listed based on their importance to classify the target variable in Table 7.

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
LG10_forecast_3_month	Transformed: forecast_3_month	1	1.0000	1.0000	1.0000
LG10_national_inv	Transformed: national_inv	8	0.6716	0.7166	1.0669
LG10_sales_1_month	Transformed: sales_1_month	5	0.2869	0.2627	0.9159
LG10_sales_3_month	Transformed: sales_3_month	2	0.2668	0.2873	1.0769
LG10_pieces_past_due	Transformed: pieces_past_due	1	0.2152	0.1936	0.8995
LG10_lead_time	Transformed: lead_time	3	0.2123	0.1693	0.7973
LG10_sales_6_month	Transformed: sales_6_month	2	0.1514	0.1410	0.9310
LG10_local_bo_qty	Transformed: local_bo_qty	1	0.1374	0.0233	0.1695
LG10_forecast_6_month	Transformed: forecast_6_month	2	0.0873	0.0883	1.0123
LG10_in_transit_qty	Transformed: in_transit_qty	1	0.0538	0.0561	1.0427

Table 7: Variable Importance of Decision Tree_1.

This model has a 85.92% accuracy, 87.73% sensitivity, 84.10% specificity and 84.66% precision based on validation model (See classification results in Table 8).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	87.6620	84.2447	4668	42.1224
1	0	12.3380	11.8571	657	5.9285
0	1	15.1641	15.7553	873	7.8776
1	1	84.8359	88.1429	4884	44.0715

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	87.2753	84.1094	3107	42.0547
1	0	12.7247	12.2631	453	6.1316
0	1	15.3344	15.8906	587	7.9453
1	1	84.6656	87.7369	3241	43.8684

Table 8: Classification Results of the Decision Tree_1 Model.

Moreover, the cumulative lift of the model is presented in the Figure 4.

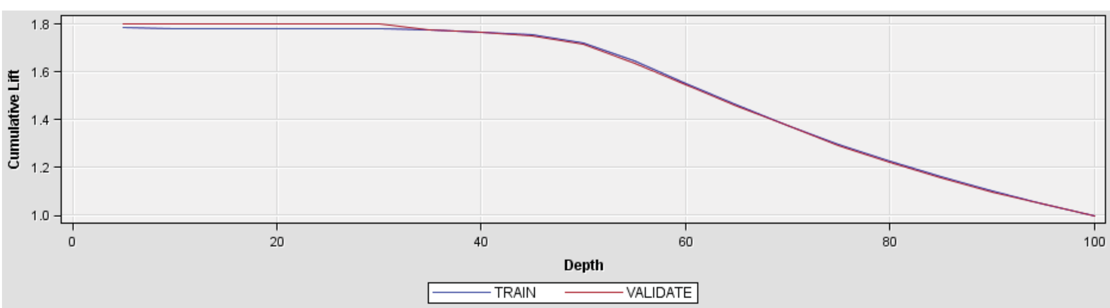


Figure 4: Cumulative Lift of the Decision Tree_1 Model.

c. Neural Network_1

Neural Networks (NN) are nonparametric models originated from the studies of the brain and nervous system. They are particularly useful in forecasting and modeling due to their data driven nature. NNs respond better to data that includes misspecification, linearity violation, outliers and re-estimation (Hill et al., 1996).

This model has an 85.77% accuracy, 87.46% sensitivity, 84.08% specificity and 84.60% precision based on validation model (See Table 9).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	87.2870	84.1364	4662	42.0682
1	0	12.7130	12.2541	679	6.1271
0	1	15.3109	15.8636	879	7.9318
1	1	84.6891	87.7459	4862	43.8729

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	87.0272	84.0823	3106	42.0411
1	0	12.9728	12.5338	463	6.2669
0	1	15.3967	15.9177	588	7.9589
1	1	84.6033	87.4662	3231	43.7331

Table 9: Classification Results of the Neural Network_1 Model.

Moreover, the cumulative lift of the model presented in the Figure 5.

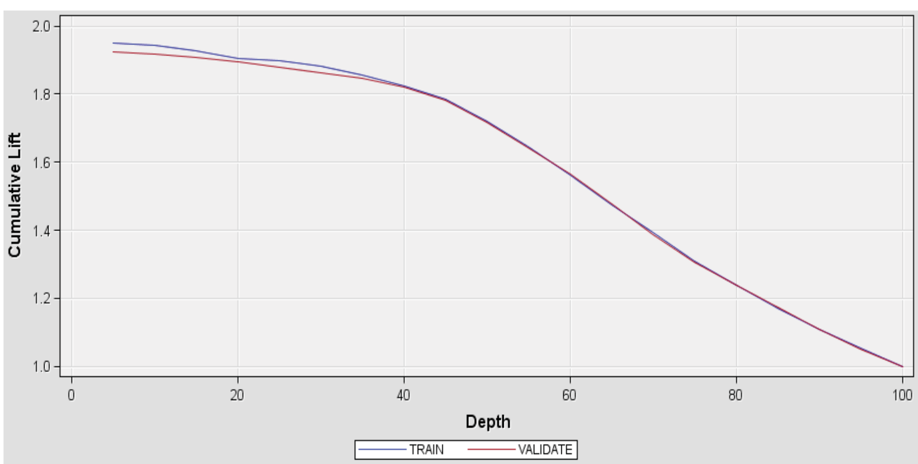


Figure 5: Cumulative Lift of the Neural Network_1 Model.

d. Autoneural_1:

This model has an 87.22% accuracy, 88.49% sensitivity, 85.95% specificity and 86.29% precision based on validation model (See Table 10).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	88.3652	85.9412	4762	42.9706
1	0	11.6348	11.3156	627	5.6578
0	1	13.6835	14.0588	779	7.0294
1	1	86.3165	88.6844	4914	44.3422

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	88.1944	85.9502	3175	42.9751
1	0	11.8056	11.5051	425	5.7526
0	1	13.7012	14.0498	519	7.0249
1	1	86.2988	88.4949	3269	44.2474

Table 10: Classification Results of the Autoneural_1 Model.

Moreover, the cumulative lift of the model presented in the Figure 6.

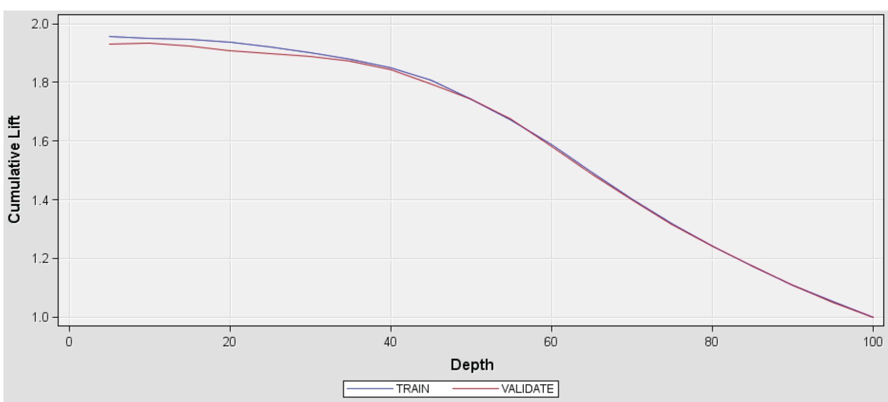


Figure 6: Cumulative Lift of the Autoneural_1 Model

e. Principal Component Analysis (PCA):

As we had a large number of input variables, we wanted to see whether a lower number of components explained a significant portion of the variance. Therefore, we have conducted PCA analysis in SAS miner, came up with 16 PCs, and used these components in our following models to improve the predictive power. We almost did not lose any variance by this reduction (99.2% variance explained). The only downside of this approach is that we do not specifically know the individual effect of each variables and composition of the components. (See the summary of the PCA in Table 11 and Figure 7).

```
Total number of input variables: 21
Maximum number cutoff of principal components: 20
Cumulative proportional eigenvalue cutoff: 0.99
Proportional eigenvalue increment cutoff: 0.001
Number of the selected principal components: 16
Total variation explained by the selected principal components: 0.992772248
```

Table 11: Summary of the PCA

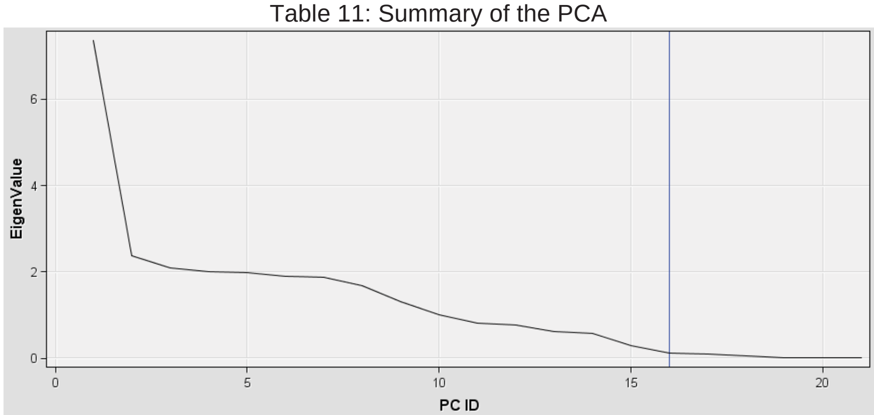


Figure 7: Suggested Number of Principal Components

f. Logistic Regression_2:

According to the logistic regression outputs provided in Table 12 below, we observed that the overall model including 16 PC s is statistically significant (p <.0001).

Likelihood Ratio Test for Global Null Hypothesis: BETA=0

-2 Log Likelihood Intercept Only	-2 Log Likelihood Intercept & Covariates	Likelihood Ratio Chi-Square	DF	Pr > ChiSq
15362.914	10058.610	5304.3041	12	<.0001

Table 12: Overall Model Evaluation Results.

Furthermore, exp (b) values (odd ratios) of significant PCs shows that with the increase in those PCs that are over 1.0, it would be more likely to face a backorder status (See the results in Table 13).

Parameter	went_on_ backorder Yes	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept	1	1	-0.1445	0.0285	25.68	<.0001		0.865
PC_1	1	1	0.1673	0.0144	135.74	<.0001	0.2480	1.182
PC_10	1	1	-0.6708	0.0402	277.79	<.0001	-0.3742	0.511
PC_11	1	1	-1.2357	0.0579	455.42	<.0001	-0.5995	0.291
PC_12	1	1	-0.7524	0.0580	168.43	<.0001	-0.3457	0.471
PC_13	1	1	-1.0314	0.0562	336.83	<.0001	-0.4410	0.357
PC_15	1	1	0.4304	0.0545	62.33	<.0001	0.1264	1.538
PC_2	1	1	0.0923	0.0206	20.06	<.0001	0.0767	1.097
PC_3	1	1	0.9144	0.0427	458.19	<.0001	0.6141	2.495
PC_4	1	1	0.6182	0.0742	69.35	<.0001	0.3661	1.855
PC_6	1	1	0.1391	0.0279	24.89	<.0001	0.0995	1.149
PC_7	1	1	0.3258	0.0229	202.20	<.0001	0.2382	1.385
PC_9	1	1	-2.9565	0.0794	1386.23	<.0001	-1.7515	0.052

Table 13: Results of Logistic Regression_2 Analysis.

This model has 81.21% accuracy, 80.59% sensitivity, 81.83% specificity and 81.60% precision based on validation model (See Table 14).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	80.5496	81.9888	4543	40.9944
1	0	19.4504	19.7979	1097	9.8989
0	1	18.3388	18.0112	998	9.0056
1	1	81.6612	80.2021	4444	40.1011

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	80.8289	81.8354	3023	40.9177
1	0	19.1711	19.4099	717	9.7049
0	1	18.3936	18.1646	671	9.0823
1	1	81.6064	80.5901	2977	40.2951

Table 14: Classification Results of the Logistic Regression_2 Model.

Moreover, the cumulative lift of the model is presented in the Figure 8.

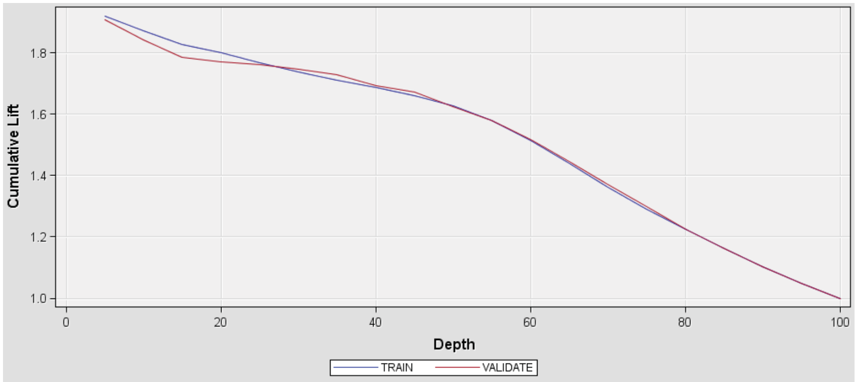


Figure 8: Cumulative Lift of the Logistic Regression_2 Model.

g. Decision Tree_2:

In the Decision Tree_2 model, we specified the assessment measure as misclassification with 2 max branch and 6 max depth. The decision tree elicited is presented in Figure 9 below.

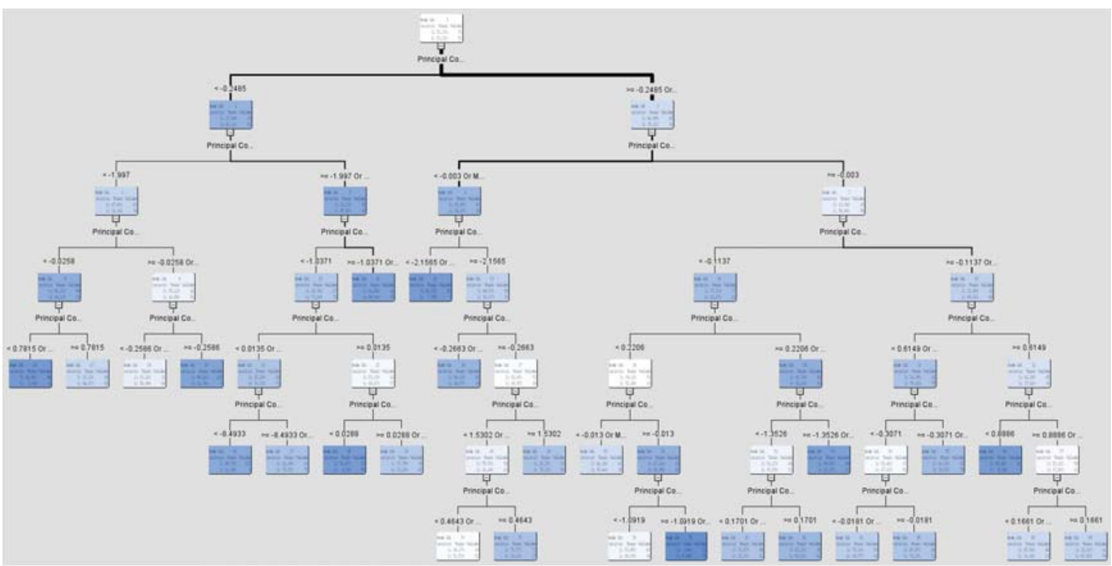


Figure 9: Decision Tree_2

The PCs are listed based on their importance to classify the target variable in Table 15 below. PC 9 has the most importance among the other PCs.

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
PC_9	Principal Component 9	4	1.0000	1.0000	1.0000
PC_14	Principal Component 14	1	0.6642	0.6268	0.9436
PC_6	Principal Component 6	1	0.4792	0.4380	0.9139
PC_1	Principal Component 1	4	0.4656	0.4365	0.9374
PC_3	Principal Component 3	3	0.3495	0.3131	0.8958
PC_16	Principal Component 16	4	0.1968	0.1504	0.7645
PC_12	Principal Component 12	1	0.1456	0.0283	0.1941
PC_2	Principal Component 2	1	0.1397	0.1254	0.8975
PC_7	Principal Component 7	2	0.1355	0.1677	1.2382
PC_8	Principal Component 8	1	0.1337	0.1074	0.8029
PC_15	Principal Component 15	1	0.1146	0.0766	0.6687
PC_4	Principal Component 4	1	0.0869	0.0900	1.0349
PC_5	Principal Component 5	1	0.0700	0.0000	0.0000

Table 15: Variable Importance

This model has a % 81.85 accuracy, % 86.70 sensitivity, % 76.98 specificity and % 79.02 precision based on validation model (See Table 16).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	86.0112	77.6755	4304	38.8378
1	0	13.9888	12.6331	700	6.3165
0	1	20.3521	22.3245	1237	11.1622
1	1	79.6479	87.3669	4841	43.6835

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	85.2774	76.9897	2844	38.4949
1	0	14.7226	13.2918	491	6.6459
0	1	20.9721	23.0103	850	11.5051
1	1	79.0279	86.7082	3203	43.3541

Table 16: Classification Results of the Decision Tree_2 Model.

Moreover, the cumulative lift of the model presented in the Figure 10.

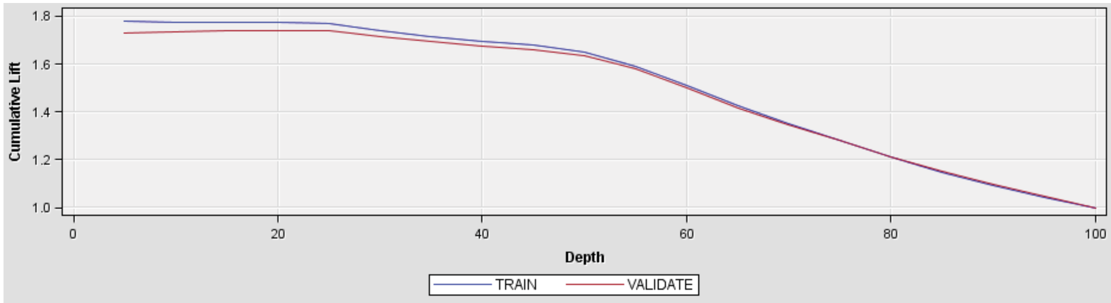


Figure 10: Cumulative Lift of the Decision Tree_2 Model.

h. Neural Network_2:

This model has 85.93% accuracy, 89.12% sensitivity, 82.75% specificity and 83.78% precision based on validation model (See Table 17).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	88.4474	82.3498	4563	41.1749
1	0	11.5526	10.7562	596	5.3781
0	1	16.5119	17.6502	978	8.8251
1	1	83.4881	89.2438	4945	44.6219

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	88.3781	82.7558	3057	41.3779
1	0	11.6219	10.8825	402	5.4413
0	1	16.2128	17.2442	637	8.6221
1	1	83.7872	89.1175	3292	44.5587

Table 17: Classification Results of the Neural Network_2 Model.

Furthermore, the cumulative lift of the model is presented in the Figure 11.

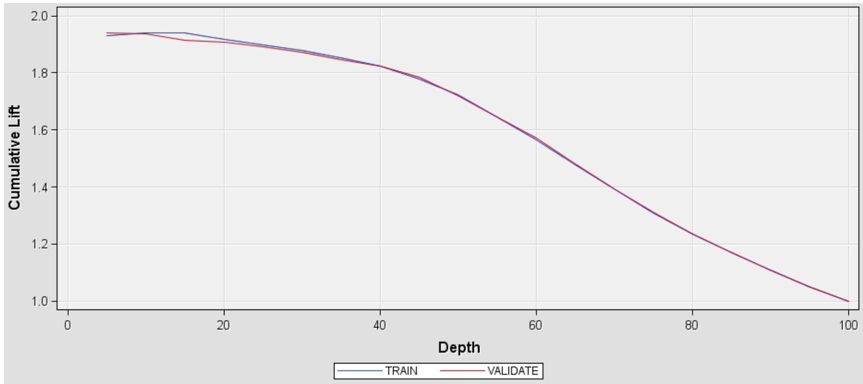


Figure 11: Cumulative Lift of the Neural Network_2 Model.

i. Autoneural_2:

This model has 84.98% accuracy, 86.22% sensitivity, 83.73% specificity and 84.12% precision based on validation model (See Table 18).

Classification Table

Data Role=TRAIN Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	86.0044	84.0642	4658	42.0321
1	0	13.9956	13.6798	758	6.8399
0	1	15.5842	15.9358	883	7.9679
1	1	84.4158	86.3202	4783	43.1601

Data Role=VALIDATE Target Variable=went_on_backorderYes Target Label=went_on_backorderYes

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	85.8690	83.7304	3093	41.8652
1	0	14.1310	13.7791	509	6.8896
0	1	15.8743	16.2696	601	8.1348
1	1	84.1257	86.2209	3185	43.1104

Table 18: Classification Results of the Autoneural_2 Model.

Furthermore, the cumulative lift of the model is presented in the Figure 12.

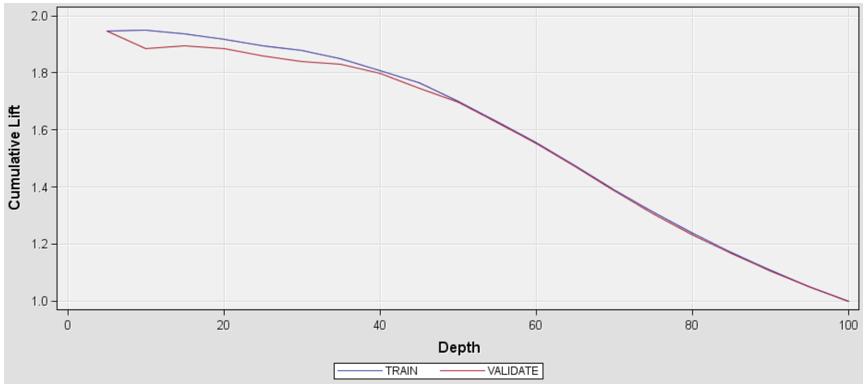


Figure 12: Cumulative Lift of the Autoneural_2 Model.

j. Model Comparison:

We compared the models (See Table 19 and Table 20) based on two different assessment criteria: misclassification and average square error (ASE).

Fit Statistics
Model Selection based on Valid: Misclassification Rate (_VMISC_)

Selected Model	Model Node	Model Description	Valid: Misclassification Rate	Train: Average Squared Error	Train: Misclassification Rate	Valid: Average Squared Error
Y	AutoNeural	AutoNeural	0.12777	0.09312	0.12687	0.09535
	Neural2	Neural Network (2)	0.14063	0.10319	0.14203	0.10315
	Tree	Decision Tree	0.14077	0.10982	0.13806	0.11255
	Neural	Neural Network	0.14226	0.10292	0.14059	0.10456
	AutoNeural2	AutoNeural (2)	0.15024	0.10742	0.14808	0.11181
	Tree2	Decision Tree (2)	0.18151	0.13368	0.17479	0.13869
	Reg	Regression	0.18286	0.13769	0.18273	0.13701
	Reg2	Regression (2)	0.18787	0.13971	0.18905	0.13849

Table 19: Comparison of the Models Based on Misclassification Rate of the Validation Models.

Fit Statistics
Model Selection based on Valid: Average Squared Error (_VASE_)

Selected Model	Model Node	Model Description	Valid: Average Squared Error	Train: Average Squared Error	Train: Misclassification Rate	Valid: Misclassification Rate
Y	AutoNeural	AutoNeural	0.09535	0.09312	0.12687	0.12777
	Neural2	Neural Network (2)	0.10315	0.10319	0.14203	0.14063
	Neural	Neural Network	0.10456	0.10292	0.14059	0.14226
	AutoNeural2	AutoNeural (2)	0.11181	0.10742	0.14808	0.15024
	Tree	Decision Tree	0.11255	0.10982	0.13806	0.14077
	Reg	Regression	0.13701	0.13769	0.18273	0.18286
	Reg2	Regression (2)	0.13849	0.13971	0.18905	0.18787
	Tree2	Decision Tree (2)	0.13869	0.13368	0.17479	0.18151

Table 20: Comparison of the Models Based on ASE of the Validation Models.

Autoneural_1 model provides the lowest misclassification rate (12%) and the lowest average squared error (9%). Autoneural_1 model also performs better than the other models in terms of accuracy, precision, specificity and sensitivity (See Table 21). In other words, the model that uses autoneural algorithm with all input variables provides the highest predictive power in this case.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Logistic Regression_1	81.71	80.50	82.91	82.49
Decision Tree_1	85.92	87.73	84.10	84.66
Neural Network_1	85.77	87.46	84.08	84.60
Autoneural_1	87.22	88.49	85.95	86.29
Logistic Regression_2	81.21	80.59	81.83	81.60
Decision Tree_2	81.85	86.70	76.98	79.02
Neural Network_2	85.93	89.12	82.75	83.78
Autoneural 2	84.98	86.22	83.73	84.12

Table 19: Performance of the Models Based on Accuracy, Sensitivity, Specificity and Precision

MANAGERIAL IMPLICATIONS / CONCLUSIONS

Backorder is a serious supply chain problem, affecting an inventory system's level of service, efficiency and customer satisfaction. The ability to detect parts with the highest probability of shortage before it occurs can be leveraged as an opportunity rather than a risk, which can lead to a competitive advantage. In this project, we have investigated different algorithms in order to come up with a forecasting model for this imbalanced data problem, where the occurrence of a backorder is rare but important.

We have used logistic regression, decision trees, neural networks and auto neural data mining techniques to predict our target variable, backorder status. In the second phase, we investigated whether using principal component analysis, which resulted in 16 principal components, would improve the predictive power of our modeling. For comparison, we used the model misspecification rate and Average Squared Error (ASE) criteria. We also considered the accuracy, sensitivity, specificity and precision of the models. Although not dramatically different, the first auto neural model comes up with the best performance of .12 misclassification rate and .09 ASE among the other models. The model also has the highest accuracy rate (87.2%), sensitivity (88.49%), specificity (85.95%) and precision rates (86.29%)

Overall, we can conclude that the use of machine learning techniques for forecasting backorders provide more precise forecasts than traditional forecasting techniques such as logistic regression. By implementing our auto neural model, businesses can accurately predict the backorder status of a product, in turn, minimize cost by more efficient inventory control, and increase profits by higher customer satisfaction. The importance level of the variables in the model also calls for attention. Operations managers should pay greater attention to forecast levels for 3 months as well as maintaining a balanced national inventory level. Defining a minimum and maximum level of national inventory level is recommended.

Future research could be directed to investigating other supply chain phenomena such as transportation management, demand forecasting, inventory levels, cash-to-cash cycles, supplier quality and production planning using machine-learning techniques to increase the robustness and generalizability of this approach.

REFERENCES

Abad, P. L. (2000). Optimal lot size for a perishable good under conditions of finite production and partial backordering and lost sale. *Computers & Industrial Engineering*, *38*(4), 457-465.

Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *The Journal of marketing*, 53-66.

Carter, C. R., & Rogers, D. S. (2008). A framework of sustainable supply chain management: moving toward new theory. *International journal of physical distribution & logistics management*, *38*(5), 360-387.

Chandra, C., & Grabis, J. (2008). Inventory management with variable lead-time dependent procurement cost. *Omega The International Journal of Management Science* , 36, 877-887.

Hariga, M., & Ben-Daya, M. (1999). Some stochastic inventory models with deterministic variable lead time. *European Journal of Operational Research*, *113*(1), 42-51.

Koh, S. (2004). MRP-controlled batch-manufacturing environment under uncertainty. Journal of the Operational Research Society , 55, 219-232.

Koh, S., Saad, S. M., & Jones, M. H. (2002). Uncertainty under MRP-planned manufacture: review and categorization. *International Journal of Production Research* , 40 (10), 2399-2421.
Montgomery, D. C., Bazaraa, M. S., & Keswani, A. K. (1973). Inventory models with a mixture of backorders and lost sales. *Naval Research Logistics Quarterly*, *20*(2), 255-263.
Ouyang, L. Y., Chuang, B. R., & Wu, K. S. (1999). Optimal inventory policies involving variable lead time with defective items. *Opsearch*, *36*(4), 374-389.
Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E., & Shankar, R. (2008). *Designing and managing the supply chain: concepts, strategies and case studies*. Tata McGraw-Hill Education.
Vijayan, T., & Kumaran, M. (2008). Inventory models with a mixture of backorders and lost sales under fuzzy cost. European Journal of Operational Research, 189(1), 105-119.

DECISION SCIENCES INSTITUTE

UNIX administrator ISP compliance: The impact of focused SETA workshops on heuristics, biases, and trust

John P. McConnell
Nova Southeastern University
Email: jm3967@mynsu.nova.edu

ABSTRACT

Information security policy compliance is crucial to the success of healthcare organizations. UNIX administrators maintain servers that house PHI/PII. The use of cognitive heuristics and biases may influence compliance behavior. This study evaluates the effect of SETA workshops on heuristics, biases, and trust using credentialed security scans on servers.

KEYWORDS: Knowledge Sharing, Privacy and Security, SETA, ISP Compliance, Heuristics, and Biases

INTRODUCTION

The healthcare industry is a complicated network of hospitals, providers, independent laboratories, payers, pharmacies, imaging centers, and public health departments centered on patients and their health (Dixon, 2016). The ability to safely and efficiently store, process, and exchange information about patient care between the healthcare industry participants is key to improving patient medical outcomes and lowering the cost of healthcare (Office of the National Coordinator for Health Information Technology, n.d.; Steinbrook, 2009; Thieme, 2016). In the United States, the implementation of information technology has been encouraged by the federal government. The United States Patient Protection and Affordable Care Act (ACA) of 2010 provided incentives to organizations to apply technology to the healthcare system. Additionally, the American Recovery and Reinvestment Act (ARRA) of 2009 allocated $145B for health care spending which included $30 billion to modernize the information technology infrastructure of health care organizations (Ajami & Bagheri-Tadi, 2013). Both laws contain provisions defining the privacy and security requirements necessary to protect patient personal health information (Steinbrook, 2009; Thieme, 2016). The goal of these laws was to encourage the use of technology to reduce healthcare costs by improving efficiency, reducing medical errors, reducing care duplication, and improving coordination of care among medical providers (Amarasingham, Plantinga, Diener-West, Gaskin & Powe, 2009; Office of the National Coordinator for Health Information Technology, n.d.). The move from paper-based health records to electronic health records that are shared among diverse organizations has resulted in a greater risk of data breaches and violations of patient's protected health information (PHI) and personally identifiable information (PII) (McFarland, 2012).

The foundation for the safeguarding of health information in the United States was established in two pieces of legislation: The Health Insurance Portability and Accountability Act (HIPAA) of 1996 and the Health Information Technology for Economic and Clinical Health Act (HITECH) of 2009. HIPAA provided the definition of PHI and set forth the minimum necessary standards for sharing any form of PHI as well as the regulations for the electronic interchange of PHI (Bernstein & Hartsell, 2013; Thieme, 2016). The HITECH Act expanded on HIPAA's definitions of security controls and further defined proper management of electronic health information, as well as providing financial incentives for electronic health record usage as well as meaningful use of health information technology (Bernstein & Hartsell, 2013; Hersh et al., 2016; Runyon, 2017).

The HIPAA Privacy Rule provided standards related to the disclosure and use of PHI, the rights of patients to access or correct their data, and patient's rights to request an audit of accesses and usage of their PHI (Avancha, Baxi, & Kotz, 2012; McFarland, 2012; Steinbrook, 2009). The goal of the Privacy Rule was to protect patient privacy rights while still allowing for the benefits of electronic health records (Avancha et al., 2012). The HIPAA Security Rule provided specifications and standards that covered entities should implement to ensure the confidentiality, integrity, and availability of PHI (Koch, 2017). Three categories of safeguards were defined to direct organizations in how best to protect PHI including physical safeguards, administrative safeguards, and technological safeguards (Avancha et al., 2012). Administrative safeguards include policies, procedures, and administrative actions related to security management, vulnerability and risk assessment, workforce security training, incident reporting, and contingency planning (Koch, 2017). Physical safeguards include facility access controls, computer controls, and device and media security controls (Avancha et al., 2012). Technical safeguards include access controls, audit controls, integrity management, authentication, and transmission controls for PHI (U.S. Department of Health and Human Services, 2013). Both the Privacy Rule and Security Rule outline civil and criminal penalties for the privacy violations (McFarland, 2012). The cost of data breaches can be significant for covered entities, but patients may also bear a significant cost when their PHI is stolen (Koch, 2017). In 2018, in the United States the average cost of a single data breach was $7.91 million which included the costs of detection, notification, response, and lost customer revenue (Ponemon Institute, 2018). In 2018, HIPAA penalties and settlements levied on nine healthcare systems and insurance companies totaled $28,683,400 (HIPAA Journal, 2019). Given the potential liabilities associated with data breaches and privacy violations it is imperative that healthcare organizations secure their computing resources by following the HIPAA and HITECH guidelines.

While many organization's servers are Windows based, a significant number of larger, back-end systems are UNIX based to capitalize on increased server processing power, reliability, security, and clustering technology (Bajgoric, 2006; Beuchelt, 2013; Hussain et al., 2015). In terms of hospital EMR systems, which house a significant amount of HIPAA protected patient data, Epic and Cerner represent 61% of the market for EMR implementations in inpatient hospitals in the United States (Newman, 2019; Shrivastava, 2018). The Epic EMR application, the industry leader for hospital EMR systems, only supports UNIX based operating systems including IBM UNIX (AIX), HP UNIX (HP-UX), and RedHat Linux (Epic, 2018; Newman, 2019). Larger Cerner EMR customers use high-end UNIX and Linux servers for the back-end databases while smaller hospitals use the Windows server based version of the Cerner EMR product (Shrivastava, 2018). As of May 2019, Linux/UNIX servers represent 69.8% of all active web servers worldwide (W3Techs, 2019). In the Amazon cloud, Linux/UNIX images represent 89.42% of the servers and Windows represents 10.54% of the servers out of a total of 963512 images (Cloud Market, 2019). When considering the total distinct known vulnerabilities from 1999 to 2019, the top six operating systems are Linux-based and Linux variants (CVE Details, 2019). The top six Linux variants have 12,635 known vulnerabilities and the top four Windows operating systems have 3,062 vulnerabilities (CVE Details, 2019). Clearly, there are significant vulnerabilities with Linux and UNIX systems that need to be addressed to protect the systems that house HIPAA protected data (Caballero, 2013; Santara, 2013).

PROBLEM

Most research in information security policy (ISP) compliance has focused on the end-user compliance intention and compliance behavior (Albrechtsen & Hovden, 2010; Dang-Pham, Pittayachawan, & Bruno, 2017; Hanus & Wu, 2016; Ifinedo, 2012; Kraemer & Carayon, 2007;

Safa, Von Solms, & Furnell, 2016). While end-users are critical to organizations reducing the threat of information security breaches, another group of specialized users has the highest privilege levels and legitimate access to the vast amount of confidential PHI and PII data stored on their systems—server systems administrators (Beuchelt, 2013; Kraemer & Carayon, 2007). Systems administrators are super users on servers and are responsible for operating system installation, configuration, patching, user management, monitoring, data backup, implementation of security controls, disaster recovery, and testing (Inshanally, 2018; Santara, 2013). Common information security threat vectors for servers include network, security, or operating system misconfiguration, unpatched operating systems or device firmware, privileged account access, and unsecured data or backups (Caballero, 2013; Donaldson, Siegel, Williams, & Asalm, 2015). From February 2017 to February 2019, server related incidents resulted in the loss of 6,651,957 individual's PHI records (U.S. Department of Health and Human Services Office for Civil Rights, 2019). Server related incidents were due to hacking and IT incidents (89%), unauthorized access and disclosure (11%), and theft (<1%) (U.S. Department of Health and Human Services Office for Civil Rights, 2019). Server breaches were responsible for 46% of all healthcare PHI breaches from February 2017 to February 2019 (U.S. Department of Health and Human Services Office for Civil Rights, 2019). These vulnerabilities and statistics demonstrate the importance for server administrators to consistently follow organizational ISP to ensure the confidentiality, integrity, and availability of healthcare systems, and the PII, and PHI contained therein.

Human factors have gained prominence as a significant risk factor for information systems security (Colwill, 2009; Ifinedo, 2014; Safa et al., 2015). To strengthen the human aspect of information security, information systems policies were developed that enhance security, decrease vulnerability to security breaches, and ensure legal compliance (Furnell & Clarke, 2012; Ng, Kankanhalli, & Xu, 2009). Unfortunately, researchers have found that employees frequently circumvented information systems policies when workload increased (Albrechtsen & Hovden, 2010; Guo, Yuan, Archer, & Connelly, 2011; Kraemer & Carayon, 2007) or when they felt the information systems policies were a nuisance (Renaud, 2012; Sedighi, van Splunter, Brazier, van Beers, & Lukosch, 2016). Siponen, Mahmood, and Pahnila (2014) identified employee failure to follow information systems policies as a key threat to the security of an organization. An additional risk is that employees can make errors due to cognitive limitations, task demands, organizational, social, or environmental factors (Dismukes, Berman, & Loukopoulos, 2007; Safa & Von Solms, 2016). Most of these studies have evaluated end-user ISP compliance. Compliance of UNIX systems administrators (UXA), however, is even more significant as the data housed on the back-end UNIX servers may contain PHI, PII, financial data, or intellectual property.

Information security knowledge is frequently scattered throughout organizations and many organizations have not developed an effective information security knowledge sharing program (Belsis, Kokolakis, & Kiountouzis, 2005; Flores, Antonsen, & Ekstedt, 2014; Safa & Von Solms, 2016). Additionally, organizational silos, where there is a rigid functional division between teams, can negatively impact social interaction and knowledge sharing (Oparaocha, 2016). Trust is key for knowledge sharing relationships (Dang-Pham, Pittayachawan, & Bruno, 2016; Safa & Von Solms, 2016). The most effective way of increasing security knowledge sharing and cyber skills is through effective security education, training, and awareness (SETA) programs (Oltsik, 2017). Wash and Cooper (2018) found SETA programs to be the most effective means of changing the security behaviors of end-users. Bauer and Bernroider (2017), in assessing the impact of information security awareness on end-user compliance behavior, found that security awareness significantly positively influenced attitude toward compliance and provided a weak negative relationship to neutralizing behaviors. These studies have focused on the effectiveness of SETA programs on ISP compliance of end-users. Although ISP compliance of server administrators is crucial to protecting the organization's data, to the author's knowledge, studying the effectiveness of targeted SETA workshops and their influence on UNIX server administrators has not been researched.

The research problem that this study will address is that some UNIX administrators may fail to completely implement organizational ISP believing the threat of privacy breaches to be primarily a problem for Windows servers. This failure may leave UNIX servers open to potential systems disruption, loss of proprietary or confidential data, harm to organizational reputation, potential loss of revenue, or financial loss due to litigation or fines (Donaldson et al., 2015; Kraemer & Carayon, 2007). It is crucial to investigate the cognitive processes (heuristics and biases) that UXA may employ which can result in inadequate information security threat appraisal, inflated coping appraisal, and could encourage ISP non-compliance. It is vital to understand of how specially designed SETA programs influence UXA use of heuristics and biases. This research is needed as implementing HIPAA mandated security and compliance with organization ISP are key to the successful protection of healthcare organizational assets, including patient PII, and PHI (Koch, 2017; Ng et al., 2009). Also, this research will help fill the gap on SETA program effectiveness with UNIX server administrators.

PURPOSE

The main goal of this research is to empirically assess how SETA workshops influence UNIX administrator's use of cognitive heuristics and biases and how those biases influence knowledge of security vulnerabilities and their severity, information security response-efficacy, information security self-efficacy, and ultimately compliance behavior with ISP and protection of PII and PHI. It is crucial that organizations comply with the U.S. federal government's guidelines in the HIPAA Privacy and Security rules to ensure the confidentiality, integrity, and availability of PHI and minimize the risk of breaches and potential fines (Avancha et al., 2012; U.S. Department of Health and Human Services, 2013). Additionally, this research will seek to understand how the focused SETA program affects UNIX administrator trust and ISP compliance behavior. Finally, by integrating Protection Motivation Theory (PMT) components, this research will assess the impact of server administrator focused SETA programs on UNIX administrator's threat appraisal (perceived threat severity, perceived threat vulnerability), coping appraisal (response efficacy and self-efficacy), and ultimately ISP compliance behavior.

RESEARCH QUESTIONS

RQ1: How do formal knowledge sharing arrangements and support for knowledge transfer, in the form of SETA workshops, influence UXA security threat appraisal and coping appraisal?
RQ2: How does UXA trust in the Information Security team influence ISP compliance?
RQ3: How does information security knowledge sharing, in the form of SETA workshops, impact UXA use of cognitive heuristics and biases?
RQ4: How does the use of cognitive heuristics and biases impact UXA threat and coping appraisals?

HYPOTHESES

H1: Formal knowledge sharing arrangements, in the form of a focused SETA workshop and ISCs, significantly negatively influences UXA use of the availability heuristic.
H2: Formal knowledge sharing arrangements, in the form of a focused SETA workshop and ISCs, significantly negatively influences UXA use of optimism bias.
H3: Formal knowledge sharing arrangements, in the form of a focused SETA workshop and ISCs, significantly negatively influences UXA use of confirmation bias.

H4: Formal knowledge sharing arrangements, in the form of a focused SETA workshop and ISCs, significantly positively influences UXA trust.
H5: Trust in the Information Security Team by UXA significantly positively influences UXA ISP compliance behavior.
H6: The availability heuristic significantly negatively influences UXA information security threat appraisals.
H7a: Optimism bias significantly negatively influences UXA information security threat appraisal.
H7b: Optimism bias significantly positively influences UXA information security coping appraisal.
H8a: Confirmation bias significantly negatively influences UXA information security threat appraisal.
H8b: Confirmation bias significantly positively influences UXA information security coping appraisal.
H9: UXA security threat appraisal significantly positively influences UXA ISP compliance behavior.
H10: UXA security coping appraisal significantly positively influences UXA ISP compliance behavior.

LITERATURE REVIEW

Security education, training & awareness

The sharing of information security knowledge, experience, and insights can improve organizational performance and help to ensure the security of data (Safa & Von Solms, 2016). Development of a formal means for information security knowledge sharing can help to foster sharing of ideas, experiences, tools, and processes to improve security and protect an organization's information systems assets (Flores et al., 2014). Making users aware of the current and evolving information security risks, threats, vulnerabilities, and their severities, the speed with which the threats propagate, and the potential impact to the organization is crucial to information security policy compliance (Guo et al., 2011; Safa et al., 2016; Siponen et al., 2014). Dang-Pham et al. (2017) found that security awareness also improved the diffusion of information security practices (knowledge sharing) throughout the organization. Safa and Van Solms (2016) found that information security knowledge sharing benefitted business, increased employee information security self-efficacy, and improved ISP compliance.

The development of an information security knowledge sharing culture is an important goal for any organization that has critical information systems assets (Flores et al., 2014; Razmerita, Kirchner, & Nielsen, 2016; Safa & Van Solms, 2016). End-user education is important and developing knowledge sharing processes that include the information security team and UXA is crucial to any organization that has UNIX servers hosting business critical data (Bauer, Bernroider, & Chudzikowski, 2017). It is beneficial to build both formal and informal knowledge sharing networks within organizations as they have been found to be significant contributors to awareness and mitigation of information security risks (Dang-Pham et al., 2017; Safa & Von Solms, 2016; Yoo, Sanders, & Cerveny, 2018). Encouraging relationships between employees across team boundaries is also helpful to developing and enabling an effective social network that fosters knowledge sharing (Bauer et al., 2017; Ifinedo, 2014; Oparaocha, 2016). Connelly and Zweig (2015) found that distrust was a predictor of knowledge hiding behaviors which are detrimental to effective knowledge sharing in organizations. Consequently, it is important to encourage trusting relationships between the information security team and the UXA for an effective information security knowledge sharing culture (Dey & Mukhopadhyay, 2018; Rutten, Blaas-Franken, & Martin, 2016). Sedighi et al. (2016) found that social interaction was a key component of knowledge sharing. Therefore, creating regular interactions between the information security team and UXA is crucial to development of reciprocal, trusting relationships that enable knowledge sharing and improved security awareness and ISP compliance behavior.

SETA programs are a means for organizations to minimize the risk of insider caused security failures (Burns, Roberts, Posey, Bennett, & Courtney, 2015). SETA programs are an important antecedent and positively influence information security behavior, and appropriately designed SETA programs can help reduce the human information security risk to organizational assets (Whitman & Mattord, 2012). Users are the weakest link for information security and SETA programs can help to reduce the potential attack surface of organizations by improving the ability of users to identify and prevent information security breaches (Furnell & Clarke, 2012; Gardner & Thomas, 2014). Posey, Roberts and Lowry (2015) found that SETA programs were positively correlated with both perceived threat severity and response efficacy indicating that they are an effective way of encouraging information security behavior and ISP compliance. SETA programs should be regular and recurring due to the dynamic nature of information security threats and vulnerabilities (Posey et al., 2015). In a study of SETA effectiveness, Yoo et al. (2018) found that the psychological flow factors including feedback, immersion, challenge, autonomy, and social interaction significantly improved psychological ownership and SETA program effectiveness. Yoo et al. (2018) suggested using relatable security scenarios that challenged employees could improve ownership and ISP compliance.

Bauer et al. (2017), in a study of ISP compliance at banks, found that developing a comprehensive, multi-modal information security awareness program was key to successfully establishing an information security culture in an organization. The goal of their research was to define specific propositions that could be used by information security management in banks to establish and maintain an effective information security awareness (ISA) program (Bauer et al., 2017). Also, Bauer et al. (2017) sought to develop an understanding of how users' perceptions of the ISA program influence their ISP compliance. The three banks studied had implemented ISA programs but with very different processes and procedures (Bauer et al., 2017). The most successful bank, in terms of information security awareness and employee engagement, conducted regular ISA campaigns using different modalities which encouraged high levels of interaction and dissemination of critical security knowledge (Bauer, et al., 2017). The regular contact included daily emails, intranet postings, weekly themes, and quizzes as well as role modeling both positive and negative behaviors (Bauer, et al., 2017). Another bank required employees to sign IS agreements and performed periodic monitoring of employees IS behavior (Bauer et al., 2017). This was viewed by some employees as a lack of trust from the organization toward the employees. Bauer et al. (2017) found that individuals used different neutralizing behaviors to justify ISP non-compliance. The third bank Bauer et al. (2017) studied had just begun an ISA program and provided little insight into SETA effectiveness. Two primary areas of design for ISA were proposed by Bauer et al. (2017): structural design and communicational design. Structural recommendations included media rich interventions that appeal to multiple learning styles; implementation of regular review and revision so that ISA programs evolve with the dynamic IS threat landscape; and customizing the ISA programs toward the recipients (Bauer, et al., 2017). The communicational recommendations Bauer et al. (2017) identified included: limited use of security vernacular; matching the security awareness messages to the recipient's personality; driving for two-way discussions including IS reflection and ISP enforcement; effectively using positive and negative role models and using role play; and constantly seeking feedback about ISA program effectiveness so that the programs remain relevant and provide an effective means of sharing IS knowledge with employees.

Albrechtsen and Hovden (2010) developed and tested information security workshops where security personnel acted as facilitators for end-users to discuss relevant security scenarios. This small group atmosphere lead to collaborative, two-way dialogs and fostered participation and collective reflection to gain insights from one another (Albrechtsen & Hovden, 2010). Safa et al. (2016) integrated social bond theory and involvement theory and found that knowledge sharing, collaboration, intervention, experience, commitment, and personal norms were all correlated to attitude toward compliance with ISP. Sedighi, et al. (2016) found reputation, reciprocity, altruism, and knowledge self-efficacy were all positively related to quantity and quality of knowledge sharing while effort and time were negatively related to both.

Organizations spend a great deal of money on SETA programs in the hopes of increasing employee ISP compliance and engagement but employees continue to cause significant security breaches due to their failure to comply with ISP (Yoo et al., 2018). Managers need to understand what psychological antecedents may improve the effectiveness of SETA programs and increase employee compliance (Yoo et al., 2018). Yoo et al. (2018) sought to identify the factors that influence psychological flow and how they impact SETA effectiveness and ultimately ISP compliance behavior. The constructs tested by Yoo et al. (2018) included: challenge, feedback, autonomy, immersion, social interaction, flow, psychological ownership, SETA effectiveness, self-efficacy, and security compliance intention. The components of flow (feedback, immersion, challenge, autonomy, and social interaction) all significantly influenced psychological ownership and SETA effectiveness (Yoo et al., 2018). An important detail found by Yoo et al. (2018) was that training must be at an appropriate level to challenge but not overwhelm the participant to support engagement. Psychological ownership and SETA effectiveness significantly influenced security behavioral intention (Yoo et al., 2018). Yoo et al. (2018) suggested using relatable scenarios to help in connecting employees to the SETA content. Yoo et al. (2018) brought the concept of flow into the knowledge management realm in terms of the impact that flow has on SETA.

Dang-Pham et al. (2017) noted that employee information security failures can lead to security breaches causing substantial financial loss for organizations. SETA programs are effective in reducing the cost of breaches but cannot eliminate the problem of employees not following ISP due to negligence or malicious intent (Dang-Pham et al., 2017). Developing a security knowledge sharing culture may help employees gain the information security knowledge, develop an informal knowledge sharing network, and further reduce breaches caused by human error (Dang-Pham et al., 2017). Dang-Pham et al. (2017) investigated the factors that help to develop an information security knowledge sharing culture through the lens of Social Network Analysis, the Theory of Planned Behavior, and Accountability Theory to identify how employees use formal and informal networks to gain and share security knowledge. Encouraging regular interactions between employees is important as some information security knowledge sharing occurs during those informal dialogs (Dang-Pham et al., 2017). Also, developing trusting relationships with individuals increased formal and informal sharing between employees (Dang-Pham et al., 2017). Educating users regarding the benefits of ISP compliance, rather than just how to comply with the policy, was more effective in supporting self-efficacy and inter-employee sharing (Dang-Pham et al., 2017).

Ifinedo (2014) integrated the Theory of Planned Behavior, Social Behavioral Theory, and Social Cognitive Theory to provide additional insights into why employees comply or fail to comply with ISP. Ifinedo (2014) used nine constructs that were developed from prior research including: attachment, involvement, commitment, personal norms, attitude toward compliance, subject norms, locus of control, self-efficacy, and behavioral intention. Socialization, personal norms, social norms, self-efficacy, and group dynamics positively influenced ISP compliance intention (Ifinedo, 2014). These results suggest that managers should encourage formal and informal socialization of employees to increase trust and establish relationships that influence positive behaviors and discourage negative or malicious behaviors (Ifinedo, 2014). Perceived control also had a positive impact on ISP compliance intention (Ifinedo, 2014). Developing a better understanding of the organizational and social factors that influence ISP compliance allows organizations to prepare for and encourage appropriate behavior to safeguard security.

By adequately managing the information security knowledge and social norms of the organization managers can better control the antecedents of positive behavior while applying social pressure to curtail negative behavior.

Chen, Chen and Wu (2018) developed a model of ISP compliance based on an Awareness-Motivation-Capacity perspective. The significant influencers of ISP compliance intention included information security awareness (awareness of the ISP and potential threats), capability to comply (self-efficacy and controllability), and motivation to comply (penalty and reward) (Chen et al., 2018). Educating employees about the importance of information security can be achieved through SETA programs (Chen et al., 2018; Dang-Pham et al., 2017). The same programs can also inform employees about the organizational ISP and introduce them to the potential security threats facing the organization (Chen et al., 2018). Engaging and audience appropriate SETA programs will also influence information security self-efficacy and ISP compliance (Chen et al., 2018; Ifinedo, 2014).

Trust

Trust is a complex concept and frequently tailored in research to the specific discipline being studied (McKnight, Cummings, & Chervany, 1998). From the perspective of a consumer using an information technology artefact, trust can help overcome risk and insecurity (McKnight, Choudhury, and Kacmar, 2000). In a seminal study on trust, McAllister (1995) identified two types of trust: affect-based trust, and cognitive-based trust. Antecedents to affect-based trust included citizenship behavior, and interaction frequency (McAllister, 1995). Antecedents to cognitive-based trust include reliable role performance, cultural-ethnic similarity, and professional credentials (McAllister, 1995). McAllister's (1995) research studied trust and its influence on interpersonal cooperation in organizations which is relevant for the current research. From McAllister (1995), reliable role performance can relate directly to the reliability of the information security team's knowledge sharing with the UXA. Trusting the information security team's recommendations on security best practices may cause concerns about information accuracy, privacy, as well as perceived reliability of the information security team's understanding of how to secure the UNIX servers (Gefen, Karahanna, & Straub, 2003).

Gefen et al. (2003) integrated a trust model into the Technology Acceptance Model (TAM) framework. While their research was related to trust in an online vendor, their work can be integrated into the current research model given that the information security team, in facilitating the SETA program, act as a resource to the UXAs, and the team's reputation and experience can influence the UXA's trust in the relevance of the training. By extended TAM, Gefen et al. (2003) found that information integrity and institutional trust played a significant role in the attitudes and intention to use. Antecedents to trust that Gefen et al. (2003) found significant included: calculative-based, institution-based structural assurances, institution-based situational normality, and perceived ease of use (from TAM). These antecedents include features relevant to the current research. Calculative-based trust includes the cognitive processes that the UXA use to evaluate the information security team's trustworthiness. Gefen et al. (2003) identified antecedents to institution-based structural assurance to include structural safeguards, legal guarantees, and regulations that influence trust. Institution-based situational normality is the social context of the information technology artifact (Gefen et al., 2003). With respect to the current research, the ISP forms the basis for institution-based structural assurance, noted by Gefen et al. (2003), and provides the guidelines, processes, and rules surrounding information system security.

Lowry, Posey, Bennett, and Roberts (2015) developed a model based on Reactance Theory and Fairness Theory that included constructs of organizational trust, SETA programs, information security policy compliance, reactive ISP non-compliance (reactive computer abuse), and explanatory adequacy. There was a significant negative relationship between organizational trust and reactive computer abuse (ISP non-compliance behavior) (Lowry et al., 2015). This implies that trust in the organization can reduce ISP non-compliance behavior. Additionally, the SETA program developed by Lowry et al. (2015) had a significant positive influence on explanatory adequacy which in turn was a significant contributor to positive trust formation. Trust was also a significant mediator between explanatory adequacy and reactive computer abuse (Lowry et al., 2015). Finally, Lowry et al. (2015) found that the SETA programs reduced employees' perceptions of lost freedom and external control experienced when enhanced security policies were implemented.

Development of interpersonal relationships and social networks are significant contributors to knowledge sharing within an organization (Oparaocha, 2016). SETA programs and communities of practice can help in both formal and informal sharing of knowledge, the development of collaborative relationships, and the building of trust (Dang-Pham et al., 2017; Oparaocha, 2016). Supportive social networks can also improve relationships and encourage fostering of both cognitive-based and affective-based trust (Dey & Mukhopadhyay, 2018; Rutten et al., 2016). Trust is key for knowledge sharing relationships and the development of those relationships through an effective SETA program can be beneficial to the organization and encourage information security engagement and compliance (Dang-Pham et al., 2016; Safa & Von Solms, 2016).

Research conducted by Dey and Mukhopadhyay (2018) sought to understand how affective organizational commitment, behavioral intention, and affective trust influence knowledge sharing behavior. Knowledge sharing is key to organizational competitiveness, but it is reliant on individuals' willingness to share explicit and implicit knowledge (Dey & Mukhopadhyay, 2018). It is important to understand the factors that motivate knowledge sharing behaviors to increase organizational effectiveness (Dey & Mukhopadhyay, 2018). The constructs selected for their model included: affective trust, knowledge sharing intention, affective organizational commitment, and knowledge sharing behavior (Dey & Mukhopadhyay, 2018). Affective commitment mediated the relationship between knowledge sharing intention and knowledge sharing behavior and made the relationship significant (Dey & Mukhopadhyay, 2018). These results suggest that it is important to increase employee affective commitment to the organization in order to increase knowledge sharing behavior. Trust is also an important mediator of the relationship between knowledge sharing intention and knowledge sharing behavior (Dey & Mukhopadhyay, 2018). Trust between employees and trust between the employee and the organization are key to willingness and effective knowledge sharing (Dey & Mukhopadhyay, 2018). Given the importance of trust and organization commitment it is important to build social bonds between employees at all levels of the organization. A collaborative SETA program can help to foster social bonds and trust and increase organizational commitment of UXA.

Trust is a key component of knowledge sharing but very little is known specifically about the impact of low levels of trust on explicit and implicit knowledge sharing (Rutten et al., 2016). The knowledge sharing constructs studied by Rutten et al. (2016) were divided into explicit knowledge sharing and implicit knowledge sharing types. Explicit knowledge sharing types included: instructions, templates, and emails (Rutten et al., 2016). Implicit knowledge sharing types included: directly contacting a person, informal news, and tips & tricks (Rutten et al., 2016). Rutten et al. (2016) found that affective-based trust had significantly more influence on knowledge sharing for both implicit and explicit types of knowledge sharing. In terms of implicit knowledge sharing, affective-based trust was significantly more influential. These results demonstrate the need for socialization and the building of trust relationships in organizations for knowledge sharing to be effective (Rutten et al., 2016). The success of knowledge sharing, and knowledge management relies on the willingness of the individuals in the organization to share information, knowledge, skills, tips, and wisdom (Rutten et al., 2016). By better understanding how cognitive and affective-based trust influence knowledge sharing the community can develop a fuller understanding of how to facilitate better communication, collaboration, and knowledge sharing.

Van Vuuren (2016) proposed a new IT Security Trust (ITST) model to help understand how trust influenced security compliance for information systems. This model linked trust, knowledge sharing, and ISP compliance to better understand how to develop an information security culture (Van Vuuren, 2016). Van Vuuren (2016) used the term Information Security DNA to define the organizational culture inclusive of information security. Their model included organizational security trust, relationship trust, self-trust, social and home trust along with components that are knowledge sharing related-SETA, knowledge, training and education, understanding and acceptance (Van Vuuren, 2016). Additionally, Van Vuuren (2016) integrated two constructs related to PMT into their model including perceived security implications and risk perception. Each form of trust, self-trust, relationship trust, organizational security trust, and information security DNA based trust influenced security awareness and security knowledge sharing (Van Vuuren, 2016). The goal, like Flores et al. (2014), was to develop an information security based organizational culture formation model that included the influence of trust (Van Vuuren, 2016). Believing that trust is the cornerstone to organizational effectiveness, knowledge sharing, information security, and relationships Van Vuuren (2016) augmented the ITST model to integrate the key behavioral trust components.

Heuristics & biases

Kahneman (2011) referred to the two cognitive systems of decision making as System One and System Two. System One is the intuitive, implicit, involuntary, and nonverbal cognitive system (Kahneman, 2003). According to Kahneman (2003), intuitive judgements may harken to evolutionary history and occur “between the automatic operations of perception and the deliberate operations of reasoning” (p. 697). The intuitions provided by System One come to mind quickly with little reflection—they are automatic once a stimulus occurs (West, Meserve, & Stanovich, 2012). For example, a loud noise or gunshot draws immediate attention as System One quickly assesses the situation and determines if fight or flight is necessary. Also, when walking down the street, someone that is approaching is assessed by System One to evaluate the situation and quickly determine if the individual should be avoided. These analyses are pre-cognitive, and frequently termed intuitions (Tversky & Kahneman, 1974). Intuitions, which rely on similarity and accessibility rather than true logic or probabilities, can be flawed due to the use of cognitive heuristics (Kahneman, 2011; Tversky & Kahneman, 1974). Heuristics are mental short-cuts used to make inferences about situations, and they require a minimal amount of information and cognitive processing power (Gilovich & Griffin, 2013; Marsh, Todd, & Gigerenzer, 2004; Roberts, 2004). As opposed to algorithmic means of solving a problem, heuristics, by their nature, do not guarantee a correct answer (Gilovich & Griffin, 2013; Roberts, 2004). Heuristics provide a means of finding an adequate solution to a problem without having to consider all possible causally relevant information (Marsh et al., 2004). They also help to reduce constrained working memory (Kahneman, 2011; Toplak, West, & Stanovich, 2011). Examples of heuristics that influence decision making include: anchoring, availability, illusion of pattern, subjective confidence, the law of small numbers, prediction by representativeness, and the illusion of understanding (Kahneman, 2011).

System Two, the reasoning and analytical system, is where deliberate thought occurs (Kahneman, 2011). System Two is activated whenever a problem presents itself to which System One cannot provide a fast and reasonable answer (Kahneman, 2011). Unfortunately, System One frequently will answer a difficult or challenging question with an associated question (heuristic) that is easier to draw from memory (Kahneman, 2003). Attribute substitution can allow System One to answer a question that was not asked resulting in faulty decision making (Gilovich & Griffin, 2013). One of System Two’s responsibilities is to monitor System

One to ensure correct decisions are made (Kahneman, 2003). To reduce cognitive load, however, System Two may accept faulty System One responses due to what Kahneman (2011) terms lazy monitoring. If System Two is activated and engaged it may reject potentially biased System One intuitions but that activation is cognitively taxing (Kahneman, 2003). Kahneman (2003) found that individuals made aware of their use of heuristics were able to correct the intuitive judgement.

Epstein (2014), in the field of cognitive psychology, defined the two cognitive systems as the experiential system and the rational system. The experiential system functions outside individual's awareness and influences interpretation of feelings, behaviors, and events (Epstein, 2014). The experiential system is non-verbal, and activation requires minimal cognitive demand (Epstein, 2014). A key feature of the experiential system relevant to the current research is that it has the potential to learn from experience (Epstein, 2014). The rational system reflects an individual's personal understanding of logic and is uniquely human (Epstein, 2014). It represents conscious reasoning, verbal thought, tends to be affect free, considers cause & effect, is slower processing, and requires higher cognitive load (Epstein, 2014).

To judge the frequency or probability of an event an individual may assess the availability of associations related to the event (Tversky & Kahneman, 1974). Rather than taking the time to find an actual probability it is easier to estimate a probability based on the ease that one recalls occurrences of a similar event - termed the availability heuristic (Kahneman, 2011; Kliger & Kudryavtsev, 2010; Tversky & Kahneman, 1974). Pachur, Hertwig, and Steinmann (2012), in assessing how the availability heuristic influenced individual's judgement of risks, found that the availability heuristic significantly influenced general perceived risk. Since the question of frequency is difficult to answer without more information an easier question is substituted—how easily can examples of the event be recalled (Kahneman, 2011). If examples come to mind easily the frequency is estimated to be high and if examples are difficult to imagine the frequency is assumed to be low (Tversky & Kahneman, 1974). System One does not have the means to properly apply probability theory and the reliance on availability as an assessment of actual probability can lead individuals to faulty evaluations of risk (Kahneman, 2011). This is relevant to the current research as perceived threat can influence ISP compliance behavior.

Toplak et al. (2011) analyzed the use of the Cognitive Reflection Test (CRT) to assess cognitive performance. They found that the CRT predicted an individual's propensity toward cognitive errors (Toplak et al., 2011). Toplak et al. (2011) studied intelligence and working memory and found both were moderately predictive of rational thinking skills and cognitive performance. The conclusion was that the quick acceptance of the System One data was primarily due to cognitive load and the miserly cognitive processing of System Two (Toplak et al., 2011). These findings are in line with Kahneman's (2011) lazy monitoring performed by System Two. Toplak et al. (2011) warned that while the intuitive processing of System One may be useful it can also be dangerous due to the propensity to over simplify problems and underestimate risk.

Ferreira, Garcia-Marques, Sherman, and Sherman (2006) investigated heuristic problem-solving skills to understand what actions might encourage System Two engagement. Using modified versions of Tversky and Kahneman's (1974) heuristic problem set, Ferreira et al. (2006) found that providing priming instructions to participants helped them to resist System One intuitions and engage System Two reasoning. Kliger and Kudryavtsev (2010) defined priming as an unconscious process that occurs when a current stimulus increases the availability (recall) of past associations. Increased use of System Two resulted in significantly improved performance on the heuristic problems (Ferreira et al., 2006). Interestingly, Ferreira et al. (2006) concluded that automatic (System One) and controlled (System Two) processing are independent but can also be used in parallel. Education and practice can improve the reliability of System One's intuitions (Kahneman, 2003). An example of learned intuition can be found in chess masters who can quickly evaluate a chess board, analyze possible outcomes, and make moves seemingly instantaneously (Kahneman, 2003). This identifies the potential to improve the intuitive responses of UXA through effective and engaging SETA programs.

Optimistic bias is another bias that can result in dangerous neglect of risks (Rhee, Ryu, & Kim, 2012). Optimistic bias leads one to assess situations in self-serving ways (Rhee et al., 2012). This fundamental underestimate of risk can enhance perceived invulnerability to negative events and lead to inappropriately low levels of safeguarding behaviors related to information security (Rhee et al., 2012). Optimistic bias is a protective measure to protect the self, and reduce both anxiety and stress (Rhee et al., 2012). Rhee et al. (2012) found, in a study of information security perceptions of technology executives, that they perceived the security risks but that optimistic bias allowed them to conclude that their organization were at a much lower risk of security breach than other organizations. Rhee et al. (2012) suggested information security training is key to reducing optimistic bias and improving security practices within organizations. Optimistic bias can also cause individuals to discount future consequences, for example, one may believe they are not at risk of intrusion and therefore not protect personal privacy (Acquisti, 2004). Interestingly, this behavior was noted in both naïve and sophisticated individuals (Acquisti, 2004).

Confirmation bias and optimism bias are closely related and can significantly influence decision making (Kahneman, 2011). With confirmation bias, one gives greater validity to information that supports rather than contradicts one's beliefs (Sternberg, 2004). Tsohou, Karyda, & Kokolakis (2015) suggested that confirmation bias may lead individuals to believe that hackers are not sophisticated or dangerous and discount the threat caused by nation states, organized crime, or terrorists. Kahneman (2011) identifies confirmation bias as a System One heuristic and it is therefore easily activated when making decisions. System Two must be engaged to contradict a System One confirmation bias but in most situations, individuals do not devote the cognitive energy to disprove their beliefs (Kahneman, 2011). The exaggeration of events caused by the news can reinforce confirmation bias and result in faulty assumption of risks—discounting risks with higher probabilities over risks that are more easily available (Kahneman, 2011).

Kahneman's work has had significant influence on the finance industry in researching investment decisions, but using his concepts in the information security area, to the author's knowledge, has been limited. Building on Kahneman's (2011) heuristics and biases work in assessing risk, this study will also assess how cognitive bias may influence UXA decisions regarding their security threat and coping appraisals. Heuristics and biases associated with the dual-process theory are applicable to the information security realm given they may influence administrators' decisions. System One's automatic and intuitive assessments can be prone to error (Kahneman, 2011). The associative System One tends to best-guess answers to questions with available data even when that data is not completely relevant to the posed question (Pennycook et al., 2013; West et al., 2012). Kahneman (2011) referred to this as the shotgun effect. Another potential area of concern is the way that System One deals with ambiguities and competitive hypotheses eliminating options before cognitive awareness (Kahneman, 2003). When System Two is not actively monitoring System One, to assess the validity of the decisions, errors may occur (Kahneman, 2011; Toplak et al., 2011). Security risks and threat vectors in information security continue to evolve (Caballero, 2013). It is crucial to have competent and engaged administrators that acknowledge the threats and mitigate the risks (Caballero, 2013). Based on the prior research noted above, simple acceptance of System One responses may cause administrators to be oblivious of the threats and vulnerabilities they face. By better understanding what heuristics impact how decisions are made, what biases may result, and learning how to encourage System Two processing, organizations may be better prepared to manage their administrators and security risks (Kahneman, 2011; Pennycook et al., 2013; West et al., 2012).

An example of the potential use of the availability heuristic may be UXA incorrectly assessing a lower security risk to their servers because they more easily recall data breaches, security alerts, or reported fixes associated with Windows servers from the news. Without a correct assessment of potential risks, the UXA may not perceive the true threat severity or threat vulnerabilities to their servers. This bias can potentially result in insufficiently secured servers, leaving them at higher risk of breach. This failure may leave UNIX servers open to potential systems disruption, loss of proprietary or confidential data, harm to organizational reputation, potential loss of revenue, or financial loss due to litigation or fines (Donaldson et al., 2015). This research is needed as implementing HIPAA mandated security and compliance with organization ISP are key to the successful protection of healthcare organizational assets, including patient PII, and PHI (Koch, 2017; Ng et al., 2009).

Protection motivation theory

Several key theories have been used to assess and predict information security behaviors. Ng et al., (2009) integrated the Theory of Planned Behavior, Healthcare Behavioral Theory, and Expectancy-Value Theory to create their framework. Siponen, Pahnila, & Mahmood (2010) integrated Protection Motivation Theory (PMT), Deterrence Theory, Theory of Reasoned Action, Diffusion of Innovation Theory, and rewards theory to predict compliance behavior. Siponen et al., (2014) integrated the Theory of Planned Behavior, Theory of Reasoned Action, Cognitive Evaluation Theory, and PMT to understand users' intention to comply with policy. Ifinedo (2012) combined PMT and Theory of Planned Behavior to assess employee compliance with ISP.

Protection Motivation Theory is frequently used to understand compliance with ISPs and security procedures (Hanus & Wu, 2016; Safa, et al., 2015; Siponen, et al., 2014). PMT was initially proposed by Rogers (1975) to help understand how health behaviors were influenced by fear appeals. Rogers (1975) theorized that relevant sources of information that influenced behavioral change include environmental data (verbal persuasion and observational learning) and intrapersonal data (personality and prior experience). These sources are evaluated through a cognitive mediation process that assesses the threat and coping potential which leads to adaptive or maladaptive coping behaviors (Rogers & Prentice-Dunn, 1997). Threat appraisal includes the positive factors of extrinsic and intrinsic rewards offset by negative factors of the perceived severity and perceived vulnerability to potential threats (Posey et al., 2015; Rogers & Prentice-Dunn, 1997). The coping appraisal is an assessment of how the individual can cope with, adapt to, and change behavior to avoid the danger (Rogers & Prentice-Dunn, 1997). The factors related to coping appraisal include an individual's self-efficacy and response efficacy (Posey et al., 2015). Response efficacy is an evaluation of the effectiveness of the proposed behavior to reduce the probability of the negative event (Rogers & Prentice-Dunn, 1997). Self-efficacy is the belief that one is capable of the adaptation necessary to mitigate the negative event (Rogers & Prentice-Dunn, 1997). Fear influences the evaluation of severity and vulnerability and indirectly influences behavioral intention (Rogers & Prentice-Dunn, 1997). Siponen et al. (2014) found perceived severity of threat and perceived vulnerability to be positively correlated with ISP compliance. Appropriate threat appraisal can be manifest through increased knowledge and awareness of information security risks, vulnerabilities, and organizational policies and procedures (Albrechtsen & Hovden, 2010; Guo, et al., 2011; Safa & Von Solms, 2016; Siponen, et al., 2014). Coping appraisal can be also be positively influenced by information security knowledge sharing and awareness (Safa, et al., 2015; Siponen et al., 2010). Rogers and Prentice-Dunn (1997) noted that there are numerous cognitive heuristics and biases that can influence both appraisal processes in the PMT model. Cognitive heuristics can lead to cognitive biases and influence daily decision-making without our awareness (Kahneman, 2011). As such, it is important to understand how heuristics and biases may influence the threat appraisal of UXA. If threats are perceived as “Windows problems” the precognitive choice to resist ISP implementation on UNIX servers may put the organization at considerable risk.

In a study of home computer users, Hanus and Wu (2016) evaluated how awareness, a potential antecedent of desktop security behavior, influenced user’s security actions. Hanus and Wu (2016) extended PMT by defining the multi-dimensional construct of awareness (threat awareness and countermeasure awareness) to understand how awareness may influence desktop security behaviors. The goal of this study was to determine if threat awareness and countermeasure awareness, as antecedents of PMT, influence desktop security behavior (Hanus & Wu, 2016). Hanus and Wu (2016) demonstrated that research into the antecedents of PMT can provide a clearer picture into security behavior. Hanus and Wu (2016) identified both threat awareness and countermeasure awareness as key points for training to improve desktop security behavior. Hanus and Wu (2016) proposed that there may be a hidden variable that influences the relationship between threat awareness and perceived vulnerability that has not been discovered through existing research. Hanus and Wu (2016) demonstrate how PMT can be extended by including awareness. Awareness can be facilitated through effective training and knowledge management within an organization (Bauer et al., 2017; Dang-Pham et al., 2017).

Bélanger, Collignon, Enget, & Negangard (2017) in a study of early adopter password compliance, found that perceived threat severity and perceived threat vulnerability were positively related to attitude toward ISP change. Organizational triggers and ISP awareness had a positive correlation with attitude and intention to comply (Bélanger et al., 2017). Interestingly, Bélanger et al. (2017) found that subjective norm and self-efficacy did not significantly influence intention to conform to ISP for early adopters. While the research model was primarily based on the Theory of Planned Behavior their inclusion of perceived threat severity and perceived threat vulnerability from PMT helped to inform their model and, as their results indicated, were crucial to understanding the antecedents to attitude and intention (Bélanger et al., 2017).

Posey et al. (2015), investigated the impact of SETA programs on PMT. Specifically, Posey et al (2015), evaluated constructs frequently unused when applying PMT to information security contexts, including response costs, intrinsic and extrinsic maladaptive behaviors, and fear. The goal of the research by Posey et al. (2015) was to fully test PMT in an information security context adding SETA as an antecedent and organizational commitment as a moderating variable to better understand information security behavior. SETA was positively correlated with both perceived threat severity and perceived response efficacy indicating that SETA programs are an effective way of encouraging information security behavior (Posey et al., 2015). Appropriately designed SETA programs can help reduce the human information security risk to organizational assets (Van Vuuren, 2016; Whitman & Mattord, 2012). Posey et al. (2015) noted that SETA programs should be regular and recurring due to the dynamic landscape of information security.

Safa et al. (2015) conducted research to identify the factors that influence user’s information security conscious behavior by integrating the Theory of Planned Behavior (TPB), PMT, and adding potential antecedent factors of information security awareness, organizational policy, experience, and involvement. Security conscious behavior by users can help to mitigate information security risk (Safa et al., 2015). The study found that by increasing a user’s awareness of risks and vulnerabilities, improvements in attitude and information security conscious behavior can be achieved (Safa et al., 2015). Additionally, engaging users in the process of securing their systems, and educating them regarding potential threats improves threat appraisal and self-efficacy and positively influences information security conscious behavior (Safa et al., 2015). Knowledge sharing and collaboration are key components of this model in terms of information security awareness, organizational policy (and communication), involvement, engagement, and behavior (Safa et al., 2015). By combining TPB and PMT with additional antecedents Safa et al. (2015) helped to provide a better understanding of the need to consider behavioral and knowledge management aspects of information security.

One of the challenges in ISP compliance studies that use PMT noted by Boss, Galletta, Lowry, Moody, and Polak (2015) is the focus on intention rather than actual behavior. Boss et al. (2015) found few studies had been conducted that evaluated actual security behavior (Boss et al., 2015). Boss et al. (2015) conducted four field experiments to evaluate PMT. In one experiment, focused on the use of data backups, Boss et al. (2015) manipulated the fear-appeal construct and found in high-fear-appeal scenarios the significant influencers of intention and actual behavior were perceived threat severity, perceived threat vulnerability, fear, response efficacy, self-efficacy, and response costs. In the low-fear-appeal scenarios only perceived threat vulnerability and response costs influenced behavioral intention and actual compliance (Boss et al., 2015). In the high-fear-appeal manipulation, behavioral intention correlated with behavior at 0.710 while in the low-fear appeal manipulation the correlation was only 0.407 indicating that the high-fear-appeal was significantly more effective in encouraging actual behavior. For future research Boss et al. (2015) recommended manipulating the fear-appeal levels, using the full framework of PMT, as well as evaluating actual behavior not just behavioral intention.

Information security policy compliance

Although organizations spend considerable money on information security technology, users are still a major source of failures that result in information security breaches costing organizations substantial financial loss (Safa et al., 2016). ISPs elucidate the required security processes employees must follow to ensure the confidentiality, integrity, and availability of organizational information technology resources (D’Arcy & Lowry, 2019; Van Vuuren, 2016). ISPs include formalized procedures, guidelines, and technical controls that employees must follow to meet organizational security requirements (Cram, Proudfoot, & D'Arcy, 2017; Lowry & Moody, 2015). The human aspects of information security must be understood to reduce the risk of information security breaches (Van Vuuren, 2016). Users’ ignorance, apathy, resistance, and mischievous nature can result in human error and cause information security breaches (Bélanger et al., 2017; Safa et al., 2016). Compliance with ISP can help to mitigate information security risk (Ifinedo, 2014). Unfortunately, employee’s noncompliance with ISP is “the key threat” for organizational information security (Siponen et al., 2014, p. 217). Given the critical nature of organizational data and the significance of human behavior in protecting data, developing an understanding of what factors encourage and discourage ISP compliance will help to protect organizations (Bélanger et al., 2017; Carlton & Levy, 2015; Van Vuuren, 2016).

Siponen et al. (2014) combined PMT, the Theory of Reasoned Action (TRA), and Cognitive Evaluation Theory (CET) to develop an integrated theory that better explains end user ISP compliance. The following constructs were included: perceived severity, perceived vulnerability, response efficacy, and self-efficacy (from PMT); normative beliefs, intention to comply, and behavioral compliance (from TRA); and rewards (from CET) (Siponen et al., 2014). The model proposed by Siponen et al. (2014) accounted for 51% of the variance explained for intention to comply with ISP. Attitude (0.420) and normative beliefs (0.327) were the highest correlations with intention to comply with ISP (Siponen et al., 2014). This implies that organizational and relational factors do have a substantial impact on ISP compliance. The other factors, which were all significant, indicate that it is important to make sure employees are aware of security vulnerabilities and risks and trained on how to comply with ISP (Siponen et al., 2014). Knowledge management, in the form of knowledge sharing, training, and collaboration, all improved intention and behavioral compliance with ISP (Siponen et al., 2014). By making users aware of the vulnerabilities, risks, and potential impacts of information security breaches through effective SETA organizations can help to reduce noncompliance with ISP and improve information security.

Safa et al. (2016) sought to identify the factors that influence ISP attitude toward compliance through the lens of Social Bond Theory and Involvement Theory. Specifically, how does information security knowledge sharing, information security intervention, and collaboration influence attitude toward compliance (Safa et al., 2016). The factors Safa et al. (2016) identified through a literature review included: information security knowledge sharing, information security collaboration, information security intervention, information security experience, attachment, commitment, personal norms, attitude towards ISP compliance, and ISP compliance behavioral intentions. Information security knowledge sharing, collaboration, intervention, commitment, and personal norms all significantly influenced attitude toward compliance with ISP (Safa et al., 2016). While attachment was not a significant influencer of attitude toward compliance with ISP, attitude toward compliance with ISP was a significant predictor of ISP compliance behavioral intention (Safa et al., 2016). Information security knowledge sharing is key to improving user awareness and results in improved attitude towards ISP compliance and behavioral intention (Safa et al., 2016). Additionally, as socialization between the information security team and users was found to contribute to compliance behavior, organization managers should encourage information security knowledge sharing and development of robust intervention programs (Safa et al., 2016).

Nurse et al. (2014), in a study to understand insider threat, developed a framework for characterizing insider attacks that included a catalyst event, actor characteristics, attack characteristics, and organizational characteristics. To help identify the propensity to attack they combined elements of precipitating event, with the individual actor characteristics of psychological state, personality characteristics, historical behavior, attitude towards work, motivation to attack, skill set, and opportunity (Nurse et al., 2014). Also included in their model were observed physical and cyber security behaviors, enterprise role, type of actor, and state of relationship (Nurse et al., 2014). Attack characteristics included the type of attack, the object, the steps, and the step goals (Nurse et al., 2014). Lastly, the organizational characteristics included in the model developed by Nurse et al. (2014) included the type of asset and the perceived vulnerability. Nurse et al. (2014) applied the framework to three real-world cases of information security insider attacks. The framework Nurse et al. (2014) developed demonstrates the complexity of understanding ISP non-compliance by insiders. Nurse et al. (2014) showed the variety and significance of actor characteristics and how those characteristics interact with organizational characteristics to result in an information security attack.

In a detailed literature review regarding organizational ISP, Cram et al. (2017) developed a research framework to understand best practices for ISP development. The framework included five influencing relationships: (1) influences on design and implementation of ISP; (2) influence on employees and organization; (3) influence of employees and organization on ISP compliance; (4) influence of ISP compliance on organization's objectives; (5) and ISP maintenance (Cram et al., 2017). Key factors and commonly used theories were identified for each influential relationship and included: Systems theory, Grounded theory, Critical Social theory, Actor-network theory, Social Cognitive theory, Deterrence theory, Protection Motivation Theory, Theory of Planned Behavior, Rational Choice Theory, Control theory, and the Theory of Organizational learning (Cram et al., 2017). Several key findings of Cram et al. (2017) are relevant to the current research. First, Cram et al. (2017) identified the factors important for ISP design and maintenance. Second, Cram et al. (2017) demonstrated the importance of security awareness, clear consequences for non-compliance, as well as the influences of personality traits and culture on interpreting and complying with ISP. Finally, their research identified key areas of research that are needed for compliance to result in organizational security objectives (Cram et al., 2017).

It can be challenging to evaluate actual ISP compliance given the risk of social desirability bias that can occur in interviews and self-reported surveys (Redmiles, Acar, Fahl, & Mazurek, 2017). The use of scenarios and hypotheticals can help to remove the focus on the participant and may provide better insight into non-conforming behavior (Crossler et al., 2013). Rather than doing point-in-time studies, Crossler et al. (2013) suggested the use of longitudinal studies and field experiments can provide a much more valid picture of information security behaviors (Crossler et al., 2013). When evaluating intention to perform security behaviors, researchers are challenged to determine if participants have responded with over or under-reported counts as compared to actual behaviors (Egleman & Peer, 2015). As intention to comply with ISP can be a more socially desirable response, care must be taken to mitigate the risk of social desirability bias when designing research studies (Redmiles et al., 2017). Given these risks, it is preferred to study actual information security behavior rather than relying on self-reported intention to comply to truly understand the behaviors and develop valid and reliable behavioral models (Crossler et al., 2013).

THEORETICAL DEVELOPMENT/MODEL

Based on the literature review, there are several notable gaps that need further investigation. First, research on ISP compliance has focused primarily on end-user compliance intention (Albrechtsen & Hovden, 2010; Dang-Pham et al., 2017; Hanus & Wu, 2016; Ifinedo, 2012; Safa et al., 2016). It is crucial, however, to understand the determinants of server administrator ISP compliance behavior to better ensure the confidentiality, integrity, and availability of PII and PHI data contained on their servers.

Second, this study will focus on formal knowledge sharing arrangements from the security knowledge sharing framework developed by Flores et al. (2014). Specifically, this research will investigate how SETA workshops, designed specifically for server administrators, affect information security knowledge sharing and ISP compliance behavior. Concentrating on this specific user group for education is unique and will help to evaluate the effectiveness of workshops that use server administrator focused scenarios.

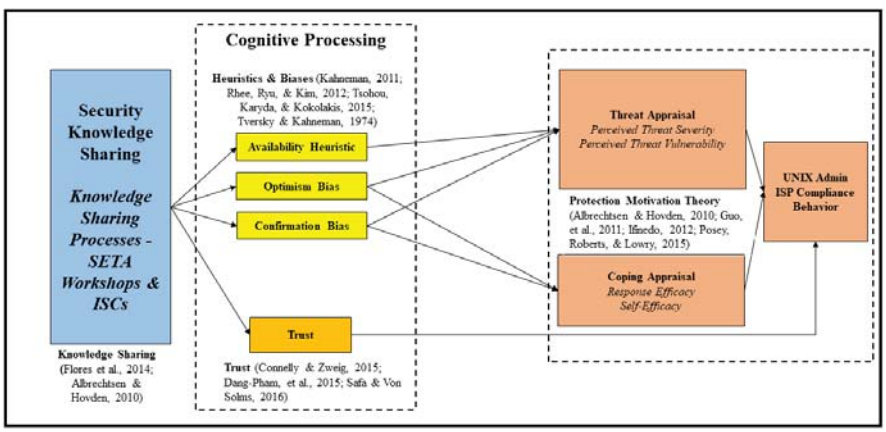
Third, trust is a crucial precursor to information security awareness, knowledge sharing, risk perception, and can be effective in reducing non-compliance (Lowry et al., 2015; Moqbel & Bartelt, 2015; Van Vuuren, 2016). It is therefore important to understand how the trust the UXA have in the information security team influences the effectiveness of the SETA programs and ISP compliance behavior. Ideally, the SETA workshops will improve the relationships between the UXA and information security team resulting in increased ongoing communication, improved awareness of threats, and ultimately enhanced compliance behavior.

Fourth, cognitive heuristics can lead to biased decision making and biased evaluation of information security threats and risks (Pachur et al., 2012; Rhee et al., 2012; Toplak et al., 2011; Tsohou et al., 2015). The inappropriate use of heuristics can prevent UXA from honestly assessing the true risks that may impact their servers potentially reducing ISP compliance behavior. SETA programs can be an effective means of reducing bias and improving security behaviors (Rhee et al., 2012). While Rogers and Prentice-Dunn (1997) mention that cognitive biases may influence both assessment processes in PMT, they did not introduce the construct in their model, and did not tie it to PMT threat appraisal. Through the SETA program, awareness should increase, encouraging a more appropriate understanding of the vulnerabilities that exist for their servers. It is crucial to minimize the use of cognitive heuristics in order to minimize biased decisions and improve appropriate threat assessment and response appraisal. The development of server administrator-oriented SETA workshops may improve awareness, reduce UXA use of cognitive heuristics, and improve threat assessment, coping appraisal, and ISP compliance behavior.

Fifth, PMT has been established as an effective model for evaluating ISP compliance intention and behavior (Boss et al., 2015; Johnston, Warkentin, & Siponen., 2015). Threat appraisal and coping appraisal have been found to be positively influenced by SETA programs (Posey et al., 2015). They have also been established as precursors of compliance intention and behavior (Boss et al., 2015; Siponen et al., 2014). While Posey et al. (2015) are the only researchers identified that have integrated SETA and PMT, their focus was limited to end-user behavioral intention. Also, Posey et al. (2015) did not include the influences of trust and heuristics and biases or SETA effectiveness.

Finally, while behavioral intention is frequently taken as an indicator of behavior, it was noted by Crossler et al. (2013), that actual behavior may not follow reported intention due to social desirability biased self-reporting. Boss et al. (2015) also found that intention differed significantly from actual implementation of security controls. This study will investigate actual UXA ISP compliance behavior by running and analyzing credentialed security scans on servers. Using credentialed security scans of security measures implemented by UXA following the SETA workshops will afford unique insights into the effectiveness of the training in terms of actual implementation of security controls and ISP compliance behavior.

The proposed research model, based on the literature review, can be found in Figure 1. It integrates the security knowledge sharing portion of the Information Security Organizational Knowledge Sharing Framework (Flores et al., 2014), cognitive heuristics and biases (Kahneman, 2011; Tversky & Kahneman, 1974; Tsohou et al., 2015), trust (Connelly & Zweig, 2015; Dang-Pham et al., 2016; Safa & Von Solms, 2016), and Protection Motivation Theory (PMT) (Albrechtsen & Hovden, 2010; Guo et al., 2011; Ifinedo, 2012; Posey et al., 2015).



Information Security Policy Cognitive Processing Model

Research Design

This quantitative research will be conducted in a pretest and posttest experimental and control group design. In this method, both the control group and the experimental group complete pretest and posttest surveys (Sekaran & Bougie, 2013). The experimental group will participate in three new SETA workshops. The independent variable for this model is security knowledge sharing from the Information Security Knowledge Sharing Framework developed by Flores et al. (2014). Security knowledge sharing is operationalized through SETA workshops designed specifically for the server systems administrators. Trust is integrated into the model as it has been demonstrated to influence the effectiveness of security education (Van Vuuren, 2016). The trust of the UXA in the information security team may influence ISP compliance. Security education has been found to positively influence threat and coping appraisal from PMT (Safa et al., 2015; Woon, Tan, & Low, 2005). The use of the availability heuristic, however, can negatively influence correct assessment of risk (Hertwig, Pachur & Kurzenhäuser, 2005; Pachur et al., 2012). The affect heuristic, optimism bias, and confirmation bias can influence both threat and coping appraisal (Tsohou et al., 2015). The inappropriate use of cognitive heuristics may result in biases which influence UXA security threat and coping appraisals and reduce overall ISP compliance (Kahneman, 2011; Pachur et al., 2012). Education has been demonstrated to influence the use of heuristics and biases and improve estimation of risk (Kahneman, 2003). The SETA workshops may reduce the use of inappropriate heuristics, afford a more appropriate estimate of threat to UNIX servers, and improve ISP compliance. PMT has been used extensively to describe the influence of threat and coping appraisal on security intention and behavior (Hanus & Wu, 2016; Safa, et al., 2015). Posey et al. (2015) demonstrated that SETA programs influenced the constructs of PMT (threat assessment and coping appraisal).

Variables

Security Knowledge Sharing - SETA Workshops

Security workshops will be designed that are appropriate to the job functions of UNIX and Windows server administrators (Beuchelt, 2013; Inshanally, 2018; US Department of Health and Human Services, 2013). The organizational Information Security and Risk Management Plan (ISRMP) will be used to prioritize the top eight behavioral information security risks facing the organization that are relevant to server administration. This document prioritizes risk as very high, high, moderate, and low and prioritizes risks based on the criticality consistent with the NIST Cyber Security Framework (NIST, 2018). By identifying and prioritizing the risks, specific behaviors can be targeted that are evaluated as having the greatest benefit to the organization and mitigate the greatest amount of risk. The specific behaviors will be matched to governmental recommendations to ensure compliance with legal requirements (Avancha et al., 2012; Koch, 2016; US Department of Health and Human Services, 2013). Workshop format will be constructed based on prior research as well as the SANS standards for developing an integrated security training, awareness, and education program (Wilson & Hash, 2003).

The SETA program will be developed and conducted as a series of small-group security workshops that include information security team members, UXA, and Windows administrators within the CIT group of the organization (Albrechtsen & Hovden, 2010). The goal of the workshops will be to encourage connection, collaboration, discussion, and information security knowledge sharing between the participants. After a scenario is reviewed during the workshop participants and discussion has completed the information security team members may suggest additional ideas, tools, or processes that had not been considered. Additionally, online learning tools will be utilized between workshops to reinforce skills learned during the workshops.

Trust

Trust is a measure of the confidence that the UXA have in the competence, reliability, dependability, and responsibility of the information security team (McAllister, 1995). Without an adequate level of trust the recommendations and warnings from the security team may be dismissed resulting in ISP non-compliance (Dang-Pham et al., 2016; Safa & Von Solms, 2016). Trust will be assessed through use of survey questions based on prior research.

Cognitive Heuristics

The cognitive heuristics evaluated in the present research include the availability heuristic, optimism bias, and confirmation bias. The use of these heuristics can influence UXA estimation of risk and vulnerability associated with their UNIX servers (Kahneman, 2011; Pachur et al., 2012; Tversky & Kahneman, 1974). Additionally, optimism bias and confirmation bias may influence UXA coping appraisal. Questions and scenarios to assess the use of the heuristics and biases will be developed based on prior research.

Threat Appraisal

Threat appraisal is an evaluation of the UXA perceived severity and perceived vulnerability to information security threats (Posey et al., 2015; Rogers & Prentice-Dunn, 1997). Perceived vulnerability and perceived severity may be influenced by cognitive heuristics which may be influenced by an increased knowledge and awareness of information security risks and vulnerabilities (Albrechtsen & Hovden, 2010; Guo, et al., 2011; Safa & Von Solms, 2016; Siponen, et al., 2014). Perceived severity and perceived vulnerability will be assessed via survey.

Coping Appraisal

Coping appraisal is an assessment of how the UXA perceives that they can cope with, adapt to, or mitigate the information security risk (Rogers & Prentice-Dunn, 1997). Coping appraisal considers the UXA's information security self-efficacy and information security response efficacy (Posey et al., 2015). Response efficacy is an assessment of how effective the proposed behavior can reduce the probability of the negative event (Rogers & Prentice-Dunn, 1997). Self-efficacy is the belief that one can make the changes needed to mitigate the risk (Rogers & Prentice-Dunn, 1997). Self-efficacy and response efficacy will be assessed via survey.

ISP Compliance Behavior

UXA ISP compliance behavior will be assessed quantitatively. A random sample of UNIX servers managed by experimental group participants will be evaluated using Tenable Nessus vulnerability management tool. The total number of critical, high, medium, and low vulnerabilities will be assessed prior to and three months following the SETA workshops. Tenable scans of servers provide a detailed report of discovered vulnerabilities related to the server operating system, networking, and applications and includes a synopsis, description, solution, risk factor, Common Vulnerability Scoring System (CVSS) base score, CVSS temporal score, references, and output for each vulnerability (Tenable, 2017). Additionally, the CVSS total score, will be used to evaluate the implementation of specific security countermeasures presented during the SETA workshops. CVSS is a common security metric used to evaluate information security (Hayden, 2010). Using credentialed security scans of security measures implemented by UXA following the SETA workshops will afford unique insights into the effectiveness of the training in terms of actual implementation of security controls and ISP compliance behavior. This research is needed as implementing mandated security and compliance with organization ISP are key to the successful protection of healthcare organizational assets, including patient PHI and PII (Koch, 2017; Ng et al., 2009).

Control variables

Age, gender, experience, and education will be used as control variables in the present research.

Population and Sampling

The population for this research is made up of the UXA in a major university and hospital system in the mid-Atlantic United States. The CIT organization manages 600+ UNIX physical and logical servers located in three data centers in two states as well as the District of Columbia. The administrators are responsible for the servers running enterprise wide applications including the electronic medical record system, pathology labs, radiology, web services, information security servers, student information systems, precision medicine systems, document management system, change control systems, as well as numerous departmental systems that house PHI and PII. Presently there are eight UXA in CIT and 51 other UXA in other teams in the organization.

Sampling Method

A listing of all active UXA will be obtained from the information security manager within the CIT organization. Administrators will be assigned to two groups using convenience sampling (for group 2) and random sampling (for group 1). Group one will be the control group and will consist of eight UXA outside of the CIT organization that will complete the existing organizational online information security compliance education module. The existing training is a ten-minute general information security online learning module required for all organization members. Group two, the experimental group, will participate in three newly designed SETA workshops developed specifically for UXA.

Subjects

The CIT UXA team is made up of individuals with a wide variety of technical experience and educational background. The team consists largely of senior administrators with 75% of UXA having 10 or more years of experience. In terms of educational attainment, 50% of the UXA have completed a college degree. With regard to gender the CIT UNIX team is 100% male. This is not unusual given the gender imbalance noted in IT (Gorbacheva, Beekhuyzen, vom Brock, & Becker, 2019). Additionally, this reflects the gender representation in the population of UXA in the institution.

Data collection techniques

Two types of data will be collected. First, a survey instrument will be used that integrates previously validated questions from prior empirical research. The survey will be developed and administered using Qualtrics. The survey instrument will assess the following constructs: trust, self-efficacy, response-efficacy, perceived vulnerability, perceived severity, information security knowledge sharing, as well as the availability heuristic, optimism bias, and confirmation bias. The second set of data points will be the analysis of security controls implemented on a sampling of servers prior to and three-months following the security workshops and online virtual modules. While a great deal of security research has focused on behavioral intention it was noted in the literature review that this is not always a true assessment of information security behavior (Crossler et al., 2013). To mitigate this possible response bias, credentialed Tenable Nessus scans will be run on a sample of UNIX servers prior to and three-months following the workshops to quantitatively determine if the administrator implemented security changes. The number of critical, high, medium, and low vulnerabilities and the total Common

Vulnerability Scoring System (CVSS) score will be used to assess compliance changes made in the three-months following the SETA training. Last, baseline data will be collected for the following: third-party and core operating systems levels (Brotherston & Berlin, 2017; CIS, 2018; Hayden, 2010); root login capabilities (Beuchelt, 2013); password settings (Beuchelt, 2013; Brotherston & Berlin, 2017); user ids with empty passwords (Beuchelt, 2013; Brotherston & Berlin, 2017); use of multifactor authentication (CIS, 2018); status of SELinux (Beuchelt, 2013; CIS, 2018); status of central log management (CIS, 2018); all running services (Beuchelt, 2013; Brotherston & Berlin, 2017; CIS, 2018); and status of local firewall (Beuchelt, 2013; CIS, 2018).

Safa and Von Solms (2016) performed confirmatory factor analysis to test their measured variables. Convergent validity was evaluated with factor loading and as a result three items were dropped (Safa & Von Solms, 2016). Composite reliability was confirmed with Cronbach's alpha and discriminant validity tests verified the independence of the remaining constructs (Safa & Von Solms, 2016). Moqbel and Bartelt (2015) assessed discriminant and convergent validity for their survey instrument by performing factor analysis and tested composite reliability by examining Cronbach's alpha. All results exceeded the recommended statistical thresholds (Moqbel & Bartelt, 2015). Moqbel and Bartelt (2015) used the square root of average variance extracted (AVE) to assess discriminant validity for all their constructs. Ifinedo (2014) examined the convergent validity, internal consistency, and discriminant validity of their survey instruments. Composite reliability was confirmed via factor loading tests (Ifinedo, 2014). Convergent validity was verified using AVE (Ifinedo, 2014). Finally, Ifinedo (2014) confirmed discriminant validity by assessing the AVE and the square root of AVE. Hanus and Wu (2016) evaluated their model using partial least squares structural equation modeling (PLS-SEM). Exploratory factor analysis resulted in a total of 37 factors in their survey instrument (Hanus & Wu, 2016). Reliability of the model was confirmed with composite reliability and Cronbach's alpha (Hanus & Wu, 2016). Convergent validity was confirmed using AVE and discriminant validity was confirmed by testing cross loadings (Hanus & Wu, 2016). Siponen et al. (2014) tested convergent validity by factor loading and AVE which resulted in six items being dropped. Discriminant validity was evaluated by assessing correlations between all constructs (Siponen et al., 2014). Additionally, the variance extracted, and the square root of variance extracted met the corresponding statistical recommendations (Siponen et al., 2014). Reliability was evaluated using Cronbach alpha and all items exceeded the recommended levels (Siponen et al., 2014). Safa et al. (2016) tested convergent validity by using factor loading resulting in two items being removed from their survey instrument. Internal consistency was evaluated Cronbach's alpha and all factors exceeded the recommended threshold (Safa et al., 2016). Discriminant validity was evaluated using the square root of variance extracted and again all items exceeded the recommended thresholds (Safa et al., 2016).

Following are the survey instrument items for each construct in the model. Modifications to existing, previously validated instruments are indicated by italics. Unless otherwise noted, all items are based on a 7-point Likert scale ranging from (1) strongly disagree to (7) strongly agree which is consistent with the prior instrument's usage.

Trust

In Safa and Von Solms (2016) instrument, "colleague's" was replaced with "information security team's" to evaluate the UXA trust in the information security team. In Moqbel and Bartelt's (2015) instrument, the phrase "personal cloud computing companies" was replaced with "the information security team members" to assess the UXA trust in the information security team.

TR1: I believe that my *information security team's* information security knowledge is reliable (Safa & Von Solms, 2016).

TR2: I believe that my *information security team's* information security knowledge is effective (Safa & Von Solms, 2016).
TR3: I believe that my *information security team's* information security knowledge mitigates the risk of information security breaches *for my servers* (Safa & Von Solms, 2016).

Self-Efficacy

For Ifinedo's (2014) instruments the references to "computer" were changed to "servers" to reflect the focus on systems administrator's enterprise servers.

SE1: I have the necessary skills to protect my *servers* from information security threats (Ifinedo, 2014).
SE2: I have the expertise to implement preventative measures to stop people from getting my organization's confidential information stored on my *servers* (Ifinedo, 2014).
SE3: I have the skills to implement preventative measures to stop people from damaging my *servers* (Ifinedo, 2014).

Response Efficacy

For Ifinedo (2012) and Hanus and Wu's (2016) instruments the references to "computer" were changed to "servers" to reflect the focus on systems administrator's enterprise servers. The addition of patient information to RE2 reflects the focus of the present research in the context of an academic teaching hospital.

RE1: Enabling the security measures on my *servers* is an effective way to deter hacker attacks (Ifinedo, 2012).
RE2: The preventative measures available to me to stop people from gaining access to my organization's *servers* and data are adequate (Ifinedo, 2012).
RE3: Frequently applying security patches on my operating system is an effective way of preventing hacker attacks on my *servers* (Hanus & Wu, 2016).

Perceived Vulnerability

For Siponen et al. (2014) and Hanus and Wu (2016) the references to "computer" were changed to "servers" to reflect the focus on systems administrator's enterprise servers. VU3 was extended to include additional types of information security threats.

VU1: My *servers* could be subjected to a serious information security threat (Siponen et al., 2014).
VU2: My organization could be subjected to a serious information security threat (Siponen et al., 2014).
VU3: I believe that trying to protect my company's *servers* and information will reduce illegal access to it (Ifinedo, 2012).

Perceived Severity

For Siponen et al. (2014) and Hanus and Wu (2016) the references to "computer" were changed to "servers" to reflect the focus on systems administrator's enterprise servers. SV4 was expanded to include a wider range of information security threats.

SV1: An information security breach in my organization would be a serious problem for me (Siponen et al., 2014).
SV2: An information security breach in my organization would be a serious problem for my organization (Siponen et al., 2014).
SV3: I believe that having my *servers* infected malware, virus, or similar infection would be a serious problem (Hanus & Wu, 2016).

Information Security Knowledge Sharing

SKS1: I frequently share my information security knowledge in my team in order to decrease information security risk (Safa et al., 2016).
SKS2: I think information security knowledge sharing helps me to understand the usefulness of information security policies in my organization (Safa et al., 2016).
SKS3: I think information security knowledge sharing is a valuable practice in my organization (Safa et al., 2016).

Cognitive Heuristics

New questions will be developed to evaluate the use of the availability heuristic, representativeness heuristic, affect heuristic, optimism bias, and confirmation bias. As the use of heuristics is a function of System One, a time limit will be used for all the cognitive heuristics questions to minimize the use of analytical (System Two) reasoning (Finucane et al., 2000; Gertner, Zaromb, Roberts, and Matthews, 2016; Kahneman, 2011). The availability heuristic impacts the estimation of probabilities and participants will be asked to rank common operating systems and applications vulnerabilities and those responses will be compared to actual statistics (Kahneman, 2011; Kliger & Kudryavtsev, 2010; Tversky & Kahneman, 1974). Confirmatory bias will be tested using a fictional scenario, similar to the technique of Fischer et al. (2011) where participants are presented a scenario, asked to make an initial decision, then provided six confirming and six disconfirming bits of additional information they can choose to review, and then asked to choose again. The level of confirmation bias will be determined by subtracting the number of disconfirming from the confirming choices selected (Fischer et al., 2011; Gertner et al., 2016). To evaluate representativeness bias, Gertner et al. (2016) used scenarios that encouraged base rate neglect where subjects are provided statistical information and then counterintuitive data. Similarly, the present research will utilize a security related scenario to evaluate use of the representativeness heuristic. Rhee et al. (2012) found that optimism bias was based on a measure of risk perception and controllability. To develop a quantitative measure for optimism bias, the results of the perceived vulnerability (VU1 and VU3), self-efficacy (SE2 and SE3), and response efficacy (RE3) questions will be used. These items indicate complete confidence that all security threats will be prevented (SE2, SE3, and RE3) and focus specifically on the administrator's servers (VU1 and VU3). The affect heuristic will be evaluated using questions associated with the administrator's perceptions (feeling) of their value to the organization as well as their feelings regarding the organizational ISP. Administrators who feel unvalued by the organization may be less prone to comply with ISP as the organization and ISP are viewed negatively—Slovic, Peters, Finucane, and MacGregor (2005) identified this is a high risk, low benefit affect-based decision. All cognitive heuristic based questions will be pilot tested with non-participants to ensure clarity. Additionally, the questions will be discussed with an expert in the field of cognitive decision making to assess the reliability and validity of the questions. Participant's responses to AV1-AV4 will be compared to the CVSS ratings which indicate actual vulnerabilities.

Availability heuristic questions:

AV1: Microsoft servers have more vulnerabilities than Linux/UNIX servers.
AV2: There are more security vulnerabilities, alerts, and patches related to Windows servers than Linux/UNIX servers.
AV3: A Windows server containing PII/PHI is likely to be breached.
AV4: A Linux/UNIX server containing PII/PHI is likely to be breached.

Optimism bias questions:

OPT1: Our organization has the means to control information security threats.
OPT2: Our organization has the ability to execute security practices to avoid information security threats.
OPT3: The likelihood that my servers will be disrupted due to information security breaches in the next 12 months is low.

Confirmatory bias scenario:

CB Screen 1: Your organization plans on implementing a new web server to provide customers' access to HIPAA protected PII data. In order to provide the highest level of security, would you recommend the web server be implemented on the Windows or UNIX/Linux operating system?
CB Screen 2: On the next screen you will be presented with additional data points you can choose to review to inform your decision. You will have 45 seconds to select as many items for review as you would like. You must choose at least one item. Each review costs you one virtual dollar. You have a total of $12 available. Once you have completed reviewing additional information please select again the operating system you would recommend for the web server. Please select continue to move to the next page.
CB Screen 3: Informational data points are below. Once you have reviewed any you wish to view you please select the operating system you would recommend.
L1 (+): Administrators benefit from the freedoms offered by the system administration methods in Linux (IONOS, 2019).
L2 (-): Porting of applications for Linux distributions is not the focus of many hardware and software companies (IONOS, 2019).
L3 (+): There are fewer demands on the hardware due to reduced operating systems overhead in Linux (IONOS, 2019).
L4 (-): Several professional programs (i.e. Microsoft Windows, Microsoft SharePoint, Microsoft Visio) do not work with Linux (IONOS, 2019).
L5 (+): Remote function access is integrated into the native operating system on Linux distributions (shell and terminal) (IONOS, 2019).
L6 (+): Licensing costs are minimal.
W1 (+): Windows server is new administrator friendly due to the intuitive operations of the graphical user interface (IONOS, 2019).
W2 (-): The licensing costs for Windows can be high and can increase with each user (IONOS, 2019).
W3 (+): Windows is compatible with popular Microsoft programs like SharePoint and Exchange (IONOS, 2019).
W4 (-): Windows servers are vulnerable to malware (IONOS, 2019).
W5 (-): The use of mandatory graphical user interface on Windows servers results in significant resource utilization for basic operating systems function (IONOS, 2019).
W6 (-): Windows servers are prone to user errors and security attacks through the integrated interface (IONOS, 2019).

Data Analysis Strategies

To test the structural model, partial least squares structural equation modeling (PLS-SEM) is frequently used (Hanus & Wu, 2016; Ifinedo, 2012; Safa & Von Solms, 2016). This method will allow for the examination of the paths of the model as well as the relationships between the variables (Hanus & Wu, 2016; Safa & Von Solms, 2016). Chi-square with degrees of freedom can be used to evaluate the global fit of the model (Safa & Von Solms, 2016). Goodness of fit index and adjusted goodness of fit index can be used to evaluate the fit between predicted values and observed values (Safa & Von Solms, 2016). To test the measurement model, confirmatory factor analysis (CFA) will be used to test the measured values against the constructs (Safa & Von Solms, 2016). CFA will be used to evaluate discriminant and convergent validity (Boss et al., 2015). Convergent validity can also be verified using average variance extracted (Hanus & Wu, 2016; Ifinedo, 2012; Posey et al., 2015). Cronbach's alpha will be used to evaluate internal consistency (Safa & Von Solms, 2016). Descriptive statistics (mean, standard deviation, skewness, and kurtosis) will be calculated for all of the constructs. Finally, paired t-tests will be used to help evaluate the influence of the SETA programs on UXA ISP compliance for the experimental group.

Implementation Plan

This survey will be administered electronically through the company's email and intranet. It will be sent to the UNIX systems administrators in the CIT group. Additionally, baseline statistics will be developed via the Tenable Nessus scans. The data point collected include: percentage of administrator's production servers sending data to centralized log management system (Splunk); percentage of administrator's production servers performing Nessus vulnerability scans; percentage of administrator's production servers performing credentialed Nessus vulnerability scans; percentage of administrator's production servers that have implemented kernel hardening; percentage of administrator's production servers with kernel at recommended patch level; percentage of administrator's production servers running antivirus software; and percentage of administrator's production servers that have implemented ISP password rules.

During the second phase of the research a set of security scenarios applicable to the business environment and related to server management will be developed. These will be utilized during a SETA program. The SETA program will be developed and conducted as a series of small-group security workshops that include information security team members and UXA within an organization. The goal of the workshops will be to encourage connection, collaboration, discussion, and information security knowledge sharing between the participants. After a scenario is reviewed with the workshop participants and discussion has completed the information security team members may suggest additional ideas, tools, or processes that had not been considered. Also, during this phase, online learning modules will be provided to participants in the experimental group to reinforce skills learned during the SETA workshops.

At the conclusion of phase three the initial questionnaire will be distributed to the UXA to evaluate any significant changes in scoring for any of the model's factors. UXA ISP compliance behavior will be assessed by analyzing seven percentages for specific security implementations for all the production UNIX servers managed by each UXA. Percentages are frequently used for analyzing and presenting security metrics as they are easily interpretable and can clearly indicate positive or negative change (Brotby & Hinson, 2013, Hubbard & Seirsen, 2016; Hayden, 2010, Jaquith, 2007). Three months following the workshops, Tenable Nessus scans will be used on the same data points to evaluate actual compliance with ISP. Discriptive statistics will be used to determine if significant changes were made.

References

Ajami, S., & Bagheri-Tadi, T. (2013). Barriers for adopting electronic health records (EHRs) by physicians. *Acta Informatica Medica*, *21*(2).

Albrechtsen, E., & Hovden, J. (2010). Improving information security awareness and behavior through dialogue, participation, and collective reflection. An intervention study. *Computers & Security, 29*(4), 432-445.

Amarasingham, R., Plantinga, L., Diener-West, M., Gaskin, D. J., & Powe, N. R. (2009). Clinical information technologies and inpatient outcomes: A multiple hospital study. *Architecture and Internal Medicine, 169*(2), 108-114.

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. 21-29. Proceedings of the 5th ACM Conference on Electronic Commerce, New York, 21-29.

Avancha, S., Baxi, A., & Kotz, D. (2012). Privacy in mobile technology for personal healthcare. *ACM Computing Surveys, 45*(1), 3:1-3:54.

Bajgoric, N. (2006). Information technologies for business continuity: An implementation framework. *Information Management & Computer Security, 14*(5), 450-466.

Bauer, S., & Bernroider, E.W.N. (2017). From information security awareness to reasoned compliant action: Analyzing information security policy compliance in a large banking organization. *DATA BASE for Advances in Information Systems, 48*(3), 44-68.

Bauer, S., Bernroider, E. W., & Chudzikowski, K. (2017). Prevention is better than cure! Designing information security awareness programs to overcome users' non-compliance with information security policies in banks. *Computers & Security, 68*, 145-159.

Bélanger, F., Collignon, S., Enget, K., & Negangard, E. (2017). Determinants of early conformance with information security policies. *Information & Management, 54*(7), 887-901.

Belsis, P., Kokolakis, S., & Kiountouzis, E. (2005). Information systems security from a knowledge management perspective. *Information Management & Computer Security, 31*(3), 189-202.

Bernstein, B. E., & Hartsell, T. L. (2013). Federal Regulations, HIPAA, and HITECH. In *The Portable Lawyer for Mental Health Professionals: An A-Z Guide to Protecting Your Clients, Your Practice and Yourself* (3rd ed., pp. 357–446). Hoboken, NJ: John Wiley & Sons.

Beuchelt, G. (2013). UNIX and linux security. In J.R. Vacca (Ed.), *Computer and Information Security Handbook* (2nd ed, pp. 165-181). Boston, MA: Morgan Kaufmann.

Boss. S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., & Polak, P. (2015). What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly, 39*(4), 837-864.

Brotby, W. K., & Hinson, G. (2013). *PRAGMATIC security metrics: Applying metametrics to information security.* Boca Raton, FL: CRC Press.

Brotherston, L., & Berlin, A. (2017). *Defensive Security Handbook.* Boston: O'Reilly.

Burns, A. J., Roberts, T. L., Posey, C., Bennett, R. J., & Courtney, J. F. (2015). *Assessing the role of security education, training, and awareness on insiders' security related behavior: An expectancy theory approach*. Proceedings of the IEEE 48th Hawaii International Conference on Systems Sciences, HI. Doi:10.1109/HICSS.2015.471.

Caballero, A. (2013). Information security essentials for IT managers: Protecting mission-critical systems. In J.R. Vacca (Ed.), *Computer and Information Security Handbook* (2nd ed, pp. 379-407). Boston, MA: Morgan Kaufmann.

Carlton, M., & Levy, Y. (2015). *Expert assessment of the top platform independent cybersecurity skills for non-IT professionals*. Proceedings of the IEEE Southeast Con, Fort Lauderdale, FL. doi: 10.1109/SECON.2015.7132932.

Chen, X., Chen, L., & Wu, D. (2018). Factors that influence employees' security policy compliance: An awareness-motivation-capability perspective. *Journal of Computer Information Systems, 58*(4), 312-324.

CIS. (2018). *CIS Controls V7.1*. East Greenbush, NY: Center for Internet Security.

Cloud Market. (2019, May 4). EC2 Statistics: Total images available. Retrieved from https://thecloudmarket.com/stats.

Colwill, C. (2009). Human factors in information security: The insider threat – who can you trust these days? *Information Security Technical Report, 14*(4), 186-196.

Connelly, C. E., & Zweig, D. (2015). How perpetrators and targets construe knowledge hiding in organizations. *European Journal of Work and Organizational Psychology, 24*(3), 479-489.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security, 32*(C), 90-101.

Cram, W. A., Proudfoot, J. G. & D'Arcy, J. (2017) Organizational information security policies: A review and research framework. *European Journal of Information Systems*, *26*(6), 605-641.

CVE Details. (2019). Top 50 products by total number of "distinct" vulnerabilities. Retrieved from https://www.cvedetails.com/top-50-products.php.

Dang-Pham, D., Pittayachawan, S., & Bruno, V. (2016). Exploring behavioral information security networks in an organizational context: An empirical case study. *Journal of Information Security and Applications, 34*(1), 46-62.

Dang-Pham, D., Pittayachawan, S., & Bruno, V. (2017). Why employees share information security advice? Exploring the contributing factors and structural patterns of security advice sharing in the workplace. *Computers in Human Behavior, 67*, 196-206.

D'Arcy, J., & Lowry, P. B. (2019). Cognitive-affective drivers of employees' daily compliance with information security policies: A multilevel longitudinal study. *Information Systems Journal, 29*(1), 43-69.

Dey, T., & Mukhopadhyay, S. (2018). Influence of behavioral intentions, affective trust and affective commitment on knowledge sharing behavior. *International Journal of Knowledge Management, 14*(2), 37-51.

Dismukes, K., Berman, B. A., & Loukopoulos, D. (2007). *The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents*. New York, NY: Routledge.

Dixon, B. E. (2016). What is health information exchange? In B.E. Dixon (Ed.), *Health Information Exchange: Navigating and Managing a Network of Health Information Systems.* Boston, MA: Academic Press.

Donaldson, S. E., Siegel, S. G., Williams, C. K., & Aslam, A. (2015). Building an effective defense. In *Enterprise Cybersecurity: How to Build a Successful Cyberdefense Program Against Advanced Threats* (pp. 133-156). New York, NY: Springer.

Egleman, S., & Peer, E. (2015). Scaling the security wall: Developing a security behavior intentions scale (SeBIS). Proceedings of the CHI Conference, 2015, Seoul, Korea.

Epic. (2018). Epic operating system settings and security setup and support guide. Retrieved from https://galaxy.epic.com/?#Browse/page=1!68!50!2220559.

Epstein, S. (2014). *Cognitive-experiential theory: An integrative theory of personality*. New York: Oxford University Press.

Ferreira, M. B, Garcia-Marques, L., Sherman, S. J., & Sherman, J. W. (2006). Automatic and controlled components of judgement and decision making. *Journal of Personality and Social Psychology, 91*(5), 797-813.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgment of risks and benefits. *Journal of Behavioral Decision Making, 13*(1), 1–17.

FIRST.Org. (n.d.a). *Common Vulnerability Scoring System v3.0: Specification Document*. Retrieved April 21, 2019, from: https://www.first.org/cvss/cvss-v30-specification-v1.8.pdf.

FIRST.Org. (n.d.b). *Common Vulnerability Scoring System v3.0: User Guide*. Retrieved April 21, 2019, from: https://www.first.org/cvss/cvss-v30-user_guide_v1.6.pdf.

Fischer, P., Kastenmüller, A., Greitemeyer, T., Fischer, J., Frey, D., & Crelley, D. (2011). Threat and selective exposure: The moderating role of threat and decision context on confirmatory information search after decisions. *Journal of Experimental Psychology: General, 140*(1), 51-62.

Flores, W. R., Antonsen, E., & Ekstedt, M. (2014). Information security knowledge sharing in organizations: Investigating the effect of behavioral information security governance and national culture. *Computers & Security, 43*, 90-110.

Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security, 31*(8), 983-988.

Gardner, B., & Thomas, V. (2014). *Building an information security awareness program: Defending against social engineering and technical threats.* New York, NY: Elsevier.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly, 27*(1), 51-90.

Gertner, A., Zaromb, F., Roberts, R. D., & Matthews, G. (2016).*The assessment of biases in cognition* (MITRE Technical Report, Case Number 16-0956). Retrieved from https://www.mitre.org/sites/default/files/publications/pr-16-0956-the-assessment-of-biases-in-cognition.pdf.

Gilovich, T., & Griffin, D. (2013). Introduction – heuristics and biases: Then and now. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The Psychology of Intuitive Judgement* (pp. 1-18). New York: Cambridge University Press.

Gorbacheva, E., Beekhuyzen, J., vom Brock, J., & Becker, J. (2019). Directions for research on gender imbalance in the IT profession. *European Journal of Information Systems, 28*(1), 43-67.

Guo, K. H., Yuan, Y., Archer, N. P., & Connelly, C. E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems, 28*(2), 203-236.

Hanus, B., & Wu, Y. (2016). Impact of users' security awareness on desktop security behavior: A protection motivation theory perspective. *Information Systems Management, 33*(1), 2-16.

Hayden, L. (2010). *IT Security Metrics: A Practical Framework for Measuring Security & Protecting Data.* New York, NY: McGraw Hill.

Hersh, W. R., Totten, A. M., Eden, K., Devine, B., Gorman, P., Kassakian, S. Z., Woods, S. S., Daeges, M., Pappas, M., & McDonagh, M. A. (2016). The evidence base for health information exchange. In B.E. Dixon (Ed.), *Health Information Exchange: Navigating and Managing a Network of Health Information Systems* (pp. 213-229). Boston, MA: Academic Press.

Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgement of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(4), 621-642.

HIPAA Journal (2019). Summary of 2018 HIPAA fines and settlements. Retrieved from https://www.hipaajournal.com/summary-2018-hipaa-fines-and-settlements/.

Hubbard, D. W., & Seirsen, R. (2016). *How to measure anything in cybersecurity risk*. Hoboken, NJ: John Wiley & Sons.

Hussain, S., Bahadur, F., Gul, F., Iqbal, A., Ashraf, G., & Nazeer, S. (2015). Survey of Windows and Linux as server operating system. *International Journal of Computer, 18*(1), 1-6.

Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security, 31*(1), 83-95.

Ifinedo, P. (2014). Information systems security policy compliance: An empirical study of the effects of socialization, influence, and cognition. *Information & Management, 51*(1), 69-79.

Inshanally, P. (2018). *CompTIA Linux+ Certification Guide.* Birmingham, UK: Packt.

IONOS. (2019). Linux vs. Windows: A comparison of the best web server solutions. Retrieved from https://www.ionos.com/digitalguide/server/know-how/linux-vs-windows-the-big-server-check/.

Jaquith, A. (2007). *Security metrics: Replacing fear, uncertainty, and doubt.* Boston, MA: Pearson Education.

Johnston, A. C., Warkentin, M., & Siponen, M. (2015). An enhanced fear appeal rhetorical framework: Leveraging threats to the human asset through sanctioning rhetoric. *MIS Quarterly, 39*(1), 113-134.

Kahneman, D. (2003). A perspective on judgement and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697-720.

Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus, and Giroux.

Kliger, D., & Kudryavtsev, A. (2010). The availability heuristic and investors' reaction to company-specific events. *Journal of Behavioral Finance, 11*(1), 50-65.

Koch, D. D. (2017). Is the HIPAA security rule enough to protect electronic personal health information (PHI) in the cyber age? *Journal of Health Care Finance, 43*(3), 1-32.

Kraemer, S., & Carayon, P. (2007). Human errors and violations in computer and information security: The viewpoint of network administrators and security specialists. *Applied Ergonomics, 38*(2), 143-154.

Lowry, P. B., & Moody, G. D. (2015). Proposing the control-reactance compliance model (CRCM) to explain opposing motivations to comply with organizational information security policies: An empirical study of the influence of counterfactual reasoning and organizational trust. *Information Systems Journal, 25*(5), 193-230.

Lowry, P. B., Posey, C., Bennett, R. J., & Roberts, T. L. (2015). Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: an empirical study of the influence of counterfactual reasoning and organisational trust. *Information Systems Journal, 25*(3), 193-230.

Marsh, B., Todd, P. M., & Gigerenzer, G. (2004). Cognitive heuristics: Reasoning the fast and frugal way. In J. P. Leighton & R. J. Sternberg (Eds.), *The Nature of Reasoning* (pp 273-287). Cambridge, UK: Cambridge University Press.

McAllister, D. J. (1995). Affect and cognitive-based trust as foundations for interpersonal cooperation in organizations. *The Academy of Management, 38*(1), 24-59.

McFarland, M. (2012). Privacy and the law. *Markkula Center for Applied Ethics at St. Clara University*. Retrieved from https://www.scu.edu/ethics/focus-areas/internet-ethics/resources/privacy-and-the-law/

McKnight, D. H., Choudhury, V., & Kacmar, C. (2000). *Trust in e-commerce vendors: A two-stage model*. Proceedings of the 21st International Conference on Information Systems, 532-536.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review, 23*(3), 473-490.

Moqbel, M. A., & Bartelt, V. L. (2015). Consumer acceptance of personal cloud: Integrating trust and risk with the technology acceptance model. *AIS Transactions of Replication Research, 1*, 1-11.

Newman, D. (2019). Top EHR vendors 2019 – Epic, Cerner, Meditech, Allscripts. *Healthcare IT Skills.* Retrieved from https://healthcareitskills.com/top-ehr-vendors-allscripts-athenahealth-cerner-epic-meditech/.

NIST. (2018). *Framework for improving critical infrastructure cybersecurity.* Retrieved April 18, 2019, from: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf.

Ng, B., Kankanhalli, A., & Xu, Y. (2009). Studying users' computer security behavior: A health belief perspective. *Decision Support Systems, 46*(4), 815-825.

Nurse, J. R. C., Buckley, O., Legg, P. A., Goldsmith, M., Creese, S., Wright, G. R. T, & Whitty, M. (2014). *Understanding insider threat: A framework for characterizing attacks*. Proceedings of 2014 IEEE Security and Privacy Workshops, 214-228.

Office of the National Coordinator for Health Information Technology. (n.d.). Federal Health IT Strategic Plan: 2015-2020. Retrieved from https://dashboard.healthit.gov/strategic-plan/federal-health-it-strategic-plan-2015-2020.php

Oltsik, J. (2017). The life and times of cybersecurity professionals. *2017 ESG & ISSA Research Report.* Retrieved from https://www.esg-global.com/esg-issa-research-report-2017.

Oparaocha, G.O. (2016). Towards building internal social network architecture that drives innovation: A social exchange theory perspective. *Journal of Knowledge Management, 20*(3), 534-556.

Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied, 18*(3), 314-330. doi: 10.1037/a0028279.

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2013). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology, 40*(2), 544-554.

Ponemon Institute. (2018). 2018 Cost of data breach study global overview. Retrieved from https://www.ibm.com/security/data-breach.

Posey, C., Roberts, T., & Lowry, P. B. (2015). The impact of organizational commitment on insiders' motivation to protect organizational information assets. *Journal of Management Information Systems, 32*(4), 179–214.

Razmerita, L., Kirchner, K., & Nielsen, P. (2016). What factors influence knowledge sharing in organizations? A social dilemma perspective on social media communication. *Journal of Knowledge Management, 20*(6), 1225-1246.

Redmiles, E., Acar, Y., Fahl, S., & Mazurek, M. (2017). A summary of survey methodology best practices for security and privacy researchers. Retrieved from https://drum.lib.umd.edu/handle/1903/19227.

Renaud, K. (2012). Blaming noncompliance is too convenient: What really causes information breaches? *IEEE Security & Privacy, 10*(3), 57-63.

Rhee, H., Ryu, Y. U., & Kim, C. (2012). Unrealistic optimism on information security management. *Computers & Security, 31*(2), 221-232.

Roberts, M. J. (2004). Heuristics and reasoning I: Making deduction simple. In J. P. Leighton & R. J. Sternberg (Eds.), *The Nature of Reasoning* (pp 234-272). Cambridge, UK: Cambridge University Press.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *Journal of Psychology, 91*(1), 93-114.

Rogers, R. W., & Prentice-Dunn, S. (1997). Protection motivation theory. In D.S. Gochman (Ed.), *Handbook of Health Behavior Research I: Personal and Social Determinants* (pp. 113-132). New York, NY: Plenum Press.

Runyon, B. (2017). *The Road to Healthcare Interoperability Goes Through the National Health Information Network*. Retrieved from https://www.gartner.com/document/3755564?ref=solrAll&refval=190178249&qid=0c8e0704c872f156c385c614d27a5787.

Rutten, W., Blaas-Franken, J., & Martin, H. (2016). The impact of (low) trust on knowledge sharing. *Journal of Knowledge Management, 20*(2), 199-214.

Safa, N. S., Sookhak, M., Von Solms, R., Furnell, S., Ghani, N.A., & Herawan, T. (2015). Information security conscious care behavior formation in organizations. *Computers & Security, 53*, 65-78.

Safa, N. S., & Von Solms, R. (2016). An information security knowledge sharing model in organizations. *Computers in Human Behavior, 57*(C), 442-451.

Safa, N. S., Von Solms, R., & Furnell, S. (2016). Information security policy compliance model in organizations. *Computers & Security, 56*, 70-82.

Santara, M. (2013). Eliminating the security weaknesses of linux and UNIX operating systems. In J.R. Vacca (Ed.), *Computer and Information Security Handbook* (2nd ed, pp. 183-196). Boston, MA: Morgan Kaufmann.

Sedighi, M., van Splunter, S., Brazier, F., van Beers, C., & Lukosch, S. (2016). Exploration of multi-layered knowledge sharing participation: The roles of perceived benefits and costs. *Journal of Knowledge Management, 20*(6), 1247-1267.

Sekaran, U., & Bougie, R. (2013). *Research Methods for Business.* West Sussex, UK: John Wiley & Sons.

Shrivastava, A. (2018). Top 10 EMR software companies in 2018. Retrieved from https://arkenea.com/blog/top-10-emr-software-companies-in-2018/.

Siponen, M., Mahmood, M. A., & Pahnila, S. (2014). Employees' adherence to information security policy: An exploratory field study. *Information & Management, 51*(2), 217-224.

Siponen, M., Pahnila, S., & Mahmood, M. A. (2010). Compliance with information security policies: An empirical investigation. *IEEE Computer, 43*(2), 64-71.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research, 177*, 1333-1352.

Steinbrook, R. (2009). Health care and the American Recovery and Reinvestment Act. *New England Journal of Medicine, 360*(11), 1057-1060.

Sternberg, R. J. (2004). What do we know about the nature of reasoning? In J. P. Leighton & R. J. Sternberg (Eds.), *The Nature of Reasoning* (pp. 433-455). New York, NY: Cambridge University Press.

Tenable. (2017). By host, with suggested remediations. Retrieved April 21, 2019, from: https://static.tenable.com/documentation/reports/html/Vulns_By_Host.html#idm730272.

Thieme, E. (2016). Privacy, security, and confidentiality: Toward trust. In B.E. Dixon (Ed.), *Health Information Exchange: Navigating and Managing a Network of Health Information Systems* (pp. 91-104). Boston, MA: Academic Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor on heuristics-and-biases tasks. *Memory & Cognition 39*(7), 1275-1289.

Tsohou, A., Karyda, M., & Kokolakis, S. (2015). Analyzing the role of cognitive and cultural biases in the internalization of information security policies: Recommendations for information security awareness programs. *Computers & Security, 52*, 128-141.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

U.S. Department of Health and Human Services. Health Insurance Portability and Accountability Act of 1996 (HIPAA): Security standard. Retrieved from https://aspe.hhs.gov/report/health-insurance-portability-and-accountability-act-1996.

U.S. Department of Health and Human Services. (2013). Summary of the HIPAA security rule | HHS.gov. Retrieved from https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html.

U.S. Department of Health and Human Services Office for Civil Rights. (2019). Breach portal: Notice to the secretary of HHS breach of unsecured protected health information. Retrieved from https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf.

Van Vuuren, I. E. (2016). IT security trust model – securing the human perimeter. *International Journal of Social Science and Humanity, 6*(11), 852-858.

W3Techs. (2019, May 4). Usage of operating systems for websites. Retrieved from https://w3techs.com/technologies/overview/operating_system/all.

Wash, R., & Cooper, M. M. (2018). *Who provides phishing training? Facts, stories, and people like me*. Proceedings of the CHI Conference, 2018, Montreal, CA.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology, 103*(3), 506-519.

Whitman, M. E., & Mattord, H. J. (2012). *Principles of Information Security* (4th ed.). Boston, MA: Course Technology.

Wilson, M., & Hash, J. (2003). *Building an information technology security awareness and training program*, (NIST Special Publication 800-50). Retrieved April 18, 2019, from: https://csrc.nist.gov/publications/detail/sp/800-50/final.

Woon, I., Tan, G., & Low, R. (2005). *A protection motivation theory approach to home wireless security*. Proceedings of the 2005 Association for Information Systems International Conference on Information Systems, Las Vegas, NV.

Yoo, C. W., Sanders, G. L., & Cerveny, R. P. (2018). Exploring the influence of flow and psychological ownership on security education, training, and awareness effectiveness and security compliance. *Decision Support Systems, 108*, 107-118.

DECISION SCIENCES INSTITUTE

Impact of Firms' Relationship with past and existing Suppliers on Future Supplier Selection Decisions: A Focus Group Study

Attique ur Rehman
PhD Scholar
Suleman Dawood School of Business
Lahore University of Management Sciences, Pakistan
Email: attique@uvas.edu.pk

Muhammad Abubakar Siddique
PhD Scholar
Suleman Dawood School of Business
Lahore University of Management Sciences, Pakistan
Email: 17080005@lums.edu.pk

ABSTRACT

Supplier selection is a critical process that directly affects overall firms' performance. There are various factors that come into play when firms are establishing supplier selection criteria. One of the key factors that directly influence firms' decision towards establishing supplier selection criteria is their relationships with current and previous suppliers. In this study, we qualitatively explore the impact of few decision biases that generates from firms' relationship with existing or previous suppliers on firms' supplier selection criteria for new suppliers. Focus group interviews were used as a tool to collect qualitative data. The data was analyzed and results revealed that the selection criteria for new supplier is affected by reference point bias in decision making based on failed relationship between firm and a supplier.

KEYWORDS: Decision making, supplier selection, firms' relationships, selection criteria

INTRODUCTION

Firms today are faced with immense competition that demands improved quality, delivery, performance and responsiveness from organizations and focusing on reduced cost at the same time. Firms are finding different ways to leverage their supply chains and particularly, evaluating role of suppliers in their activities. Prahalad and Hamel (1993) identified the increasing role outsourcing of non-core activities of the firms. This helps organizations to focus more on their core activities to add value. Role of supplier is integral towards performance and attractiveness of entire supply chain. Supplier selection is important based on the fact that "it commits resources while simultaneously impacting such activities as inventory management, production planning and control, cash flow requirements, and product quality" (Narasimhan, 1983). There exist a range of supplier selection criteria, supplier evaluation and selection techniques including Data Envelopment Analysis (DEA), mathematical programming (including linear programming, integer linear programming, integer non-linear programming, goal programming, multi-objective programming), Analytic Hierarchy Process (AHP), case base reasoning, Analytic Network Process (ANP), Fuzzy set theory, Simple multi-attribute rating technique, genetic algorithm (Ho, Xu, & Dey, 2010). Supplier selection decision not only depends on cost and quality concerns but also various risk factors. Our study will mainly focus on the supplier selection criteria and sub-criteria. Ojala and Hallikas (2006) analyzed supplier investment risks, whereas Olson and Wu (2006) focused on broader categories of supply chain risks (internal and external) and their

controllability. Various studies have modeled uncertainty and risks in selection of supply chain partners. (Cohen & Lee, 1988; Graves & Willems, 2000; Lee & Billington, 1993; Thomas & Griffin, 1996) but there is lack of literature that explain biases in future supplier selection criteria based on existing and previous supplier relationship. In this study, we explore the impact of firms' current and previous supplier relationship on future supplier selection criteria. The key research question is "What are the decision-making biases that come into play while selecting future suppliers based on firms' relationship with existing and previous suppliers?"
We use focus group to explore the degree to which decision biases come into play while establishing supplier selection criteria based on previous and current supplier relationships. Following is literature review conducted to engrain the phenomena of supplier selection and decision biases through previous studies. Literature also helps to come up with the scope of discussion in focus group study. Next is the methodology section that gives the rationale and procedure of using focus group in this study followed by the data analysis and discussion sections. Conclusion section gives an overall summary of the whole study, processes involved and ultimate findings.

LITERATURE REVIEW

A critical decision in any of firms strategy is to decide what to buy and what to make (Platts, Probert, & Canez, 2002). When producing a specific product within a firm is beneficial and when it is feasible to outsource it. These are the decisions that force any firm to go out of its boundary and interact with other players in the market i.e. suppliers. Now on what basis firm should make this decision is answered by Transaction Cost Economic theory. This theory states that economic cost of transaction will decide if a particular transaction is worth taking or not (Williamson, 1981). This cost will also identify the structure of governance as well that whether it should be market, hierarchy or alliance (Williamson, 1975). According to McIvor (2009) there are four basic hurdles and difficulties in conducting transactions. These hurdles are bounded rationality, opportunism, small numbers bargaining, and information impactedness. He referred to bounded rationality as "the cognitive limitations of the human mind, which increases the difficulties of understanding fully the complexities of all possible decisions." Opportunism was defined as "decision makers acting with guile, as well as out of self-interest." Small numbers bargaining was defined as "the degree to which the buyer has alternative sources of supply to meet its requirements." And lastly, information impactedness was defined as "the presence of information asymmetries between the buyer and supplier, which means that either party may have more knowledge than the other." Further moving on he explained that these costs and transactional difficulties are further enhanced in the presence of asset specificity, uncertainty and infrequence. So, procurement becomes a critical decision for any organization when deciding on buy vs make problem.
Procurement is one of the critical activities and widely research topic in supply chain literature. Procurement function, in today's business environment has a pivotal role to play in firm's strategy formulation (Moses, 2011). Managing supply chain partners effectively directly affects firm's performance (Collins, Worthington, Reyes, & Romero, 2010). Supplier selection is one of critically researched domain of procurement literature. Too many criteria for supplier evaluation have been used in different context in the literature (Kumar Kar & K. Pani, 2014) but the basis on which these criteria are filtered and used is still vague. Some classic work of Busch (1962) and Dickson (1996) focused on supplier selection criteria like "quality", "delivery schedule" and "warranties". Further studies also highlighted the importance of "price", "technical" and "production" capability, "financial position", "vendor reputation", "post-sale services" and "past business records". After 1990s there was a shift from quantitative criteria to qualitative criteria for supplier evaluation. (Kumar Kar & K. Pani, 2014). As the onset of globalization come into

play, further development in criteria occurred including “geographic locations”, “exchange rates”, “tariffs and customs”, trade restrictions”, “quality management” and other environmental factors. (Braglia & Petroni, 2000; Choi & Hartley, 1996; Cusumano & Takeishi, 1991; Narasimhan, Talluri, & Mendez, 2001; Talluri & Narasimhan, 2005).

Research has shown that supplier’s knowledge can affect buyer’s ability to learn and bring change (Jarratt, 2004), however this mainly depends upon the degree of positive relationship that exist between buyer and supplier. Integrating with supplier depends mainly on the success of buyer-supplier relationship (Vanpoucke, Vereecke, & Wetzels, 2014) which is largely based on supplier’s selection and evaluation process.

programming (linear programming, integer linear programming, integer non-linear programming, goal programming, multi-objective programming), Data Envelopment Analysis (DEA), Analytic Hierarchy Process (AHP), case base reasoning, Analytic Network Process (ANP), Fuzzy set theory, Simple multi-attribute rating technique, genetic algorithm (Ho et al., 2010). Kannan, Haq, Sasikumar, and Arunachalam (2008) used interpretive structural modeling and AHP for analyzing green suppliers. In another article, Hsu and Hu (2009) evaluated suppliers incorporating hazardous substance management through ANP. Büyüközkan and Görener (2015) used Fuzzy AHP for sustainable supplier selection with incomplete information. In a study by Shaw, Shankar, Yadav, and Thakur (2012), fuzzy AHP and multi-objective linear programming was used to select suppliers for developing low carbon supply chain. Sarkis and Dhavale (2015) focused on Bayesian framework and Monte Carlo simulations for supplier selection for sustainable operations.

In spite the fact that use of supplier selection and evaluation criteria is highly case specific, the buyers’ decision is not always objective. AHP gained importance because it has the ability to address human subjectivity in decision making and it can bring out non deterministic requirements (Saaty, 1994). AHP helps to decompose the problem into a hierarchy of sub-problems that are more easily comprehended and can be analyzed independently. AHP can also measure the degree of consistency in priorities which making a supplier selection. Our study focuses on the fact that this change in priorities and supplier selection criteria is highly influenced by a buyer’s relationship with existing or past suppliers. Decision biases come into play when setting priorities in selection criteria for a particular supplier and it directly affects degree of risk an organization is willing to take being a buyer. This is where decision theory comes into play.

As explained by Transaction Cost Economic theory, bounded rationality of human beings increases the difficulties in transaction. Hence role of decision making and rationality needs to be explored in the area of supply chain. According to H. A. Simon (1955) human beings are rational decision makers and they always act in their self-interest to get the best outcome of their choice and preferences. This was the basic assumption of “homo economicus” and it was also prevailed in the field of supply chain management in the form of institutional economic theory and transaction cost economics (Grover & Malhotra, 2003; McNally & Griffin, 2004; Rindfleisch & Heide, 1997). Even though basic assumption of human rationality has ruled the Supply Chain Management, but a new field of research has emerged within Supply Chain Management which has challenged this basic assumption and is known as Behavioral Supply Chain Management. It is defined by Carter, Kaufmann, and Michel (2007) as “The Study of how judgement in supply management decision-making deviates from the assumptions of homo economicus”. In his article, he reviewed a wide-ranging literature which is based on judgement and decision-making biases and integrated it by making taxonomies which can then be used in the field of supply management by supply managers and researchers. Literature has provided ample proof that human beings frequently violate the assumption of rationality and make biased decisions leading to results which are not fully optimal (Fischhoff, Slovic, & Lichtenstein, 1978;

Kahneman & Tversky, 1972; H. Simon, 1957; Thaler, 1985). Hence role of biases in Supply Chain decision-making cannot be ignored.

There are many areas where role of biases can be investigated in supply chain management. But supply chain management begins from selecting suppliers. So, supplier selection criteria and decision-making process is one the central task in supply management. It has been mainstream topic of research in the domain of supply chain management in which suppliers are selected on the principles of rationality drawn from transaction cost economics theory (Choi & Hartley, 1996; Grover & Malhotra, 2003; Huang & Keskar, 2007; McNally & Griffin, 2004; Weber, Current, & Benton, 1991). Even though majority of research are based on mathematical optimization of selection of supplier and its decision-making, but involvement of human beings makes it prone to use of heuristics and biases in decision making (Carter et al., 2007; Chan, 2003; Cook, 1992; Das & Teng, 1999; Karpak, Kumcu, & Kasuganti, 1999; Kaufmann, Michel, & Carter, 2009). It results in suboptimal solutions because one may discard other alternatives available and instead just accept satisficing solution which H. Simon (1957) explained as to choose "good enough" instead of optimal solution (Carter et al., 2007).

There are many biases which can be relevant in decision-making process of supplier selection. Carter et al. (2007) studied around 76 biases and using cluster analysis he came up with 9 categories of biases which are relevant in the context of supply chain management. Out of these 9 categories, Reference Point Bias is one of bias which can potentially affect the decision-making process of supplier selection. For simplicity and focus, in this paper we will stick to this bias only and will explore its effect on supplier selection criteria. According to Carter et al. (2007) "The reference point bias occurs when evaluations and adjustments from an initial position or reference point are usually insufficient". So, it can be argued that when selecting specific supplier and manager or company may create a reference point which can be its previous suppliers and it starts evaluating and adjusting the new supplier according to its experience with the previous supplier and this evaluation and adjustments can be insufficient according to the reference point bias.

According to Tversky and Kahneman (1974) the basic heuristic used for simplification of complex phenomenon by human beings is to start with an initial point and then alter and assess their opinions based on it. So suppliers may base their initial point as previous suppliers and make adjustments based on that experience but researchers have indicated that this alteration is usually insufficient to make optimal decision (Slovic, Fischhoff, & Lichtenstein, 1977). Similarly, when it comes to negotiating with a new supplier then organization can make the initial offer based on the previous supplier so it will lead to biased negotiation between the two parties because previous studies has shown in many experiments that final agreement in negotiations are biased towards the initial offer (Galinsky & Mussweiler, 2001). Moreover, it has also been reported that even when initial point is selected arbitrary still human beings fall prey to this bias leading to inefficient decision making so in supplier selection case even if there is no previous supplier experience but still organization may start comparing it to randomly initial points (Epley & Gilovich, 2005). Carter et al. (2007) explained this phenomenon nicely with an example related to buyer supplier relationship. According to him "For instance, a buyer might demand only incremental improvements in price levels from a supplier, because the current price level "anchors" the buyer judgment. In reality, however, the supplier's price may be far too high." This shows that Reference Point bias have the potential to effect supplier selection criteria because reference point can only be considered as rational if initial point is best indication of future selection which they are not (Hogarth, 1987).

THEORETICAL DEVELOPMENT/MODEL

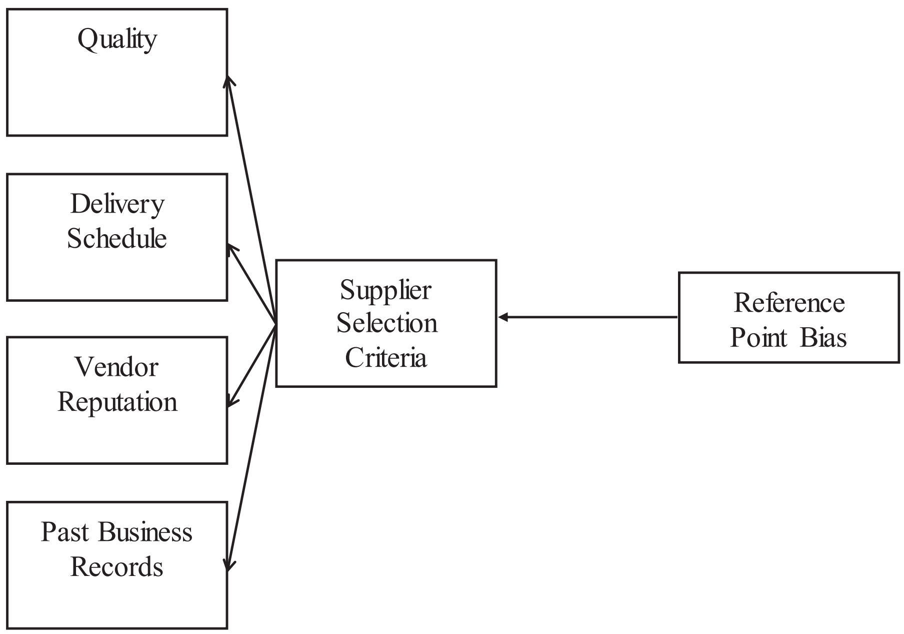


Fig 1

METHODOLOGY

Literature review and focus group discussion was adopted as a method to conduct this study. Focus group is among the top used techniques used in social sciences as a tool to execute research (Alfalla-Luque, Medina-Lopez, & Schrage, 2013). Many studies in the area of Supply Chain Management has used focus group as a research tool for example (Oehmen, De Nardo, Schönsleben, & Boutellier, 2010; Tate, Ellram, Bals, Hartmann, & Van der Valk, 2010) and it has also been used in general Operations Management field as well (Binder & Edwards, 2010; Dey, Hariharan, & Clegg, 2006). Focus group is about a well-organized discussion among a group of 6-12 members on a specific topic (Kitzinger, 1999; Tate et al., 2010). Focus group was useful for our study because we were doing exploratory study to find what biases come into play in supplier selection criteria. Focus group discussions characterize specific themes that are explored in depth (Bryman, 2015). Gibbs (1997) highlighted the importance of use of focus group as a tool that may reveal respondents' attitudes, beliefs, experiences and reactions that might not be possible to cater through other means of data collection like surveys.

Experts from different industries were selected for focus group. These experts are currently or have already worked for more than 6 years, in procurement department of their companies which was responsible for supplier selection. Brief profile of the experts is given below in the table:

No.	Name	Company	Designation	Education	Experience
1	Muhammad Bilal	Coca Cola	Procurement Specialist	MBA	7 years
2	Anayat Ali	Coca Cola	Procurement Manager	MBA	8 years
3	Haris Ali	HUBCO	Procurement and Planning Manager	Mechanical Engineer	10 years
4	Sheikh Latif	Packages Ltd.	Planning and Production Manager	MBA	15 years
5	Syed Muhammad Ahmed	Ali Murtaza and Co.	Procurement Manager	Master's in economics	8 years

Initially, a telephonic interview was conducted with these experts and objectives of the study were conveyed to them. Questions were explained that would be discussed in focus group discussion and their consent was taken if they are ready to participate in the study. It was also inquired that if they believe that they are eligible to answer the required questions.

Tool

An interview guide was developed including few leading questions that cover the scope of selection criteria and different biases that may come into play while applying those criteria. Questions along with interview guide are attached in Appendix 1.
In the focus group, the experts were first about the general importance of supplier in overall supply chain. Everyone was of the view that relationship with supplier is a strategic key to success of whole supply chain. After the initial brainstorming on the role of supplier, focus group was moved towards the supplier selection criteria and role of bias in decision making. Review of the objectives was also repeated in focus group discussion. The focus group was managed and moderated by one of the researchers. Participants were also requested that they should give an overall view about supplier instead of specific point of view from the perspective of their individual companies to make it more general. A qualitative analysis was conducted on the discussion held during focus group and results from preliminary literature review were also incorporated.

Procedure

The focus group discussion was planned using a semi-structured interview guide and some key topics were drawn in the guide. Before discussing the decision biases, the respondents were asked about the key criteria they use for selection of suppliers in their respective organizations. Respondents came up with different responses. Thus in the arguments section, both before and during discussion responses of the respondents are discussed. Two researchers were present during focus group discussion; one leading the discussion and other took notes. The interview lasted for around ninety minutes. Interview was recorded and transcribed for the purpose of analysis. Nvivo 11 was used to analyze the themes present in the transcribed interview.

Data Analysis

Transcribed material were analyzed through thematic analysis (Boyatzis, 1998). Initially six themes were identified for coding 1-6. Further transcribed material was scrutinized and all statements were reviewed to be associated with one or more themes. This allowed extracting subthemes. Table I shows the themes for code 1-6:

	Name
	Supplier selection criteria and effective procurement
	supplier screening process
	Bias in selecting new supplier
	Effect of previous supplier on new supplier selection
	Relationship sensitivity
	Common factors of failed supplier relationship

Table I

RESULT

The thematic analysis of transcribed material expanded into six themes with various sub-themes shown in table II. Two themes related to supplier selection criteria and supplier screening process whereas four themes concerned biases in decision making process, effect of previous supplier relationship on new supplier selection, supplier sensitivity and reasons behind failed supplier relationships.

Sr. No.	Themes	Sub-themes
1.	Supplier selection Criteria and effective procurement	1a. Competitiveness and cost 1b. Depends upon market, category of product and technical specifications
2.	Supplier screening process	2a. Research and Development 2b. Service level 2c. Suppliers' experience
3.	Bias in selecting new supplier	Bad relationship affects decision
4.	Effect of previous supplier on new supplier selection	4a. Intuition 4b. More subjective 4c. Increased resistance 4d. Reduced leverages to new suppliers 4e. Reference points
5.	Relationship sensitivity	5a. Industry and nature of supplies 5b. Integrity 5c. Compliance 5d. Sharing knowledge
6.	Common factors of Failed supplier relationships	6a. Delayed payments 6b. Delayed deliveries

Table II

Supplier Selection and Effective Procurement

The initial responses covered both positive and negative aspects of supplier selection criteria. Supplier selection was discussed to be the most critical aspect of procurement process and it was also revealed that firms might bear high cost for overlooking this important dimension. Negative aspects revealed that it might be take ample time in trying to risk averse in this regard. So it is a trade-off between the risk a firm might be willing to take and cost they are interested in spending in selecting appropriate supplier. Vendors play a strategic role; world is transforming into a global village so organizations cannot rely on one supplier because they need competitiveness and cost cutting.
There is enormous literature available on supplier selection. Weber et al. (1991) carried out an extensive literature on supplier selection criteria and noted that quality, net price and delivery ranked at the top by 13 articles. Production facilities and capabilities were also present in the list of important criteria. 26 criteria were summarized into 8 factors including finances, consistency, relationship, flexibility, technological capability, reliability, price and customer service (Choi & Hartley, 1996).
The respondents in focus group also conform to the supplier selection criteria as highlighted in the literature discussed. Supplier selection and effective procurement largely depend upon the market, technical specifications of the products and a set of pre-qualifications including quality, suppliers' financial performance, logistics (delivery schedule) and price
Role of supplier selection in effective procurement also varies from industry to industry. There are some industries where there is monopoly of suppliers and organizations are left with little or no leverage on the suppliers. Almost all respondents agreed to some key supplier selection criteria as quality, price, financials, sustainability, and above all technical specifications. Technical qualification is one of the most important criteria a supplier has to meet in supplier selection process of an organization. One of the respondents commented:
"It works like filtering layers. The first filter is technical qualification. For example, there are some technical qualifications a supplier has to meet in my organization like few ISO standards, OHSAS etc. After these technical qualifications have been met, then we consider the second layer that is much flexible and might vary from supplier to supplier. For example, a supplier might be charging higher prices but giving value added services so we would be willing to pay high price for value added services."

Supplier Screening Process

The second theme that was coded as a part of thematic analysis is suppliers' screening process. Two important considerations are there to discuss in supplier screening process. At the first level market intelligence system plays a pivotal role in supplier screening process in an organization. Industries in Pakistan are highly fragmented and it is very easy to gather intelligence about repute of various suppliers in the market. Secondly, firms' own knowledge management systems are important in screening the key suppliers from a range of available suppliers in the market. Procurement function also generates information that helps future decisions.
Internet has also played a significant role in finding out suppliers that qualify firms' criteria. The discussion also revealed that background check on the supplier is also an important element in the screening process. Organizations need to check litigations on suppliers before extending business terms with them. All the respondents agreed to the fact that their respective organizations have an appropriate supplier screening process in practice.

Bias in Selecting Supplier

The most critical stage of focus group interview starts with this theme. Initially, the discussion was mainly centered at identifying the key criteria for supplier selection and supplier screening process. After icebreaking and preliminary discussion, the participants in the discussion were introduced to the phenomena of decision biases. Moderator presented a situation to the respondents in which their relationship with the supplier is affected due to some reason and now they are considering to switch supplier. So whether or not, the selection criteria or screening process will remain the same or it will be altered owing to some cognitive bias. There was a mix of responses from the participants in the focus group. Some participants commented that there exists a positive relationship between relationships with supplier and bias in future supplier selection and some participants responded that it purely depends on the circumstances and the degree of relationship sensitivity between buyer (organization) and supplier. This discussion allowed a new theme to emerge (relationship Sensitivity) discussed further.
The key argument raised in discussing decision biases is that the organization (procurement function) will become more sensitive in selecting new supplier. Definitely changes will be incorporated in the process of supplier selection because every organization wants to avoid the same mistake happening twice.

Effect of Previous Supplier Relationship on new Supplier Selection

A more critical theme “Effect of previous supplier relationship on new supplier selection” emerged that highlighted different sub-themes in the coding process. All participants in the focus group discussion agreed to the fact that previous suppliers’ relationship definitely affects the selection of new suppliers.
“You might be more considerate about something you overlooked. It is a continuous development and learning process. If you have to select a new supplier, the very first challenge is to break you comfort zone. If you willing for this, you have to reform your criteria and SOPs for supplier selection.”
Another comment was that human element has to be reduced while refining the supplier selection criteria, which is not completely possible. There will be a human element and an unintentional decision bias might exist while selecting new supplier. In fact Chen and Chao (2012) identified a highly structured mechanism to evaluate suppliers based on multi-criteria decision making (MCDM) but the baseline criteria on which this model works is highly subjective and prone to human error. Another sub-theme that emerged from this theme is reduced leverage to the new supplier. Respondents believed that there are different leverages a firm may extend to their suppliers in terms of flexible lead time or timely payment schedules. When switching suppliers, the firm might reduce these leverages by focusing on strict lead times or demanding extended payment schedules. Trust is one of the reasons that will play a pivotal role in differentiating firms’ relationship with previous and new supplier.
In order to explore the phenomena of decision biases, the moderator also asked about some of the bad experiences, the participants had with their suppliers that prove to be an anchor or reference point when making a decision of selecting new supplier. This phenomena is well grounded in literature as Behavioral Supply Management (BSM) and has raised the level of attention given to the deviations from standard assumptions of rational paradigms in decision making (Kaufmann, Carter, & Buhrmann, 2012). In many situations, reference-dependent valuation affects decision makers’ behaviors that diverge from rationally optimized decision otherwise.

Relationship Sensitivity

As discussed above, bias in supplier selection also emerged a new theme that is sensitivity in buyer-supplier relationship. The phenomenon of sensitivity was discussed by majority of the participants in focus group that may affect supplier selection decision. Degree of sensitivity in relationship depends on nature of supplies. For routine supplies, the relationship is quite sensitive as there a number of alternatives available. For strategic suppliers, the degree of sensitivity is quite low. Sensitivity also comes into play when buyer and supplier share the same process. As in case of outsourcing where an outsourced partner completes a part of a firm's process. In this case the relationship will be entirely different as the firm and its supplier are sharing common objectives. Sensitivity also depends on risk involved. Dupont, Bernard, Hamdi, and Masmoudi (2018) studied the impact of delivery failure risk on supplier selection decision. Higher the risk involve, greater would be the sensitivity of the relationship between buyer and supplier.

Common Factors of Failed Buyer-Supplier Relationship

The last theme that emerged from coding was the key factors that contribute towards failure of buyer supplier relationship. There are two key factors from buyer (firm) and supplier's end: Delayed payments and delayed deliveries respectively. Lack of trust is also an important reason that contributes towards a failed buyer-supplier relationship. Kim and Choi (2015) theorized two orthogonal aspects in buyer-supplier relationships: (1) relational posture (means to what extent the firms regard each other as cooperative partners) and (2) relational intensity (means to what extent the operations of buyer and supplier are linked together. Based on these dimensions, a typology was developed that defined four types of buyer-supplier relationships: Deep, Sticky, Transient and Gracious. A sticky relationship is based on lack of trust and information sharing and most probably results in failure.

DISCUSSION

The in-depth examination of impact of decision biases on supplier selection unfolded different dimensions explained above in various themes and sub-themes. Supplier selection impacts effective procurement and becomes a reason for firms' effectiveness. There exists a significant positive relationship between procurement process and firm performance (Othman, Abd Rahman, Sundram, & Bhatti, 2015). Effective supplier selection is imperious for cost saving and thus offering better quality to customer (Tracey & Leng Tan, 2001). Screening suppliers is also an important aspect of procurement process. Conforming to the outcomes of our focus group discussion on supplier screening process, Rezaei, Fahim, and Tavasszy (2014) proposed a two layer model for supplier selection airline industry. The first phase is conjunctive screening phase in which the initial set of potential suppliers is reduced to a smaller set prior to a comprehensive analysis for final selection. Then, in the second phase, a quantitative model is used to evaluate suppliers against different criteria and sub-criteria.
Bias in decision making process comes into play at two stages: while establishing the selection criteria and while making final selection of the supplier. The interviews revealed strong concerns about interplay pf bias and supplier selection criteria. Bias

largely depends upon the degree of relationship sensitivity between buyer and supplier. Sensitivity can be affected by various factors like risk, price, lead time and quality. When establishing supplier selection criteria, it is evident from focus group interviews that reference points deeply affect supplier selection process and may deviate decision maker from highly rational and optimizing approach towards more intuitive and subjective decision making. Failed relationships become a trigger towards relationship sensitivity in future suppler relationship and can also be a source of decision bias especially when there is a need to make immediate switching decision. The failed relationship will become a reference point and new supplier will be screened and evaluated on the basis of that reference point. We use insights from transaction cost economics theory that explains that decisions are driven through bounded rationality and information impactedness. Our results conforms with the McIvor (2009) four elements in transaction cost economics theory.
Our qualitative analysis showed that there is indeed an impact of previous buyer supplier relationship on establishing supplier selection criteria for a new supplier and decision can be biased based on degree of sensitivity of the relationship with the previous supplier.

CONCLUSION

The qualitative study explored the key supplier selection criteria that may be affected by any decision-making bias based on existing and past supplier relationship and relationship sensitivity. The results through focus group discussion revealed that there is indeed an impact of decision bias on supplier selection criteria and it largely depends on intuition and subjectivity of the decision maker. The participants in focus group shared their own experiences in their respective organizations that further solidify our propositions. Thematic analysis was conducted in Nvivo 11 and based on the transcribed interviews, six themes emerged with various sub-themes.
One major limitation of this study is that the results are reported from the outcome of discussion with one focus group. There are two key process involved in supplier selection: supplier selection criteria and supplier evaluation. Supplier evaluation is majorly based on quantitative modeling techniques. This study focused primarily on supplier selection criteria.

BIBLIOGRAPHY

Alfalla-Luque, R., Medina-Lopez, C., & Schrage, H. (2013). A study of supply chain integration in the aeronautics sector. *Production Planning & Control, 24*(8-9), 769-784.

Binder, M., & Edwards, J. S. (2010). Using grounded theory method for theory building in operations management research: a study on inter-firm relationship governance. *International journal of operations & production management, 30*(3), 232-259.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*: sage.

Braglia, M., & Petroni, A. (2000). A quality assurance-oriented methodology for handling trade-offs in supplier selection. *International Journal of Physical Distribution & Logistics Management, 30*(2), 96-112.

Bryman, A. (2015). *Social research methods*: Oxford university press.

Busch, G. (1962). New twist on supplier evaluation. *Journal of Purchasing, 55*(1), 102-103.

Büyüközkan, G., & Görener, A. (2015). Evaluation of product development partners using an integrated AHP-VIKOR model. *Kybernetes, 44*(2), 220-237.

Carter, C. R., Kaufmann, L., & Michel, A. (2007). Behavioral supply management: a taxonomy of judgment and decision-making biases. *International Journal of Physical Distribution & Logistics Management, 37*(8), 631-669.

Chan, F. S. (2003). Interactive selection model for supplier selection process: an analytical hierarchy process approach. *International Journal of Production Research, 41*(15), 3549-3579.

Chen, Y.-H., & Chao, R.-J. (2012). Supplier selection using consistent fuzzy preference relations. *Expert systems with applications, 39*(3), 3233-3240.

Choi, T. Y., & Hartley, J. L. (1996). An exploration of supplier selection practices across the supply chain. *Journal of operations management, 14*(4), 333-343.

Cohen, M. A., & Lee, H. L. (1988). Strategic analysis of integrated production-distribution systems: Models and methods. *Operations research, 36*(2), 216-228.

Collins, J. D., Worthington, W. J., Reyes, P. M., & Romero, M. (2010). Knowledge management, supply chain technologies, and firm performance. *Management Research Review, 33*(10), 947-960.

Cook, R. L. (1992). Expert systems in purchasing: applications and development. *Journal of Supply Chain Management, 28*(4), 20-27.

Cusumano, M. A., & Takeishi, A. (1991). Supplier relations and management: a survey of Japanese, Japanese-transplant, and US auto plants. *Strategic management journal, 12*(8), 563-588.

Das, T., & Teng, B. S. (1999). Cognitive biases and strategic decision processes: An integrative perspective. *Journal of Management Studies, 36*(6), 757-778.

Dey, P. K., Hariharan, S., & Clegg, B. T. (2006). Measuring the operational performance of intensive care units using the analytic hierarchy process approach. *International Journal of Operations & Production Management, 26*(8), 849-865.

Dickson, G. W. (1996). An analysis of vendor selection systems and decisions.

Dupont, L., Bernard, C., Hamdi, F., & Masmoudi, F. (2018). Supplier selection under risk of delivery failure: a decision-support model considering managers' risk sensitivity. *International Journal of Production Research, 56*(3), 1054-1069.

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making, 18*(3), 199-212.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance, 4*(2), 330.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: the role of perspective-taking and negotiator focus. *Journal of personality and social psychology, 81*(4), 657.

Gibbs, A. (1997). Focus groups. *Social research update, 19*(8), 1-8.

Graves, S. C., & Willems, S. P. (2000). Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management, 2*(1), 68-83.

Grover, V., & Malhotra, M. K. (2003). Transaction cost framework in operations and supply chain management research: theory and measurement. *Journal of operations management, 21*(4), 457-473.

Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of operational research, 202*(1), 16-24.

Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision*.

Hsu, C.-W., & Hu, A. H. (2009). Applying hazardous substance management to supplier selection using analytic network process. *Journal of Cleaner Production, 17*(2), 255-264.

Huang, S. H., & Keskar, H. (2007). Comprehensive and configurable metrics for supplier selection. *International journal of production economics, 105*(2), 510-523.

Jarratt, D. (2004). Conceptualizing a relationship management capability. *Marketing Theory, 4*(4), 287-309.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology, 3*(3), 430-454.

Kannan, G., Haq, A. N., Sasikumar, P., & Arunachalam, S. (2008). Analysis and selection of green suppliers using interpretative structural modelling and analytic hierarchy process. *International Journal of Management and Decision Making, 9*(2), 163-182.

Karpak, B., Kumcu, E., & Kasuganti, R. (1999). An application of visual interactive goal programming: a case in vendor selection decisions. *Journal of Multicriteria Decision Analysis, 8*(2), 93.

Kaufmann, L., Carter, C. R., & Buhrmann, C. (2012). The impact of individual debiasing efforts on financial decision effectiveness in the supplier selection process. *International Journal of Physical Distribution & Logistics Management, 42*(5), 411-433.

Kaufmann, L., Michel, A., & Carter, C. R. (2009). Debiasing strategies in supply management decision-making. *Journal of Business Logistics, 30*(1), 85-106.

Kim, Y., & Choi, T. Y. (2015). Deep, sticky, transient, and gracious: An expanded buyer–supplier relationship typology. *Journal of Supply Chain Management, 51*(3), 61-86.

Kitzinger, J. (1999). Introduction: the challenge and promise of focus groups' in R. Barbour and J. Kitzinger (eds) Developing Focus Group Research: Politics, Theory and Practice: London: Sage.

Kumar Kar, A., & K. Pani, A. (2014). Exploring the importance of different supplier selection criteria. *Management Research Review, 37*(1), 89-105.

Lee, H. L., & Billington, C. (1993). Material management in decentralized supply chains. *Operations research, 41*(5), 835-847.

McIvor, R. (2009). How the transaction cost and resource-based theories of the firm inform outsourcing evaluation. *Journal of operations management, 27*(1), 45-63.

McNally, R. C., & Griffin, A. (2004). Firm and individual choice drivers in make-or-buy decisions: a diminishing role for transaction cost economics? *Journal of Supply Chain Management, 40*(4), 4-17.

Moses, A. (2011). Cross-functional make or buy decision process ownership. *Management Research Review, 34*(9), 1042-1060.

Narasimhan, R. (1983). An analytical approach to supplier selection. *Journal of Supply Chain Management, 19*(4), 27-32.

Narasimhan, R., Talluri, S., & Mendez, D. (2001). Supplier evaluation and rationalization via data envelopment analysis: an empirical examination. *Journal of Supply Chain Management, 37*(2), 28-37.

Oehmen, J., De Nardo, M., Schönsleben, P., & Boutellier, R. (2010). Supplier code of conduct–state-of-the-art and customisation in the electronics industry. *Production Planning & Control, 21*(7), 664-679.

Ojala, M., & Hallikas, J. (2006). Investment decision-making in supplier networks: Management of risk. *International journal of production economics, 104*(1), 201-213.

Olson, D. L., & Wu, D. (2006). Simulation of fuzzy multiattribute models for grey relationships. *European Journal of operational research, 175*(1), 111-120.

Othman, A. A., Abd Rahman, S., Sundram, V. P. K., & Bhatti, M. A. (2015). Modelling marketing resources, procurement process coordination and firm performance in the Malaysian building construction industry. *Engineering, Construction and Architectural Management, 22*(6), 644-668.

Platts, K., Probert, D., & Canez, L. (2002). Make vs. buy decisions: A process incorporating multi-attribute decision-making. *International journal of production economics, 77*(3), 247-257.

Prahalad, C., & Hamel, G. (1993). The core competence of the corporation, Harvard Business School Pub. *Corp., Boston, MA*.

Rezaei, J., Fahim, P. B., & Tavasszy, L. (2014). Supplier selection in the airline retail industry using a funnel methodology: Conjunctive screening method and fuzzy AHP. *Expert systems with applications, 41*(18), 8165-8179.

Rindfleisch, A., & Heide, J. B. (1997). Transaction cost analysis: Past, present, and future applications. *The journal of marketing*, 30-54.

Saaty, T. L. (1994). How to make a decision: the analytic hierarchy process. *Interfaces, 24*(6), 19-43.

Sarkis, J., & Dhavale, D. G. (2015). Supplier selection for sustainable operations: A triple-bottom-line approach using a Bayesian framework. *International journal of production economics, 166*, 177-191.

Shaw, K., Shankar, R., Yadav, S. S., & Thakur, L. S. (2012). Supplier selection using fuzzy AHP and fuzzy multi-objective linear programming for developing low carbon supply chain. *Expert systems with applications, 39*(9), 8182-8192.

Simon, H. (1957). Models of man Wiley. *New York*.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics, 69*(1), 99-118.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual review of psychology, 28*(1), 1-39.

Talluri, S., & Narasimhan, R. (2005). A note on" A methodology for supply base optimization". *IEEE Transactions on Engineering Management, 52*(1), 130-139.

Tate, W. L., Ellram, L. M., Bals, L., Hartmann, E., & Van der Valk, W. (2010). An agency theory perspective on the purchase of marketing services. *Industrial Marketing Management, 39*(5), 806-819.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science, 4*(3), 199-214.

Thomas, D. J., & Griffin, P. M. (1996). Coordinated supply chain management. *European Journal of operational research, 94*(1), 1-15.

Tracey, M., & Leng Tan, C. (2001). Empirical analysis of supplier selection and involvement, customer satisfaction, and firm performance. *Supply Chain Management: An International Journal, 6*(4), 174-188.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science, 185*(4157), 1124-1131.

Vanpoucke, E., Vereecke, A., & Wetzels, M. (2014). Developing supplier integration capabilities for sustainable competitive advantage: A dynamic capabilities approach. *Journal of operations management, 32*(7-8), 446-461.

Weber, C. A., Current, J. R., & Benton, W. (1991). Vendor selection criteria and methods. *European Journal of operational research, 50*(1), 2-18.

Williamson, O. E. (1975). Markets and hierarchies–Free Press. *New York*.

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American journal of sociology, 87*(3), 548-577.

Appendix I:

Interview Protocol Form (Focus Group)

Interview Protocol

Institutions/Company Name:
____

Interviewee (Title and Name): _____________________________________

Post Interview Comments or Leads:

____

"Supplier Selection and Decision Biases" Interview

Introductory Protocol

To facilitate our note-taking, we would like to audio tape our conversations today. Please sign the protocol form. For your information, only researchers on the project will be privy to the tapes. which will eventually be transcribed for the study. Essentially, (1) all information will be held confidential, (2) your participation is voluntary and you may stop at any time if you want to. Thank you for your willingness to participate.

We have planned this interview to last no longer an hour. During this time, we have few questions that we would like to cover.

Signature (consent to do the interview)

____

Introduction

We are students of Lahore University of Management Sciences (LUMS) and currently studying a course on Decision Behavior as a mandatory requirement of our PhD program. You have been selected to speak with us today because you have been identified as someone who has a great deal of experience in procurement decisions in an organization. Our research focuses on understanding the criteria based on which a firm choses it supply chain partners (vendors). What role is played by different biases in decisions making process and what are the challenges you face while establishing criteria for procurement decisions.

Vendor Selection Criteria:

Upfront Questions:

1. What do you think is the role of selection criteria in selecting appropriate vendors for the organization?
2. Prioritize some of the key supplier selection criteria while making a decision.
   Probes: [What is your rational behind this prioritizing? Has this criteria sustained or has it changed over time?]
3. Do you think you selection criteria are influenced by any bias?
   Probe: [What bias you might think has come into play while making vendor selection decision?]
4. How frequently your organization changes vendors?
   Probe: [To what extent vendor selection criteria changes as you switch from one supplier to another].
5. How sensitive is an organization's relationship with particular vendors?
   Probe: Do you think this sensitivity of the relationship affects your future vendor selection?
6. Suppose you have 1-2 minute to conclude our talk about role of supplier relationship that might bias your decision for future supplier selection, how would you conclude that?

Back-end questions:

1. What are supplier selection criteria in general?
2. What biases might come into play based on buyer-supplier relationship?
3. How these biases affect future supplier selection?
4. What are prioritizing principles of supplier selection criteria in absence of any decision bias?

5. What are prioritizing principles of supplier selection criteria in presence of any decision bias?

DECISION SCIENCES INSTITUTE

Digitalizing Supply Chains: Potential Business Benefits and Challenges

Abubaker Haddud
Southwest Minnesota State University
Email: abu.haddud@smsu.edu

Anshuman Khare
Athabasca University
Email: anshuman.khare@fb.athabascau.ca

ABSTRACT

Digital transformation continues to change business landscapes and it is important to understand aspects related to such transformation. We present a research examining potential benefits and challenges associated with digitalizing supply chains. We analyze survey data collected from 74 respondents examining predetermined 10 potential benefits and 10 potential challenges. We ranked these benefits and challenges to highlight their degree of importance. This paper can greatly benefit parties interested in understanding aspects related to supply chain digitalization and how to effectively manage digital transformation within the entire value chain. Such understanding helps managers identify key business areas digitalizing supply chains directly impacts.

KEYWORDS: Digital supply chain, Digital transformation, Benefits, Challenges

INTRODUCTION

Businesses use more and more of smart technologies within different business functions, and supply chain management is one of the key areas that is directly impacted by such adoption. Gartner provides definitions to ‘digitization’ and ‘digitalization. Digitization is defined as “the process of changing from analog to digital form, also known as digital enablement” (Gartner IT Glossary - 1). Whereas, digitalization is defined as **“**the use of digital technologies to change a business model and provide new revenue and value-producing opportunities; it is the process of moving to a digital business (Gartner IT Glossary - 2). In this study, we focus on the ‘digitalization’ processes exploring aspects related to more use of digital technologies within supply chains. More adoption of smart technologies leads to more level of digitalization of different supply chain functions. Recent developments concerning digitalization are expected to play an increasingly significant role in the management and design of global supply chains (Klötzer and Pflaum, 2017). Therefore, organisations should not consider “if” they set-up a digital strategy and get their organisation and environment ready for the digital transformation rather take the required actions as soon as possible and bring the digital transformation to life (Bienhaus and Haddud, 2018). Thus, it is very important that businesses understand key aspects related to digital transformation and where within the business it is occurring most. Supply chain management is a key business function that has seen a great deal of technology utilizations. Companies continue to seek ways to achieve competitive advantages through the adoption of more smart technologies within their internal supply chains as well as with their extended external chains that involve strategic business partners. Therefore, it is important to study the impact digitalizing supply chain would have on the operations and performance of the entire supply chain and to identify key potential business benefits and key potential challenges associated with such trends. This paper seeks to answer the following two questions:

1. What are the key potential business benefits that are likely to result from digitalizing supply chains?
2. What are the key potential challenges that businesses are likely to face when digitalizing supply chains?

Answers to these questions will provide a useful reference that decision makers can use to help them decide what emerging technologies to use and where to implement them in areas related to main elements of value chains. The answers also provide a list of potential challenges that managers need to be aware of when embarking on digitalizing supply chain initiatives that help with the implementation and operations of new acquired technologies. This paper is structured as follows: the next section provides a literature review about supply chain management, digital transformation, and digital supply chain. After that, the research methodology and method are presented. Next, results and discussions of the collected data are presented covering the two main themes of this paper as 'potential benefits' and 'potential challenges. Finally, conclusions, recommendations, and future research suggestions are included in the last section of the paper.

LITERATURE REVIEW

Supply Chain Management

Supply chain management deals with the movement, and storage, of raw materials, work-in-progress, and finished goods from the point of origin to the point of customer buying and receiving the finished goods. Supply chains are divided into two main parts: internal supply chains that consist of activities and units within a firm, and external supply chain that includes activities based in external business parties such as suppliers, distributors, retails, and customers. Integration and coordination between the internal and external parts of supply chains is crucial to the performance of supply chains as a whole value network.

Supply Chain Management is more than just managing the follow of materials. It is the oversight of material, information, and financial flows inside and outside an organization. Flows move from supplier to manufacturer to wholesaler to retailer to consumer (and back). It is concerned with planning, implementing and controlling the flows of raw materials, inventory and finished goods from the point of origin to the point of consumption, and covers purchasing, the manufacturing process and customer delivery. Lately, with the emphasis on "green," one would incorporate environmental considerations at each stage of supply chain management including post-consumer disposal. We think the information flow part is important when we discuss digitalization. One way of effectively integrating the different parts of value chains is through the use of smart technologies that ensure effective coordination and smooth operations. In this paper, we seek to identify key potential benefits and challenges related to digitalizing supply chains so that decisions are made to further improve the performance of entire supply chains (internal and external) which will lead to better overall organizational performance.

Digital Transformation

Digital transformation can be simply defined as the use of technology to radically improve performance or reach of companies (Westerman et al., 2014) and it is currently an important trend that many industries are embracing, and it creates some disruptions to the business norms. Digital transformation is a transformation that impacts business activities, processes, competencies, and models to fully leverage the changes and opportunities brought by digital technologies (Demirkan et al., 2016). Digital transformation is normally driven by internal factors such as: the continuous need to improve processes and entire workplace. Also, to effectively integrate information systems within the company and with external partners to enhance overall management of processes and to reduce costs. Externally, companies continue to invest in digital transformation to meet ever-changing customer demands and to meet supply chain management needs (Pflaum et al., 2019). Supply chain management is one of the key business areas that have seen a huge adoption of different technologies and it contributes greatly to the digital transformation of businesses as a whole. The next section discusses aspects of digital supply chains and the main theme of this paper is to identify key businesses benefits and challenges associated with digitalizing supply chains.

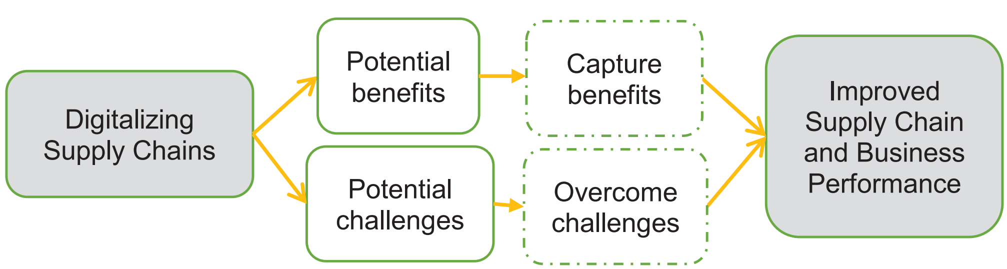
Digital Supply Chains

Digital Supply Chain (DSC) is defined as “an intelligent process that leverages innovative solutions with emerging technological means to generate new forms of revenue and business value for organizations” (Büyüközkan and Göçer, 2017). Digital supply chain relies on key drivers that consist of technologies, digitization, integration, collaboration and coordination (Iddris, 2018). The transformation process into a digital supply chain requires three key enablers: a clear definition, new capabilities, and a supportive environment (Alicke et al., 2017). The main goal of the digital supply chain is to deliver the right product into the customer's predefined spot as quickly as possible (Pereira et al., 2017). Therefore, digitalizing supply chain is no easy task and it requires effective strategies and robust implementation measures in place. In order to gain the necessary support for implementing digital supply chain initiatives, it is crucial that decision makers appreciate the potential benefits such move would bring to the business. This paper provides insights about 10 potential benefits businesses may gain from digitalizing supply chains and identifying the key impacted areas. Such information can greatly help interested decision makers with appreciating the impact more digitalizing in supply chains would have on their management and performance. Furthermore, this paper provides information about 10 key potential challenges that businesses are likely to face when acquiring the necessary capabilities and technologies to digitalize supply chains. Such information is crucial in helping decision makers understand the readiness status of key business areas and supply chain elements in order to gauge the possible occurrence of any of the highlighted potential challenges. Thus, preventive measures and other actions can be taken ahead of any digital supply chain initiatives to ensure a smooth and effective implementation and management of any new initiatives.

CONCEPTUAL RESEARCH FRAMEWORK

Figure 1 shows the conceptual framework for this research. It illustrates that digitalizing supply chains initiatives will lead to businesses gain potential benefits. Once these benefits are capitalized and captured, the overall supply chain and business performance improves. It also illustrates that digitalizing supply chains initiatives will likely face a number of potential challenges. Handling and overcoming such challenges will lead to better supply chain management and business performance. The two boxes with dashed lines are not the focus of this study. These areas are relevant to the study but need future research attention.

Figure 1 Conceptual research framework



METHODS

An online survey was utilized to collect the primary data for this research capturing the perceptions of respondents about aspects related to digitalizing supply chains. This research study was approved by Southwest Minnesota State University's Institution Review Board in September 2018 and the used online survey consisted of six main parts (see Figure 2). Three parts are included in this paper (parts 1, 3, and 4). First, General Information section (7 questions). Second, the respondents were asked to rate their perception about 10 potential benefits related to digitalizing supply chains. Third, the respondents were asked to rate their perception about 10 potential challenges related to digitalizing supply chains.

Figure 2 Survey development and structure

Main parts of used survey					
1	2	3	4	5	6
General Information (*7 questions*)	Influence of Technological Trends on SCM (*7 examined technologies*	Benefits from Digitizing Supply Chains (*10 examined benefits*)	Challenges Associated with Digitizing Supply Chains (*10 examined challenges*)	Impact of Digitizing Supply Chains on Five Lean Manufacturing Tools (*5 examined items under each tool = 25 items in total*)	Potential Impacts of Digitizing Supply Chains on The Seven Types of Manufacturing Wastes (*5 examined items under each waste = 35 items in total*)

The developed survey statements related to the examined potential benefits and challenges were extracted from information presented in a number of refereed academic journal articles as presented in Table 1.

Table 1: Sources used to develop the examined survey statements [benefits and challenges]

Examined benefits	**Source**	**Examined challenges**	**Source**
Improved visibility across chains and better traceability of processes and products	Musa et al., 2014	Challenges related to integrating different systems across the network	Majeed and Rupasinghe, 2017; Murphy, 2017
Better information-sharing and supply chain integration	de Mattos and Barbin Laurindo, 2015	Managing the transformation process from an existing company to its digitalized equivalent	Lee and Lee, 2015; Pflaum et al., 2017; Ryan and Watson, 2017

Quicker responses to changes in demand, effective inventory and logistics management, and reduction in supply chain complexity	Garrett, 2014; Gao et al., 2015	Lack of robust preparation and understanding	Howells, 2016
Supply chains become able to react more effectively to changes in the marketplace	Haddud and Khare, 2018.	Required financial investments from all supply chains players to design and deploy technologies and solutions necessary to digitize supply chains.	Haddud et al., 2017
Better demand forecasting accuracy and optimization of ordering process.	Bonnes, 2014; Ittmann, 2015	Employees' resistance to new technologies and practices	Haddud et al., 2017
Improved operational efficiencies, improved service quality, and greater resource efficiency.	Tseng et al., 2011; Ngai et al., 2011	Challenges with supply chain management integration requirements.	Haddud et al., 2017
Supply chains become more responsive to resolve supply chain-related operating problems, e.g., complex distribution problems	Haddud and Khare, 2018.	Access to/ availability of the required skilled human resources	Ryan and Watson, 2017
Supply chain designs become; flexible, open, agile, and collaborative	Petrick and Simpson, 2013	Risks and variabilities that may arise from the use of ecosystems	Lee and Lee, 2015; Riggins and Wamba, 2015
More effective delivery route planning and inventory planning	Haddud and Khare, 2018.	Diversity of adopted business models within the different supply chain entities and organizations.	Haddud et al., 2017
Enhanced employees' internal collaborations, communications, and engagement levels	Haddud et al., 2016	Supply chains in general not ready for such a change and lack the required architecture	Bughin et al., 2015; Jin et al., 2014; Li et al., 2015

Data were collected in November and December 2018 using an online survey created on Google Forms. Respondents were either authors or co-authors of papers in the Decision Science Institute (DSI) annual conference proceedings and the Emerald's Journal of Manufacturing Technology Management from 2015 to 2018. The email invitations were sent to 923 respondents and 79 completed surveys were received. This makes the response-rate as 8.55%. However, five of the received surveys were incomplete and were excluded from the data analysis process. Therefore, the total number of usable received survey responses was 74. The profile of respondents in this study is shown in Table 2. The respondents' years of work experiences was self-reported into four categories. Also, the respondents were categorized into three main areas of expertise and the remaining had different fields and were placed under one category as ‘other’. Furthermore, the respondents either worked for academia, industry, or both. The respondents’ levels of knowledge about ‘supply chain management’, ‘lean operations’, and ‘digital supply chains’ were also probed. Therefore, and based on the self-reported levels of knowledge the respondents had about these themes, it is fair to say that the respondents who completed this survey are knowledgeable about the examined themes and the perceptions they provided can be considered valid, reliable, and credible.

Table 2: Profile of Participants

	Question	Frequency (n=74)	Percentage (%)
Years of work experience in current field	• 1 - 5	13	17.56
	• 6 - 10	20	27.02
	• 11 - 20	19	25.67
	• 21 +	22	29.72
Field and main area of expertise	• Supply Chain and Operations Management	32	43.24
	• Lean Manufacturing	9	12.16
	• Manufacturing/production/logistics	19	25.67
	• Other	14	18.91
Sector	• Academia	52	70.27
	• Industry	10	13.51
	• Both	12	16.21
Level of knowledge – used scale (Low) 1 – 7 (High)		***Mean values***	
	• Knowledge of Supply Chain Management	5.66	
	• Knowledge of Lean Operations	5.46	
	• Knowledge of Digital Supply Chains	4.73	

The IBM Statistical Package for the Social Sciences (SPSS) version 24 was used to analyze the collected data in this study. First, statistical results will be presented in Table 4 about the 10 examined potential benefits. Second, the same statistical results are also presented in Table 5 for the 10 examined potential challenges. Item mean values, standard deviation, Cronbach’s alpha, and factor loadings were all calculated and presented for each examined survey statement and were included in the two tables.

RESULTS AND DISCUSSIONS

Reliability and Validity of Used Survey Constructs

Reliability is defined as the proportion of variance in observed test score that is related to true scores (Cronbach, 1951; McDonald, 1999; Lo and Yeung, 2006) of the scale items in the same construct (DeVellis, 1991; Hinkin, 1995). Therefore, Cronbach alpha values were used to examine the consistency and reliability of the items used in each of the two used survey constructs. As shown in Table 3, the Cronbach alpha coefficient values for the used potential benefits construct ranged from 0.895 to 0.903 (see Table 4 for all item values) and the potential challenges construct had values ranged from 0.655 to 0.715 (see Table 5 for all item values). Rivard and Huff (1988) suggest that Cronbach's values exceeding alpha coefficient of 0.7 thresholds provide reliability evidence for internal consistency of the measurement scales.

Although an acceptable reliable coefficient is normally 0.7 or higher, lower thresholds are sometimes used in the literature (Santos, 1999). For new instruments, constructs with reliability values as low as 0.5 are also acceptable (O'Leary-Kelly and Vokurka, 1998). The Cronbach alpha coefficient values for all the used items in the two used constructs were all above 0.655. Because the closer to 1 the Cronbach alpha coefficient value is, the more reliable the used construct becomes (Cohen et al., 2003), it can be concluded that the reliability of used survey constructs be considered as a reliable research tool. Therefore, we believe that each of the examined items under the two used survey constructs are adequate and provide a thorough picture of the examined benefits and challenges associated with supply chain digitalizing. Also, each of the used items, within each construct, is directly related to the main theme (either benefits or challenges) of the construct and all of the used statements are inter-related. Thus, we relate the high Cronbach alpha coefficient values of the two used constructs to these two main factors.

Table 3 Used survey constructs' validity and reliability tests results

Used Scale	Number of items	Items factor loadings ranges	Cronbach's α value ranges
• Potential benefits	10	0.486 to 0.633	0.895 to 0.903
• Potential challenges	10	0.577 to 0.795	0.655 to 0.715

Given the fact that the survey instrument used in this study was developed for this study and it had not been used before, an exploratory factor analysis (Principal Component) was used to examine its construct validity. This analysis helps to determine how, and to what extent, each item within the two used constructs is linked to their underlined factors (Yee-Loong Chong et al., 2009). Hair et al., (2006) state that the rule-of-thumb is that factor analysis values greater than 0.30 should be considered significant, values greater than 0.40 should be considered more important, and values that are 0.50 or greater should be considered very significant. As shown in Table 3, item's factor loading values ranged from 0.486 to 0.633 for the used 10 items under the potential benefits construct (see Table 4 for all item values). Also, item's factor loading values for the used 10 items in the potential challenges construct ranged from 0.577 to 0.795 (see Table 5 for all item values). With the exception of one item, under the used potential benefits construct, that had a factor loading value of 0.486, the remaining items within the same construct and all of the other 10 items within the potential challenges construct had factor loading values above 0.5. Therefore, the construct validity of the used two constructs in this study is very significant.

Results and Discussions for Potential Benefits

This research is looking at how digitalizing supply chains impacts the performance and management of supply chains. The research also seeks to identify key benefits businesses are likely to gain from such digitalization. 10 pre-determined potential benefits were examined in this study using a five-point Likert scale. The included 10 examined benefits statements in Table 4 were the used survey statements presented to the targeted respondents to respond to. As it is presented in Table 4, the highest rated potential benefit was 'improved visibility across chains and better traceability of process and products' with an overall mean value of 4.75. Enhancing supply chain visibility has been one of the capabilities that businesses continuously seek to improve, and it is very crucial to the performance of value networks (Busse et al., 2017; Maghsoudi and Pazirandeh; 2016). The adoption of different technologies within supply chains has a strong potential in improving the supply chain visibility and to provide businesses with traceability capabilities that further enhance the management of supply chains (Dubey et al., 2018; Raman et al., 2018; Srinivasan and Swink, 2018). Our research has confirmed the cruciality of digitalizing supply chain management on improving supply chain performance and management through more effective visibility capabilities.

The second top rated potential benefit was 'better information-sharing and supply chain integration' with a mean value of 4.49. It is important to mention that data and information are seen as crucial business assets and businesses continue to seek better ways to capture valuable data and, more important, analyze it to extract meaningful information that can help in making crucial business decisions. Without the use of smart and advanced technologies to capture and analyze data, such process would be very challenging. Thus, digital transformation and the adoption of different technologies, particularly within supply chains, helps capture and share data and information across the entire value chain that not only includes businesses, but it extended to included suppliers, distributers, retailers, and customers. In this study, we were able to confirm the importance of such benefit businesses would gain from more digitalization of supply chains. The third highest rated potential benefit was 'quicker responses to changes in demand' (mean=4.44) and the fourth top rated potential benefit was 'supply chains react more effectively to marketplace changes' (mean=4.42). The ability of businesses to effectively respond to ever-changing marketplace challenges is very crucial (Moyano-Fuentes et al., 2016). Businesses will need to create a good balance between responsiveness and efficiency (Hugos, 2018). Adoption of different smart technologies within supply chains is likely to improve companies' responsiveness capabilities (De Propris and Pegoraro, 2019; Kayikci, 2018). Our study confirmed the importance of such benefit to supply chain management. In the fifth place, respondents rated 'Better demand forecasting accuracy and optimization of ordering process' with mean value of 4.36. Big data analytics leads to better demand forecasts (Lee, 2017); leading retailers have been using point-of-sales data to optimize forecasting (Aryal et al., 2017). In this study, we confirmed the importance of digitalizing supply chains in helping companies plan more accurate demand forecasts and make better procurement decisions.

Table 4: Statistical results for examined benefits

Examined benefits	Mean	Std. Deviation	Cronbach's Alpha	Factor loadings	Rank	Impact on
Improved visibility across chains and better traceability of processes and products	4.75	0.619	0.902	0.486	1	Visibility
Better information-sharing and supply chain integration	4.49	0.71	0.898	0.535	2	Visibility
Quicker responses to changes in demand, effective inventory and logistics management, and reduction in supply chain complexity	4.44	0.687	0.896	0.607	3	Agility
Supply chains become able to react more effectively to changes in the marketplace	4.42	0.686	0.899	0.536	4	Agility
Better demand forecasting accuracy and optimization of ordering process.	4.36	0.674	0.895	0.633	5	Efficiency

Improved operational efficiencies, improved service quality, and greater resource efficiency.	4.34	0.731	0.9	0.52	6	Efficiency
Supply chains become more responsive to resolve supply chain-related operating problems, e.g., complex distribution problems	4.32	0.724	0.896	0.597	7	Agility
Supply chain designs become; flexible, open, agile, and collaborative	4.29	0.716	0.896	0.595	8	Agility
More effective delivery route planning and inventory planning	4.23	0.755	0.9	0.515	9	Efficiency
Enhanced employees' internal collaborations, communications, and engagement levels	3.85	0.938	0.903	0.517	10	Collaboration

The sixth, seventh, and eighth ranked potential benefits were 'Improved operational efficiencies, improved service quality, and greater resource efficiency' (mean=4.34); 'Supply chains become more responsive to resolve supply chain-related operating problems, e.g., complex distribution problems' (mean=4.32); and 'Supply chain designs become; flexible, open, agile, and collaborative' (mean=4.29). Technology trends such as Industrial Internet of Things and Industry 4.0 envision improved operational efficiency (Preuveneers et al., 2017); P&G and Wal-Mart had significantly improved operational performance through the use of Broadcast Driver Architecture and other IT tools (Pu et al., 2018). In this study, we further confirm the importance of digitalizing supply chains on improving supply chain's and organizational efficiencies. The ninth ranked potential benefit was 'More effective delivery route planning and inventory planning' (mean=4.23). The use of different smart technologies leads to better inventory management. For example, one of the key benefits from adopting blockchain technology is better inventory management without the need for double-verifications (Wang et al., 2019); use of RFID technologies has led to better inventory management (Nair and Anbuudayasankar, 2016). Therefore, there is a clear proportional relationship between the level of supply chain digitalization and inventory management improvement and this research has confirmed the importance of this benefit.

Finally, the tenth ranked potential benefit was 'Enhanced employees' internal collaborations, communications, and engagement levels' (mean=3.85). Digitalizing supply chains contributes to what is known as 'closed-loop supply collaboration' and the utilization of more digital technologies is likely to lead to a better collaboration performance. It is expected that more supply chain digitalization will strengthen the collaboration between members of the entire value networks. However, this potential benefit was ranked the least among the 10 examined potential benefits and perhaps further studies may investigate this point further.

Results and Discussions for Potential Challenges

The purpose of using this construct was to identify key potential challenges that businesses are likely to face when adopting more digitalization within their supply chains. 10 predetermined challenges were included in this construct and respondents were asked to report their level of agreement with the examined statements. The included 10 examined challenges statements in Table 5 were the used survey statements presented to the targeted respondents to respond to As Table 5 shows, the highest rated potential challenge was 'challenges related to integrating different systems across the networks with an overall mean value of 4.43. This remains as a key challenge related to technology diffusion that businesses continue to face when implementing new technological solutions into their business environments (Valmohammadi, 2016; Bröring et al., 2017; Haddud et al., 2017). In this study, we confirmed the importance of such a challenge and thus businesses need to be aware of this potential issue so that necessary actions are taken to overcome any possible integration challenges. 'Managing the transformation process from an existing company to its digitalized equivalent' potential challenge was rated as the top second challenge (mean=4.29). Transforming a traditional supply chain into a digital ecosystem is a complicated process that requires a lot of planning and execution strategies. In this study, this challenge seemed very crucial to the process of digitalizing supply chains and this goes in line with other studies indicating a similar challenge such as (Lavikka et al., 2017; Schallmo and Williams; 2018; Wiseman, 2019). Furthermore, 'Lack of robust preparation and understanding' can be seen as part of challenges associated to managing digital transformation and this challenge was ranked in the third place with a mean value of 4.22. Some businesses find themselves forced to react to a certain technological trend that necessitates acquiring certain technologies in order to catch up with other market leaders in adopting similar technologies. In some cases, some businesses may lack the required preparation and understanding of how to acquire and implement these new technologies. This is likely to cause difficulties in integrating the new technologies and may result in serious challenges. In this study, respondents perceived such a challenge as one of the top three challenges associated with digitalizing supply chains that businesses need to be aware of. The fourth ranked potential challenge was related to required financial investment from all supply chain players (mean=4.21). In order for supply chains to digitize in an effective manner, all the involved parties should possess certain technical capabilities that allow effective technology implementation and operations. Some parties, particularly smaller suppliers, may lack the financial and human skills required to invest in new emerging technologies (Martinsuo and Luomaranta, 2016) and such technical improvements can lead to huge financial burdens (Kumar and Shoghli, 2018).

Table 5: Statistical results for examined challenges

Examined challenges	Mean	Std. Deviation	Cronbach's Alpha	Factor loadings	Rank	Challenge core source
Challenges related to integrating different systems across the network	4.43	0.601	0.696	0.645	1	Technology diffusion
Managing the transformation process from an existing company to its digitalized equivalent	4.29	0.615	0.704	0.613	2	Technology diffusion
Lack of robust preparation and understanding	4.22	0.755	0.715	0.577	3	Human factor
Required financial investments from all supply chains players to design and deploy technologies and solutions necessary to digitize supply chains.	4.21	0.838	0.706	0.749	4	Technology diffusion
Employees' resistance to new technologies and practices	4.06	0.82	0.692	0.795	5	Human factor
Challenges with supply chain management integration requirements.	4.04	0.701	0.658	0.679	6	Technology infusion

Access to/ availability of the required skilled human resources	3.94	0.837	0.682	0.739	7	Human factor
Risks and variabilities that may arise from the use of ecosystems	3.85	0.744	0.664	0.649	8	Technology infusion
Diversity of adopted business models within the different supply chain entities and organizations.	3.76	0.864	0.703	0.651	9	Business models
Supply chains in general not ready for such a change and lack the required architecture	3.69	1.002	0.655	0.647	10	Business models

In the fifth place came the potential challenge of ‘Employees’ resistance to new technologies and practices’ (mean=4.06). By default, employees are resistant to change particularly when they are comfortable with what they do, and because any change would require changing their routine. Employees’ resistance to change has been identified as one of the key challenges related to supply chain initiatives (Wiseman, 2019). This study has confirmed the importance of such a potential challenge to change so that businesses take this into account when planning more digitalization of supply chains. The sixth ranked potential challenge was ‘challenges with supply chain management integration requirements’ (mean=4.04). Supply chain integration refers to integrating regional supply chains into an integrated global supply network (Agrawal and Narain, 2018). Digitalization of information flows and system integration makes it much more challenging in terms of data security and protection of intellectual property (Wei et al., 2019). Therefore, in this study, we confirmed this challenge and believe that supply chain digitalization processes will likely face a similar challenge and businesses should consider this carefully before embarking on any new digitalization initiatives.

The remaining four potential challenges had mean values below 4.0. These results imply that these examined potential challenges are not crucial anymore. For example, ‘Access to/ availability of the required skilled human resources’ had a mean value of 3.94 and perhaps the reason why this potential challenge was ranked relatively low was because the marketplace has seen entrance of more skilled human resources due to universities and other educational institutions had introduced programs that equip graduates with the necessary skills required for todays’ technology sectors. Furthermore, in the eighth place, ‘Risks and variabilities that may arise from the use of ecosystems’ had a mean value of 3.85 and we think this is because more businesses are using ecosystems and it is becoming a norm rather than a ‘feared hype’. In the ninth place, ‘diversity of adopted business models within the different supply chain entities and organization’ had a mean value of 3.76. Most businesses follow common business structures that fit the nature and size of their business and these do not differ greatly from the business models of other business partners. Thus, any supply chain digitalizing initiatives implemented across the same value chains are not likely to face a challenge related to business models compatibility.

Finally, the least perceived crucial potential challenge was ‘supply chains in general not ready for such a change and lack the required architecture’. This result was no surprise because many businesses have implemented some degree of technological solutions within their supply chains. Embracing more emerging technologies usually builds on existing capabilities and infrastructures and this was not perceived to be a major challenge to more digitalization of supply chains.

CONCLUSIONS

Concluding Remarks

The two main themes of this paper were; 1) identifying key potential benefits from digitalizing supply chains; and 2) identifying key potential challenges associated with digitalizing supply chains. The results revealed that businesses may gain the following main benefits: 1) Supply chains become more visible and this leads to better processes and products traceability. 2) digitalizing supply chains leads to better information-sharing and supply chain integration. 3) supply chains become able to react more effectively to changes in the marketplace that allows quicker responses to changed demand through better demand forecast. 4) digitalizing supply chains improves overall operational efficiencies. 5) more effective inventory management and supply chains become more flexible. As indicated in Table 4, these benefits have direct impacts on supply chain's; agility, efficiency, visibility, and collaboration.

On the other hand, digitalizing supply chains is likely to face several potential challenges that may include: 1) challenges related to integrating different systems across the network. 2) difficulties with managing the transformation process and lack of robust preparation and understanding. 3) investing in new technologies may lead to financial burden for small businesses. 4) employees' resistance to new technologies and practices. 5) supply chain management integration requirements. As indicated in Table 5, these challenges can be categorized into four main groups; technology diffusion, technology infusion, business models, and human factor.

Recommendations

It is recommended that businesses continue to explore the benefits that emerging technologies would bring to their current supply chain structures and operations and decide what technologies to acquire. By using the identified 10 potential business benefits in this study, businesses can identify the business areas that require improvements and then decide what technologies to adopt. On the other hand, it is very important that decision makers be aware of the associated potential challenges that are likely to face when embarking on supply chain digitalization initiatives so that necessary measures are taken in order to avoid the occurrence of any of these challenges and to ensure a smooth implementation and management of new technologies.

Future Research Suggestions

It is recommended that similar studies are conducted to explore the impact of digitalization on other business functions to further identify where more technologies can be adopted within businesses and to further identify any possible challenges associated with the process. It is also recommended that similar research studies are conducted as case studies exploring similar aspects on companies that have greatly invested in technology adoptions and that are market leaders in this field. It is also important to learn aspects related to implementing digitalization initiatives and to identify key best practices. Furthermore, our study explored potential benefits and challenges related to digitalization of supply chains and we recommend that other studies are conducted to explore the development of digitalization frameworks that help businesses with the implementation and management of new technologies. More importantly, to explore how to actualize and maximize the benefits and to minimize the challenges. It would be interesting to survey a larger sample size and perhaps use more open-ended questions where respondents can report other potential benefits and challenges in addition to the predetermined ones in this

study. We believe that it would also be interesting to replicate the study across different industries and sectors to gauge if there are any key similarities or differences with how businesses approach digitalization initiatives.
Finally, digitization allows for more accurate tracking of parts and processes. While here we look at the supply chain moving in one direction, a bigger topic of “de-manufacturing” can be researched which looks at the complete life cycle of a product. This will contribute to closer analysis of green supply chains.

REFERENCES

Agrawal, P. and Narain, R., 2018, December. Digital supply chain management: An Overview. In *IOP Conference Series: Materials Science and Engineering* (Vol. 455, No. 1, p. 012074). IOP Publishing.

Alicke, K., Rexhausen, D. and Seyfert, A., 2017. Supply chain 4.0 in consumer goods. *Operations as a competitive advantage in a disruptive environment. McKinsey & Company*, pp.41-51.

Aryal, A., Liao, Y., Nattuthurai, P. and Li, B., 2018. The emerging big data analytics and IoT in supply chain management: a systematic review. *Supply Chain Management: An International Journal (early cite)*.

Bienhaus, F. and Haddud, A., 2018. Procurement 4.0: factors influencing the digitisation of procurement and supply chains. *Business Process Management Journal*, *24*(4), pp.965-984.

Bonnes, K., 2014. Predictive Analytics for Supply Chains: a Systematic Literature Review. Available from: http://referaat.cs.utwente.nl/conference/21/paper/7436/predictive-analytics-for-supply-chains-a-systematic-literature-review.pdf. Accessed on 8 August 2017.

Bröring, A., Schmid, S., Schindhelm, C.K., Khelil, A., Käbisch, S., Kramer, D., Le Phuoc, D., Mitic, J., Anicic, D. and Teniente, E., 2017. Enabling IoT ecosystems through platform interoperability. *IEEE software*, *34*(1), pp.54-61.

Bughin, J., Chui, M. and Manyika, J., 2015. An executive’s guide to the Internet of Things. *McKinsey Quarterly, McKinsey & Company*, *2*(9), pp.89-105.

Busse, C., Schleper, M.C., Weilenmann, J. and Wagner, S.M., 2017. Extending the supply chain visibility boundary: utilizing stakeholders for identifying supply chain sustainability risks. *International Journal of Physical Distribution & Logistics Management*, *47*(1), pp.18-40.

Büyüközkan, G. and Göçer, F., 2017, June. An extension of MOORA approach for group decision making based on interval valued intuitionistic fuzzy numbers in digital supply chain. In *2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS)* (pp. 1-6). IEEE.

Cohen, J., Cohen, P., West, S.G. and Aiken, L.S., 2003. *Applied multiple regression/correlation analysis for the behavioral sciences*.3rd ed., Lawrence Erlbaum Associates, London.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), pp.297-334.

de Mattos, C.A. and Barbin Laurindo, F.J., 2015. Collaborative platforms for supply chain integration: Trajectory, assimilation of platforms and results. *Journal of technology* management *& innovation*, *10*(2), pp.79-92.

Demirkan, H., Spohrer, J.C. and Welser, J.J., 2016. Digital innovation and strategic transformation. *IT Professional*, *18*(6), pp.14-18.

De Propris, L. and Pegoraro, D., 2019. Technological Disruptions and Production Location Choices. In *The Changing Strategies of International Business* (pp. 221-240). Palgrave Macmillan, Cham.

DeVellis, R.F., 2016. *Scale development: Theory and applications* (Vol. 26). Sage Publications, Newbury Park, CA.

Dubey, R., Luo, Z., Gunasekaran, A., Akter, S., Hazen, B.T. and Douglas, M.A., 2018. Big data and predictive analytics in humanitarian supply chains: enabling visibility and coordination in the presence of swift trust. *The International Journal of Logistics Management*, *29*(2), pp.485-512.

Garrett, B., 2014. 3D printing: new economic paradigms and strategic shifts. *Global Policy*, *5*(1), pp.70-75.

Gartner IT Glossary - 1, 2019. Available from: https://www.gartner.com/it-glossary/digitalization/. Accessed on 19 March 2019.

Gartner IT Glossary - 2, 2019. Available from: https://www.gartner.com/it-glossary/?s=digitization. Accessed on 19 March 2019.

Gao, W., Zhang, Y., Ramanujan, D., Ramani, K., Chen, Y., Williams, C.B., Wang, C.C., Shin, Y.C., Zhang, S. and Zavattieri, P.D., 2015. The status, challenges, and future of additive manufacturing in engineering. *Computer-Aided Design*, *69*, pp.65-89.

Haddud, A., DeSouza, A., Khare, A. and Lee, H., 2017. Examining potential benefits and challenges associated with the Internet of Things integration in supply chains. *Journal of Manufacturing Technology Management*, *28*(8), pp.1055-1085.

Haddud, A., Dugger, J.C. and Gill, P., 2016. Exploring the Impact of Internal Social Media Usage on Employee Engagement. *Journal Social Media for Organizations*, *3*(1), pp.1-17.

Haddud, A. and Khare, A., 2018. The Impact of Digitizing Supply Chains on Lean Operations. In *Marktorientiertes Produkt-und Produktionsmanagement in digitalen Umwelten* (pp. 27-46). Springer Gabler, Wiesbaden.

Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. 2006. *Multivariate Data Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hinkin, T.R., 1995. A review of scale development practices in the study of organizations. *Journal of management*, *21*(5), pp.967-988.

Howells, R. 2016. The Supply Chain Goes Digital: Opportunities And Challenges. Available from: http://www.digitalistmag.com/digital-supply-networks/2016/05/09/the-supply-chain-goes-digital-opportunities-challenges-04194124. Accessed on 10 August 2017.

Hugos, M.H., 2018. *Essentials of supply chain management*. John Wiley & Sons.

Iddris, F., 2018. Digital Supply Chain: Survey of the Literature. *International Journal of Business Research and Management*, *9*(1), pp.47-61.

Ittmann, H.W., 2015. The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management*, *9*(1), pp.1-9.

Jin, J., Gubbi, J., Marusic, S. and Palaniswami, M., 2014. An information framework for creating a smart city through internet of things. *IEEE Internet of Things Journal*, *1*(2), pp.112-121.

Kayikci, Y., 2018. Sustainability impact of digitization in logistics. *Procedia manufacturing*, *21*, pp.782-789.

Klötzer, C. and Pflaum, A., 2017. Toward the development of a maturity model for digitalization within the manufacturing industry's supply chain. In *50th IEEE Hawaii International Conference on System Sciences (HICSS),* (pp. 4210-4219).

Kumar, A. and Shoghli, O., 2018. A review of IoT applications in Supply Chain Optimization of Construction Materials. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 35, pp. 1-8). IAARC Publications.

Lavikka, R., Hirvensalo, A., Smeds, R. and Jaatinen, M., 2017, September. Transforming a supply chain towards a digital business ecosystem. In *IFIP International Conference on Advances in Production Management Systems* (pp. 295-301). Springer, Cham.

Lee, I., 2017. Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, *60*(3), pp.293-303.

Lee, I. and Lee, K., 2015. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, *58*(4), pp.431-440.

Li, S., Da Xu, L. and Zhao, S., 2015. The internet of things: a survey. *Information Systems Frontiers*, *17*(2), pp.243-259.

Lo, V.H. and Yeung, A., 2006. Managing quality effectively in supply chain: a preliminary study. *Supply Chain Management: An International Journal*, *11*(3), pp.208-215.

Maghsoudi, A. and Pazirandeh, A., 2016. Visibility, resource sharing and performance in supply chain relationships: insights from humanitarian practitioners. *Supply Chain Management: An International Journal*, *21*(1), pp.125-139.

Majeed, A.A. and Rupasinghe, T.D., 2017. Internet of Things (IoT) Embedded Future Supply Chains for Industry 4.0: An Assessment from an ERP-based Fashion Apparel and Footwear Industry. *International Journal of Supply Chain Management*, *6*(1), pp.25-40.

McDonald, R.P., 2013. *Test theory: A unified treatment*. Psychology Press. Mahwah, NJ: Erlbaum.

Moyano-Fuentes, J., Sacristán-Díaz, M. and Garrido-Vega, P., 2016. Improving supply chain responsiveness through advanced manufacturing technology: the mediating role of internal and external integration. *Production Planning & Control*, *27*(9), pp.686-697.

Murphy, N. 2017. The Biggest Driver of Digital Transformation in the Supply Chain: End-to-End Visibility. Available from: https://evrythng.com/the-biggest-driver-of-digital-transformation-in-the-supply-chain-end-to-end-visibility/. Accessed on 10 August 2017.

Musa, A., Gunasekaran, A. and Yusuf, Y., 2014. Supply chain product visibility: Methods, systems and impacts. *Expert Systems with Applications*, *41*(1), pp.176-194.

Nair, P.R. and Anbuudayasankar, S.P., 2016. Tackling supply chain management through RFID: Opportunities and challenges. In *Proceedings of the International Congress on Information and Communication Technology* (pp. 467-475). Springer, Singapore.

Ngai, E.W., Chau, D.C. and Chan, T.L.A., 2011. Information technology, operational, and management competencies for supply chain agility: Findings from case studies. *The Journal of Strategic Information Systems*, *20*(3), pp.232-249.

O'Leary-Kelly, S.W. and Vokurka, R.J., 1998. The empirical assessment of construct validity. *Journal of operations management*, *16*(4), pp.387-405.

Pereira, T., Barreto, L. and Amaral, A., 2017. Network and information security challenges within Industry 4.0 paradigm. *Procedia Manufacturing*, *13*, pp.1253-1260.

Petrick, I.J. and Simpson, T.W., 2013. 3D printing disrupts manufacturing: how economies of one create new rules of competition. *Research-Technology Management*, *56*(6), pp.12-16.

Pflaum, A., Bamberg, O.F.U., Prockl, G., Bodendorf, F. and Chen, H., 2019. The Digital Supply Chain of the Future: From Drivers to Technologies and Applications. In *The 52nd Hawaii International Conference on System SciencesHawaii International Conference on System Sciences* (pp. 5153-5154). Hawaii International Conference on System Sciences (HICSS).

Pflaum, A., Bodendorf, F., Prockl, G., & Chen, H. (2017, January). Introduction to The Digital Supply Chain of the Future: Technologies, Applications and Business Models Minitrack. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. pp: 4179-4181

Preuveneers, D., Joosen, W. and Ilie-Zudor, E., 2017. Trustworthy data-driven networked production for customer-centric plants. *Industrial Management & Data Systems*, *117*(10), pp.2305-2324.

Pu, G., Li, Y. and Bai, J., 2018, August. The effect of big data analytics on firms sustainable competitive advantage of quality: a theory framework. In *Proceedings of the 2018 International Conference on Information Management & Management Science* (pp. 33-37). ACM.

Raman, S., Patwa, N., Niranjan, I., Ranjan, U., Moorthy, K. and Mehta, A., 2018. Impact of big data on supply chain management. *International Journal of Logistics Research and Applications*, *21*(6), pp.579-596.

Riggins, F.J. and Wamba, S.F., 2015, January. Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics. In *IEEE 48th Hawaii International Conference on System Sciences (HICSS)* (pp. 1531-1540).

Rivard, S., & Huff, L. (1988). Factors of success for end-user computing. Communication of the ACM, *31*(5), pp. 552-561.

Ryan, P.J. and Watson, R.B., 2017. Research Challenges for the Internet of Things: What Role Can OR Play?. *Systems*, *5*(1), p.24.

Santos, J.R.A., 1999. Cronbach's alpha: A tool for assessing the reliability of scales. *Journal of extension*, *37*(2), pp.1-5.

Srinivasan, R. and Swink, M., 2018. An investigation of visibility and flexibility as complements to supply chain analytics: An organizational information processing theory perspective. *Production and Operations Management*, *27*(10), pp.1849-1867.

Tseng, M.L., Wu, K.J. and Nguyen, T.T., 2011. Information technology in supply chain management: a case study. Procedia-Social and Behavioral Sciences, 25, pp.257-272.

Valmohammadi, C., 2016. Examining the perception of Iranian organizations on Internet of Things solutions and applications. *Industrial and Commercial Training*, *48*(2), pp.104-108.

Wang, Y., Singgih, M., Wang, J. and Rit, M., 2019. Making sense of blockchain technology:(How) will it transform supply chains?. *International Journal of Production Economics*. 2011, pp.221-236.

Wei, S., Ke, W., Lado, A.A., Liu, H. and Wei, K.K., 2019. The Effects of Justice and Top Management Beliefs and Participation: An Exploratory Study in the Context of Digital Supply Chain Management. *Journal of Business Ethics*, pp.1-21.

Westerman, G., Bonnet, D. and McAfee, A., 2014. *Leading digital: Turning technology into business transformation*. Harvard Business Press.

Wiseman, C., 2019. Establishing the Program Management Office: A Key Enabler for Digital Supply Chain Transformation. In *Technology Optimization and Change Management for Successful Digital Supply Chains* (pp. 107-127). IGI Global.

Yee-Loong Chong, A., Ooi, K.B., Lin, B. and Yi Tang, S., 2009. Influence of interorganizational relationships on SMEs'e-business adoption. *Internet Research*, *19*(3), pp.313-331.

DECISION SCIENCES INSTITUTE

Maturity Model for Digitalization: Insights from a Developing Country Context

Fatima Gillani
Lahore University of Management Sciences
15080009@lums.edu.pk

Kamran Ali Chatha
Lahore University of Management sciences
kamranali@lums.edu.pk

Shakeel Sadiq Jajja
Lahore University of Management Sciences
ssj@lums.edu.pk

ABSTRACT

With the advent of digital technologies, digitalization is no more an option for firms but a compulsion for survival. The extant practitioner and academic literature frequently state the confusion on how successful digital transformation of firms should take place. Hence, it is important to identify the specific organizational and technological elements which need to be fortified to achieve digitalization. This study proposes a Digitalization Maturity Model (DMM) and identifies six organizational dimensions critical for a successful digital transformation which mature across four stages. The proposed framework will be validated from practitioners through semi structured interviews and an updated version will be finalized after incorporation of suggestions and feedbacks. The final maturity model will provide firms a comprehensive digital strategy for their successful digital transformation.

KEYWORDS: Digitalization, digital transformation, maturity model

INTRODUCTION

The next big thing for the business world is the rise of digital technologies and their diffusion in companies which is consequently changing the traditional business landscape. Digitalization refers to “the integration of digital technologies to transform activities, processes, actors, and goods from analogue to digital to facilitate new forms of value creation” (Hagberg et al., 2016). Value creation by these technologies can be leveraged when the firm has a clear and well-designed digital strategy. This digital strategy is crucial for firms as it directs them towards the right path and guides them about the appropriate set of actions which need to be carried out to achieve digitalization (Hess et al., 2016). Existing studies mention the use of relevant technologies (Liao et al., 2017; Rüßmann et al., 2015) and their benefits (Cenamor et al., 2017; Gunasekaran et al., 2018; Hannola et al., 2018) yet, is scarce on the guidance of its realization and implementation of digital strategy in firms (El Sawy et al., 2016; Hess et al., 2016). The shift in role of employees, changes in organization structure, enhancement of skills and competences, emergence of systems thinking and an alignment of the IT-business strategies call for a new way of strategizing (Lee et al., 2012; Tripsas, 2009).

The extant literature repeatedly mentions the need of an implementation framework of digital technologies (Strozzi et al., 2017). The lack of clarity in terms of the firms digital transformation journey has been mentioned in several studies (Horváth et al., 2019; Yeh et al., 2018).

Zangiacomi et al. (2017) states that the digital transformation of an organization is a complex process and the lack of clarity about the elements involved can result in failed transformations. The specific aspects which need to be focused, what capabilities shall be developed, which competences are required and which organizational form shall best serve the digital transformation need to be investigated and identified (El Sawy et al., 2010; Kahre et al., 2017; Wu et al., 2016). Firms require guidelines as to which set of tasks, activities and technologies they should focus on in order to achieve digitalization (Boutetière et al., 2018; Breunig et al., 2016; El Sawy et al., 2016). Keeping in view this important research gap this study aims to answer the following research questions:

RQ1: What are the different stages of the digital transformation journey an organization goes through?

RQ2(a): What are the critical organizational factors which are modified at each stage of the transformation journey?

RQ2(b): What changes need to be made in the organizational elements to lead a successful digital transformation?

The maturity model we propose intends to identify different hierarchical progression stages for all firms which lie on the continuum from traditional state to digitalized state. This study aims to provide firms with a guideline for their transformation to the digitalized state. The aim is to identify those dimensions or aspects of digitalization which need to be focused and developed so that a firm achieves digital maturation. The need for a maturity model is of critical importance at this point in time as firms are in need to know a systematic evolution path which would guide their digital transformation (Kane et al., 2015). Maturity models serve as an assessment tool where firms can assess their current state and take guidance on their to-be state through the evolution shown (Becker et al., 2009).

LITERATURE REVIEW

The importance of Maturity Models has been well versed in the extant literature. MM aids in the transformational journey of organizations providing them with a stepwise pathway towards their final goal. They serve as assessment and improvement tools which provide an as-is picture of the firm and also the way forward. It is a framework which describes the evolution of a particular concept across different levels. The different levels represent the extent of sophistication and maturity for each particular knowledge area (Guédria et al., 2015). These knowledge areas vary according to the domain being studied and are derived from the relevant literature. This multistage development serves to provide a guideline for effective hierarchical progression which cannot be undone or reversed easily (Poeppelbuss et al., 2011).

The concept of MM has been mostly used in the IS literature for describing the maturity of different concepts in the field. A few examples are the Nolan's Growth Stage Model (Nolan, 1973), Capability Maturity Model (CMM) (Paulk et al., 1993), Business Process Management Maturity Model (BPMM) which was a variation of the CMM and focused on how firms could achieve excellence in their business processes (De Bruin et al., 2005; Röglinger et al., 2012; Tarhan et al., 2016). Since then, a lot of work has been done in different disciplines which study and propose the maturity of different aspects such as supply chain management maturity (Lockamy III et al., 2004), IT-Business alignment maturity (Luftman, 2004), e-business maturity (Gottschalk, 2009), e-government maturity (Andersen et al., 2006; Darem et al., 2010; Kim et al., 2010), Interoperability maturity (Guédria et al., 2015), IT architecture maturity (Eckhardt et al., 2014; Ross, 2003), IS/ICT management maturity (Curry et al., 2012; Renken, 2004), Enterprise Resource Planning System use (Holland et al., 2001; Nightingale et al., 2002) and informatics capability (Cosic et al., 2012; Liaw et al., 2017). The operations and supply chain literature has not appreciated the use of maturity models and therefore has a dearth of such studies. Only a few studies use the maturity model concept and describe the maturity of supply chain management processes (Lockamy III & McCormack, 2004) and maturity of supply chain analytics in logistics management (Wang et al., 2016).

Maturity Models and Digitalization

In light of the above maturity models, this study proposes a digitalization maturity model and highlights the essential elements which play a pivotal role in the digital transformation. Digitalization refers to the "integration of digital technologies to transform activities, processes, actors, and goods from analogue to digital to facilitate new forms of value creation" (Hagberg et al., 2016). The business world of today competes in the ability of firms to utilize the relevant information and create value out of it. On part of the firms, this requires the right strategies and the right mix of people, processes, technologies and organization design, all of which complement one another to achieve the desired value (Yeow et al., 2018). The new era requires new digital capabilities which aid in making use of digital technologies and extracting benefit from them.

In the realm of digitalization, a few industry reports have proposed different variations of maturity models for the digitalization of businesses (Deloitte, 2018; Leadership, 2019; Solis et al., 2016; tmforum, 2019). Similarly, a conference article by Shahiduzzaman et al. (2018) also talks about the maturity of digital businesses. These studies have identified a multitude of basic dimensions which are important for digitalization by providing a birds-eye view of the changes these dimensions go through. However, these studies are descriptive and do not provide enough guidance about the stages of digital maturity for a business. The proposed Digitalization Maturity Model (DMM) provides insights on the dimensions, stages of progression and the cumulative nature of value created and capability development at each stage of digitalization.

Methodology

The methodology for this study consists of two parts. In the first part, the literature on the maturity model and digitalization is reviewed to propose a Digitalization Maturity Model. While the second part will cater the validation of the proposed maturity model. Description of both the parts is given below.

1. Proposing a Digitalization Maturity Model

The aim of this study is to develop a maturity model for the digitalization of organizations which could guide firms in their digital transformation journeys irrespective of which level the firm is at. We aim to propose strategic guidelines for transformation which could help digitally naive firms, firms with mediocre digital proficiency or firms which are about to reach their digital maturity. The two important components of the maturity model i.e. levels and dimensions are literature driven and are specific to each domain (Van Looy et al., 2013). The dimensions for the maturity model were finalized based on the review of digitalization and maturity model literature. A similar approach has been followed in earlier studies of maturity model development (Gottschalk, 2009; Guédria et al., 2015; Holland & Light, 2001; Kim & Grant, 2010; Rosemann et al., 2004).

The selection of dimensions for a maturity model depends on the specific discipline which is being studied. The literature on that discipline informs about the factors which are crucial for its development and progression (Pöppelbuß et al., 2011). The two major benchmark maturity models which serve as a guide for others are the Capability Maturity Model and the Business Process Management Maturity Model. Studies in the IS domain usually follow the CMM while in management literature the BPMM is usually followed. The dimensions can be segregated into the technology-related dimensions and non-technology related dimensions.

However, these dimensions are yet to be identified for the digitalization of organizations and their hierarchical trajectory needs to be investigated. The adoption of digital technologies does not assure a firms digitalization rather it's the collection of different resources and activities which need to be orchestrated such that the firms' objectives are achieved (Loonam et al., 2018; Sirmon et al., 2011). Multiple studies have proposed infrastructure, technology, people & skills, organization changes, strategy and culture as the most important dimensions which aid in shaping the digitalization of firms (Ardolino et al., 2018; Brettel et al., 2014; Brynjolfsson et al., 2000; Gölzer et al., 2017; Kamble et al., 2018; Li, 2018; Oswald et al., 2017; ÖZTÜRK, 2017; Wulf et al., 2017). The aforementioned studies provide guidance and a direction for digital transformation of firms and inform the critical aspects which need to be developed and enhanced to achieve digitalization.

2. Model Validation

To test the proposed DMM we will conduct structured interviews with practitioners from industry. These practitioners will be from the top management of firms who are involved in the strategic decision making and digital transformation of their firms. As the model is not constrained to a specific industry and can be applicable to all types of organizations. We intend to conduct interviews with 15-20 managers belonging to the Manufacturing Sector, Banking, Telecommunication and Software Industry. We will use the snowball approach to identify respondents suitable to validate the proposed DMM. The managers will be given the model and explained the transformation from Level 1 to Level 4 across all dimensions and asked if the model applies. The suggestions and recommendations from these managers will be incorporated in the proposed model to come up with a final version of the DMM which will incorporate industrial feedback and validation. If necessary, NVIVO software tool will be used to segregate and categorize the content of interview feedback.

DIGITALIZATION MATURITY MODEL

The DMM we propose constitutes of four levels of maturity across six dimensions derived from the review of digital transformation and maturity model literature which has been mentioned above. Based on the aforementioned factors, we select the dimensions of our Digitalization Maturity Model (DMM) to be Human Capital, Organization structure, Operations, Digital Technology, Intra-enterprise integration and Data. A brief description of each of the dimensions is given below. A level-wise description of each dimension and how it evolves is given in table 1.

Dimensions of Digitalization Maturity

Human capital is defined as "the knowledge that employees take with them when they leave the firm, including individual knowledge, skills and abilities" (Kim & Grant, 2010). Human Capital is a crucial element for any change management journey as employees are the ones who must manage the change and adapt according to the transformation requirements. *Organization Structure* depicts the "power and responsibility structure formed in the managing process" (Ramezan, 2011). The organization structure plays a critical role in leveraging the use f its resources which are cumulatively at work to create value for the firm. *Operations* are defined as "the set of activities that creates value in the form of goods and services by transforming inputs into outputs" (Heizer et al., 2017). As the firm evolves and adapts to the changing market dynamics the way operations are conducted within an organization also changes. *Digital Technology* refers to those technologies which characterize the digitalization era such as Internet of Things, Additive Manufacturing, Cyber-Physical systems, Cloud and Big Data Analytics (Rüßmann et al., 2015). *Intra-Enterprise Integration* is represented as “the systems being connected, data is shared in the organization, applications can interoperate and the business processes are coordinated with IS and other processes” (Giachetti, 2004). *Data* is defined as “elementary descriptions of objects’ properties (e.g. activities, events and/or transactions) generated by observation with different forms” (Ardolino et al., 2018).

Levels of Digitalization Maturity

Level 1 - Computerization is defined as the use of computers and telecommunication technologies in organizations which changes the roles & responsibilities and levels of organizational efficiency (Kraut et al., 1989). It enables firms to carry out minor and very basic tasks with the help of the internet and few computers in the organization.
Level 2 – Interaction depicts the upgradation of the technology and characterizes the interaction the new technology initiates among the different levels of the organization. The technology at this level facilitates connectivity and interaction within and between hierarchical levels.
Level 3 – Integration depicts the use of technologies which facilitate decentralization in organizations and enable integration. The functional silos eliminate, and the processes of each department become interdependent.
Level 4 – Continuous Improvement is the final stage of the maturity model depicts the digitalization of the firm and is characterized by defined and repeatable processes which are highly dependent on IT systems in the firm. This stage is characterized by the conversion of raw data into actionable information with the help of analytics capability and technology. The IT systems in the firm develop the ability to foresee internal and external changes and respond proactively. Similarly, the employees at this stage are specialized in their specific knowledge domains and are always at an outlook of improving processes and taking technological initiatives for the betterment of the firms' survival and profitability.

Levels / Dimensions	**Level 1** ***(Computerization)***	**Level 2** ***(Interaction)***	**Level 3** ***(Integration)***	**Level 4** ***(Continuous Improvement)***
Capability	Uninterrupted Routine Tasks - Traditional way of work	Connectivity - Technical interconnectivity	Visibility - IT capabilities leveraged	Digital Dexterity - Business intelligence
Digital Technology	Internet - Traditional state with minimal technology	Internet of Things - tags, sensors, RFID start being used for specific tasks	Cloud Computing - Cloud serves as a data storage mechanism	BDPA - Utilization of data through different tools
Human Capital	Digital Novice - Production workers/frontline workers - High cynicism	Digital Competent - Problem Solvers - Gain knowledge and skills for technical initiatives	Digital Proficient - Employees are able to manage machine and systems	Digital Expert - knowledge workers - Data scientists and software engineers
Operations	Mechanized - Mostly manual operations	Islands of Automation - IT incorporated in business processes to some extent	Initial Digitalization - Majority tasks and processes are automated	Digitalization - Automatic monitoring of processes

Gillani et al., Digitalization Maturity Model

Data	Data Generation - Information generation is minimal - No storage	Data Acquisition - Data acquired from web and other sources - Data transmitted to some adhoc storage	Data Storage - Processes for data storage and processing - Information generated in the organization is digitized	Data Analysis - Data is inspected, transformed and analyzed to extract value from it
Organization	Mechanistic - Higher levels of hierarchy - Vertical decision making	Coordination - Coordination between departments increases for technical issues	Decentralization - Cross functional teams are formed - Decision making allocated to subunits to some extent	Organic - Employees are empowered - Data driven culture - Minimal hierarchy
Intra-enterprise integration	Network Level - Physical heterogeneity of machines or devices - Integration goal in connectivity	Data Level (vertical integration) - Integration goal is data sharing between subsystems and organizational units	Application Level (horizontal integration) - Integration goal is interoperability	Business Process Level - Tight business integration is achieved - Applications support
Value Created	Local Optimization - optimize a specific task	Interoperability - IT efficiency - Reduced cost	Integration - Process optimization	Strategic Agility - Speed to market - responsiveness

Table 1: Proposed Digitalization Maturity Model

References

Andersen, K. V., & Henriksen, H. Z. (2006). *E-government maturity models: Extension of the Layne and Lee model.* Government information quarterly, *23*(2), 236-248.

Ardolino, M., Rapaccini, M., Saccani, N., Gaiardelli, P., Crespi, G., & Ruggeri, C. (2018). *The role of digital technologies for the service transformation of industrial companies.* International Journal of Production Research, *56*(6), 2116-2132.

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). *Developing maturity models for IT management.* Business & Information Systems Engineering, *1*(3), 213-222.

Boutetière, H., Montagner, A., & Reich, A. J. M. c. (2018). Unlocking success in digital transformations. Retrieved from mckinsey.com website: https://www.mckinsey.com/business-functions/organization/our-insights/unlocking-success-in-digital-transformations

Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). *How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective.* International Journal of Mechanical, Industrial Science and Engineering, *8*(1), 37-44.

Breunig, M., Kelly, R., Mathis, R., & Wee, D. J. M. Q., April. (2016). Getting the most out of Industry 4.0. https://www.mckinsey.com/business-functions/operations/our-insights/industry-40-looking-beyond-the-initial-hype

Brynjolfsson, E., & Hitt, L. M. (2000). *Beyond computation: Information technology, organizational transformation and business performance.* Journal of Economic perspectives, *14*(4), 23-48.

Cenamor, J., Sjödin, D. R., & Parida, V. (2017). *Adopting a platform approach in servitization: Leveraging the value of digitalization.* International Journal of Production Economics, *192*, 54-65.

Cosic, R., Shanks, G., & Maynard, S. (2012). *Towards a business analytics capability maturity model.* Paper presented at the ACIS 2012: Location, Location, Location: Proceedings of the 23rd Australasian Conference on Information Systems 2012.

Curry, E., Guyon, B., Sheridan, C., & Donnellan, B. (2012). *Developing a sustainable IT capability: Lessons from Intel's journey.* MIS Quarterly Executive, *11*(2).

Darem, A. B., & Ahmed, A. (2010). *e-Government maturity models.* Asian Journal of Development Matters, *4*(2), 29-36.

De Bruin, T., & Rosemann, M. (2005). *Towards a business process management maturity model.*

Deloitte. (2018). Digital Maturity Model

Achieving digital maturity to drive growth: Deloitte.

Eckhardt, A., Laumer, S., Maier, C., & Weitzel, T. (2014). *The transformation of people, processes, and IT in e-recruiting: Insights from an eight-year case study of a German media corporation.* Employee Relations, *36*(4), 415-431.

El Sawy, O. A., Kræmmergaard, P., Amsinck, H., & Vinther, A. L. (2016). *How LEGO Built the Foundations and Enterprise Capabilities for Digital Leadership.* MIS Quarterly Executive, *15*(2).

El Sawy, O. A., Malhotra, A., Park, Y., & Pavlou, P. A. (2010). *Research commentary—seeking the configurations of digital ecodynamics: It takes three to tango.* Information Systems Research, *21*(4), 835-848.

Giachetti, R. E. (2004). *A framework to review the information integration of the enterprise.* International Journal of Production Research, *42*(6), 1147-1166.

Gölzer, P., & Fritzsche, A. (2017). *Data-driven operations management: organisational implications of the digital transformation in industrial practice.* Production Planning & Control, *28*(16), 1332-1343.

Gottschalk, P. (2009). *Maturity levels for interoperability in digital government.* Government information quarterly, *26*(1), 75-81.

Guédria, W., Naudet, Y., & Chen, D. (2015). *Maturity model for enterprise interoperability.* Enterprise Information Systems, *9*(1), 1-28.

Gunasekaran, A., Yusuf, Y. Y., Adeleye, E. O., & Papadopoulos, T. (2018). *Agile manufacturing practices: the role of big data and business analytics with multiple case studies.* International Journal of Production Research, *56*(1-2), 385-397.

Hagberg, J., Sundstrom, M., & Egels-Zandén, N. (2016). *The digitalization of retailing: an exploratory framework.* International Journal of Retail & Distribution Management, *44*(7), 694-712.

Hannola, L., Richter, A., Richter, S., & Stocker, A. (2018). *Empowering production workers with digitally facilitated knowledge processes–a conceptual framework.* International Journal of Production Research, 1-15.

Heizer, J., Render, B., Munson, C., & Sachan, A. (2017). Operations management: sustainability and supply chain management, 12/e: Pearson Education.

Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2016). *Options for Formulating a Digital Transformation Strategy.* MIS Quarterly Executive, *15*(2).

Holland, C. P., & Light, B. (2001). *A stage maturity model for enterprise resource planning systems use.* ACM SIGMIS Database: the DATABASE for Advances in Information Systems, *32*(2), 34-45.

Horváth, D., & Szabó, R. Z. (2019). *Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities?* Technological Forecasting and Social Change, *146,* 119-132.

Kahre, C., Hoffmann, D., & Ahlemann, F. (2017). *Beyond business-IT alignment-Digital business strategies as a paradigmatic shift: A review and research agenda.*

Kamble, S. S., Gunasekaran, A., & Sharma, R. (2018). *Analysis of the driving and dependence power of barriers to adopt industry 4.0 in Indian manufacturing industry.* Computers in Industry, *101,* 107-119.

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). *Strategy, not technology, drives digital transformation.* MIT Sloan Management Review and Deloitte University Press, *14,* 1-25.

Kim, D.-Y., & Grant, G. (2010). *E-government maturity model using the capability maturity model integration.* Journal of Systems and Information Technology, *12*(3), 230-244.

Kraut, R. E., Dumais, S. T., & Koch, S. J. C. o. t. A. (1989). *Computerization, productivity, and quality of work-life. 32*(2), 220-238.

Leadership, D. (2019). Digital Maturity Framework. from https://www.digitalleadership.ltd/digital-maturity-framework

Lee, J., & Berente, N. (2012). *Digital innovation and the division of innovative labor: Digital controls in the automotive industry.* Organization science, *23*(5), 1428-1447.

Li, F. (2018). *The digital transformation of business models in the creative industries: A holistic framework and emerging trends.* Technovation.

Liao, Y., Deschamps, F., Loures, E. d. F. R., & Ramos, L. F. P. (2017). *Past, present and future of Industry 4.0-a systematic literature review and research agenda proposal.* International Journal of Production Research, *55*(12), 3609-3629.

Liaw, S.-T., Kearns, R., Taggart, J., Frank, O., Lane, R., Tam, M., Dennis, S., et al. (2017). *The informatics capability maturity of integrated primary care centres in Australia.* International journal of medical informatics, *105,* 89-97.

Lockamy III, A., & McCormack, K. (2004). *The development of a supply chain management process maturity model using the concepts of business process orientation.* Supply Chain Management: An International Journal, *9*(4), 272-278.

Loonam, J., Eaves, S., Kumar, V., & Parry, G. (2018). *Towards digital transformation: Lessons learned from traditional organizations.* Strategic Change, *27*(2), 101-109.

Luftman, J. (2004). Assessing Business-IT Allignment Maturity *Strategies for information technology governance* (pp. 99-128): Igi Global.

Nightingale, D. J., & Mize, J. H. (2002). *Development of a lean enterprise transformation maturity model.* Information Knowledge Systems Management, *3*(1), 15-30.

Nolan, R. L. (1973). *Managing the computer resource: a stage hypothesis.* Communications of the ACM, *16*(7), 399-405.

Oswald, G., & Kleinemeier, M. (2017). *Shaping the Digital Enterprise*: Springer.

ÖZTÜRK, D. (2017). *Technological transformation of manufacturing by smart factory vision: industry 4.0.* International Journal of Development Research, *7*(11), 17371-17382.

Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). *The capability maturity model for software.* Software engineering project management, *10,* 1-26.

Poeppelbuss, J., Niehaves, B., Simons, A., & Becker, J. (2011). *Maturity models in information systems research: Literature search and analysis.* CAIS, *29*(1), 1-15.

Pöppelbuß, J., & Röglinger, M. (2011). *What makes a useful maturity model? a framework of general design principles for maturity models and its demonstration in business process management.* Paper presented at the ECIS.

Ramezan, M. (2011). *Intellectual capital and organizational organic structure in knowledge society: How are these concepts related?* International Journal of Information Management, *31*(1), 88-95.

Renken, J. (2004). *Developing an IS/ICT management capability maturity framework.* Paper presented at the Proceedings of the 2004 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries.

Röglinger, M., Pöppelbuß, J., & Becker, J. (2012). *Maturity models in business process management.* Business Process Management Journal, *18*(2), 328-346.

Rosemann, M., De Bruin, T., & Hueffner, T. (2004). *A model for business process management maturity.* ACIS 2004 Proceedings, 6.

Ross, J. (2003). *Creating a strategic IT architecture competency: Learning in stages.*

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). *Industry 4.0: The future of productivity and growth in manufacturing industries.* Boston Consulting Group, *9*.

Shahiduzzaman, M., & Marek, K. (2018). *Digital organisation: A value centric model for digital transformation.*

Sirmon, D. G., Hitt, M. A., Ireland, R. D., & Gilbert, B. A. (2011). *Resource orchestration to create competitive advantage: Breadth, depth, and life cycle effects.* Journal of management, *37*(5), 1390-1412.

Solis, B., & Szymanski, J. J. A., a Prophet Company. (2016). *The six stages of digital transformation maturity.*

Strozzi, F., Colicchia, C., Creazza, A., & Noè, C. (2017). *Literature review on the 'Smart Factory'concept using bibliometric tools.* International Journal of Production Research, *55*(22), 6572-6591.

Tarhan, A., Turetken, O., & Reijers, H. A. (2016). *Business process maturity models: A systematic literature review.* Information and software technology, *75,* 122-134.

tmforum. (2019). Digital Maturity Model

An online tool to navigate the maze of digital transformation. from https://www.tmforum.org/digital-maturity-model-metrics/

Tripsas, M. (2009). *Technology, identity, and inertia through the lens of "The Digital Photography Company".* Organization science, *20*(2), 441-460.

Van Looy, A., De Backer, M., Poels, G., & Snoeck, M. (2013). *Choosing the right business process maturity model.* Information & Management*, 50*(7)*,* 466-488.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. J. I. J. o. P. E. (2016). *Big data analytics in logistics and supply chain management: Certain investigations for research and applications. 176,* 98-110.

Wu, L., Yue, X., Jin, A., & Yen, D. C. (2016). *Smart supply chain management: a review and implications for future research.* The International Journal of Logistics Management*, 27*(2)*,* 395-417.

Wulf, J., Mettler, T., & Brenner, W. (2017). *Using a Digital Services Capability Model to Assess Readiness for the Digital Consumer.* MIS Quarterly Executive*, 16*(3).

Yeh, C.-C., & Chen, Y.-F. (2018). *Critical success factors for adoption of 3D printing.* Technological Forecasting and Social Change*, 132,* 209-216.

Yeow, A., Soh, C., & Hansen, R. (2018). *Aligning with new digital strategy: A dynamic capabilities approach.* The Journal of Strategic Information Systems*, 27*(1)*,* 43-58.

Zangiacomi, A., Oesterle, J., Fornasiero, R., Sacco, M., & Azevedo, A. (2017). *The implementation of digital technologies for operations management: a case study for manufacturing apps.* Production Planning & Control*, 28*(16)*,* 1318-1331.

DECISION SCIENCES INSTITUTE

Towards best practices in manufacturing data analytics

Christoph Wunck
Emden/Leer University of Applied Sciences, Emden, Germany
OFFIS Institute for Information Technology, Oldenburg, Germany
Iowa State University, Department of Computer Science
Email: christoph.wunck@hs-emden-leer.de

ABSTRACT

This study identifies best practices for the operational implementation of data analytics in manufacturing. Case studies from literature and personal research are evaluated to find potential areas where useful practices have been observed. The results show that best practices can be identified in analytics architecture, interdisciplinary communication, and systems modeling.

KEYWORDS: Best practices, manufacturing analytics, system modeling, interdisciplinary communication, analytics architecture

INTRODUCTION

Data analytics has become a major source of competitive advantage in many industries. The concept of Analytics Capability has been defined as an important strategic resource to improve operational performance (Srinivasan & Swink, 2018). Usually big data architectures, algorithms and methods are regarded as generally applicable, once the data have been transferred into the analytics applications' data stores. Sensors are seen as a source of big data, as well as web pages, log files, e-mail, and documents (Katal et al. 2013). The manufacturing industry in particular is the focus of numerous activities when it comes to the extraction of valuable information from data collected at the shop floor (Rameshwar et al. 2019). However, the manufacturing domain differs from other domains particularly due to the following characteristics (Wunck & Baumann, 2017):

- The machines, tools and devices, i.e. the physical representations of the data sources, can have an expected useful life of more than 30 years.
- Due to the long operating life, manufacturers need a migration path to implement innovative applications without being forced to prematurely renew their machines.
- Legacy machines need to be upgraded with additional sensors, processors and interfaces. This upgrade process is known as a retrofit process.

Other important characteristics of manufacturing systems are:

- Manufacturing systems have been expressly designed for a purpose. Causal relationships between system objects are usually clear and need not be rediscovered by analytics algorithms.
- Manufacturing systems have a distinct boundary between the physical processes themselves and the logical event stream derived from process behavior. The physical

processes transform material and energy during a certain timespan and can be modeled by differential equations, while changing conditions in the processes or external stimuli are best modeled by discrete events. Data analytics applications should exploit this wealth of a-priori-information rather than discard it.
- The ubiquitous availability of cheap IoT (Internet of Things) devices with integrated sensors, processors and wireless networking gave rise to a multitude of small engineering startups offering retrofitting services for traditional manufacturers. Many of these engineering firms have a strong background in computer science, but lack profound knowledge from both digital signal processing and manufacturing engineering.
- Domain specialists such as manufacturing engineers need to verify requirements for analytics software written by data scientists with a computer science background. This process is difficult as both parties have different viewpoints on a manufacturing system.

When it comes to applications like data mining or knowledge discovery, the strong physical foundation of manufacturing processes requires certain precautions, which oftentimes are not known to data scientist coming from the enterprise or business intelligence domain. This study aims to outline some of the necessary preconditions for data analytics implementations in manufacturing to be successful.

LITERATURE REVIEW

The literature review focuses on data analytics implementations in the manufacturing domain and aims to identify potential knowledge areas that are unique as compared to other application domains.

Common applications for knowledge discovery and analysis in manufacturing include (Polczynski & Kochanski, 2010): Detection of root causes of deteriorating product quality, identification of critical and optimal manufacturing process parameters, prediction of effects of manufacturing process changes, and identification of root causes and prediction of equipment breakdown. Works by Gröger et al. (2012) find data mining applications in manufacturing limited to quality analysis, failure analysis, maintenance analysis, and production planning analysis.

A conceptual architecture for data-driven manufacturing process optimizations is proposed by Gröger et al. (2012). Their Advanced Manufacturing Analytics Platform comprises three layers, data integration, process analytics, and process optimization. This corresponds roughly to the three layers of data discovery, data integration, and data exploitation suggested by Miller & Monk (2013). For generic IoT applications, a model with five layers, sensor, gateway, middleware, application, and business has been proposed (Wu et al. 2010). Wunck & Baumann (2017) introduce an architecture model from the perspective of value added by information processes. Their IVCOR model is adapted from the SCOR process model and utilizes the analogies between physical products and information products.

Rabelo et al. (2003) emphasize the strong connection between systems dynamic modeling by differential equations and discrete event modeling supply chain simulations. Their hybrid modeling approach for supply chain simulation combines dynamic and discrete system models, which is a modeling approach also useful for localized manufacturing systems. Castillo & Smith (2002) review various formal modeling methodologies for control of manufacturing cells. They conclude that language-based methodologies (e.g. process algebra, temporal logic) are not friendly to potential users, as these are more abstract than machine-based methodologies (e.g. finite automata, Petri nets). Negahban & Smith (2014) focus on the application of discrete event simulation to a broad and extensive range of manufacturing system design and operation problems.

The following references illustrate a broad range of data processing methods. Sun et al. (2018) combine sophisticated signal processing algorithms (compressed sensing) with deep learning to monitor and predict the condition of bearings in rotary machinery systems. They present an innovative method for automatic feature extraction and classification, which traditionally is a tedious manual process. Their method uses raw sensor data, sampled according to the Nyquist-Shannon sampling theorem. Zhang et al. (2018) propose a generic data analytics system for manufacturing production. They address small manufacturers who cannot afford a professional data analytics team. Users are not expected to have prior knowledge or experience in data analytics. Mulrennan et al. (2018) model plant energy profiles from historical production schedule data. The data is stored in various sources and does not conform to a common sampling rate or data type. Evans & Boreland (2018) present a study aimed to demonstrate statistical models that can identify variance components in photovoltaic manufacturing. Kibira et al. (2015) integrate data analytics and simulation methods to support manufacturing decision making. They discuss various standards that might support manufacturing analytics at different levels to measure, collect, represent, and exchange data. Sardana et al. (2015) predict the health of a spindle and its bearings using fast fourier transform analysis and a graph clustering algorithm. A wavelet-based method for feature extraction in injection molding is described by Wunck (2017).

Katal et al. (2013) recommend good practices on a very abstract level for big data applications regarding data dimensions, unique keys, data structure, data access technology, privacy concerns, data quality, scalability, and interdisciplinary cooperation. Best practices for managing predictive models in a production environment are summarized by Chu et al. (2007) from the perspective of a software company.

IMPLICATIONS

Based on the findings in the literature and various industry projects carried out by the author, several candidate areas for best practices have been identified. For each area, recommendations and explanations are given.

Analytics Architecture

It is a common characteristic of current Big Data concepts to assume that all data generated by data sources (e.g. sensors in manufacturing) will be transported to cloud storage before any data processing is performed (Wunck & Baumann 2017). Apart from performance issues, this practice contravenes one of the most important engineering principles: the principle of abstraction. Abstraction applied to data analytics implies the existence of layers with different algorithmic tasks corresponding to physical layers in the manufacturing system.

In the physical layer, signals are the dominating data source. Many important signals are analog signals read by suitable sensors. Equidistant sampling as mandated by signal processing theory (e.g. Nyquist-Shannon theorem) yields highly redundant data streams. Sampling in longer time intervals however bears the risk of falsified data due to aliasing effects.

<u>Proposals:</u>

- Data analytics should follow an architecture model that takes into account the different time scales, real time requirements and data volumes in manufacturing systems. Useful layers in an architecture model are the signal layer and the event layer. Signals are processed according the principles of digital signal processing, events are derived from signals and are processed in a higher layer.

- Sensor signals from analog process variables should be sampled with the lowest possible sample rate that conforms to the laws of digital signal processing. After sampling, a data compression method should be applied to the sampled data (e.g.wavelet-based compression or other methods based on orthogonal transformations).

Interdisciplinary Communication

According to Katal et al. (2013), business leaders and IT leaders should work together to yield more business value from the data. While this is certainly true, the interdisciplinary communication in the manufacturing domain requires more than goodwill of the stakeholders involved. The manufacturing domain specialists may be viewed as the customers of an analytics service to be delivered by the data scientists. It is the customers' task to validate the requirements posed on the analytics programs to be developed. If the requirements are phrased as domain requirements, a great risk of misunderstanding on the developers' side exists. If the requirements are too formal, the customers are likely to not validate the requirements correctly. Formal specification, though, is a valuable tool towards fault analysis and error-free implementation, both of which are crucial in manufacturing. As was stated by Castillo & Smith (2002), formal methods such as finite automata are more understandable for manufacturing engineers than other formal methods. Timed regular expressions (Wunck 2015) and domain specific languages (Wunck 2016) turn out to be helpful in bridging the gap between manufacturing domain requirements and formal specifications.

<u>Proposal:</u>

- Requirements for analytics tasks should be specified using a formal approach that is both understandable to manufacturing engineers and useful for analytics implementations.

System Modeling

Manufacturing systems can be modeled either on a physical level (differential equations or similar tools) or on an event stream level. Due to the widespread use of sensor-equipped IoT devices, the physical level and the system dynamics associated with it become more important. Once analog values from physical process variables are used as data sources for analytics programs, sampled system laws such as the Nyquist-Shannon theorem must be observed. This includes the careful design of the measuring chain regarding eigenfrequencies of the systems and cutoff frequencies of low pass filters. Otherwise measured values might be falsified, leading to false conclusions later on.

Many physical processes in manufacturing are feedback-controlled. Some controller algorithms switch to different control strategies during different phases or segments of the process. If analytics is applied using raw sensor data from a process, it might be helpful to precisely identify those segments and apply individual algorithms tailored to the respective segments. For example, in injection molding the machine's screw velocity is under closed-loop control during injection, while in the holding/cooling phase the machine switches to closed-loop control of hydraulic pressure. Usually, different sample rates are used for the different segments (Wunck & Ruoff, 2016).

Engineering control theory sets great value on the stability of computer controlled systems. Stability means that a system's output experiences only small deviations when its inputs are changed by a small amount. System dynamics and thus stability can be modified by feedback loops. Feedback of information might threaten system stability and must therefore be carefully analyzed. Unstable systems have a tendency to oscillate. Measures to prevent unwanted oscillation are damping the system to reduce its response times. The well-known bullwhip effect observed in distribution channels and supply chains is an example of an unstability in a system of firms.

Proposals:

- Manufacturing systems should be modeled using a hybrid approach, which takes into account both the dynamic properties and the discrete event stream.
- Physical manufacturing processes should be partitioned into segments with different sample rates or control strategies before applying analytics at a higher level.
- Analytic procedures that seek to compensate observed deviations by feedback must be carefully analyzed regarding system dynamics and potential instability.

DISCUSSION AND CONCLUSIONS

This paper presented an attempt to collect some best practices for manufacturing analytics. These might be helpful for both domain specialists and analytics implementers. Analytics methods and algorithms are indeed generic and universally applicable. However, the manufacturing domain has some intricacies that must be taken into account for analytics projects to deliver on their promises. Future works might be directed to identify additional areas for best practices in manufacturing.

REFERENCES

Castillo I., & Smith J.S., (2002). Formal Modeling Methodologies for Control of Manufacturing Cells: Survey and Comparison. Journal of Manufacturing Systems 21(1), 40-57.

Chu, R., Duling, D., & Thompson, W. (2007). Best Practices for Managing Predictive Models in a Production Environment. SAS Global Forum 2007, Data Mining and Predictive Modeling.

Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T., (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. British Journal of Management, Vol. 30, 341–361.

Evans, R., & Boreland, M. (2018). Multivariate Data Analytics in PV Manufacturing – Four Case Studies Using Manufacturing Datasets. IEEE Journal of Photovoltaics 8(1), 38-47.

Gröger, C., Niedermann, F., & Mitschang, B. (2012). Data Mining-driven Manufacturing Process Optimization. Proceedings of the World Congress on Engineering (WCE).

Katal, A., Wazid, M., & Goudar, R.H. (2013). Big Data: Issues, Challenges, Tools and Good Practices. Sixth international conference on contemporary computing (IC3), IEEE, 404-409.

Kibira, D., Hatim, Q., Kumara, S., & Shao, G. (2015). Integrating Data Analytics and Simulation Methods to Support Manufacturing Decision Making. Proceedings of the 2015 Winter Simulation Conference, IEEE, 2100-2111.

Miller, H.G., & Monk, P. (2013). From data to decision: a value chain for big data. IT Professional 15, 57-59.

Mulrennan, K., Donovan, J., Tormey, D., & Macpherson, R. (2018). A data science approach to modeling a manufacturing facility's electrical energy profile from plant production data. Proceedings of the IEEE 5th International Conference on Data Science and Advanced Analytics, 387-391.

Negahban, A., & Smith, J.S. (2014). Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems 33, 241-261.

Polczynski, M., & Kochanski, A. (2010). Knowledge Discovery and Analysis in Manufacturing. Quality Engineering, 22(3), 169-181.

Rabelo, L., Helal, M., Son, Y.-J., Jones, A., Min, J., & Deshmukh, A. (2003). A Hybrid Approach to Manufacturing Enterprise Simulation. Proceedings of the 2003 Winter Simulation Conference, 1125-1133.

Sardana, D., Bhatnagar, R., Pavel, R., & Iverson, J. (2015). Data Driven Predictive Analytics for A Spindle's Health. Proceedings of the IEEE International Conference on Big Data, 1378-1387.

Sun, J. Yan, C., & Wen, J. (2018). Intelligent Bearing Fault Diagnosis Method Combining Compressed Data Acquisition and Deep Learning. IEEE Transactions on Instrumentation and Measurement 67(1), 185-195.

Srinavasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. Production and Operations Management, Vol. 27, No. 10, 1849-1867.

Wu, M., Lu, T.-J., Ling, F.Y., Sun, J., & Du, H.-Y. (2010). Research in the architecture of Internet of things. Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), 484-487.

Wunck, C. (2015). Monitoring Factory Floor Events in Injection Moulding Shops Using Timed Regular Expressions. Institute of Industrial Engineers (IIE) Annual Conference & Expo 2015.

Wunck, C. (2016). Implementation of Mobile Event Monitoring Agents for Manufacturing Execution and Intelligence Systems Using a Domain Specific Language. Proceedings of the 2016 International Conference on Industrial Engineering and Operations Management (IEOM), 332-339.

Wunck, C. (2017). Feature extraction methods for quality attributes prediction in injection moulding. Journal of Computational Methods in Sciences and Engineering 17(S1), S71-S83.

Wunck, C., & Baumann, S. (2017). Towards a process reference model for the information value chain in IoT applications. 2017 IEEE European Technology and Engineering Management Summit (E-TEMS), 1-6.

Wunck, C. & Ruoff, M. (2016). Wavelet-based contour synthesis to simulate injection molding process signals. Proceedings of the IEEE 14th International Conference on Industrial Informatics (INDIN), 851-855.

Zhang, H., Wang, H., Li, J., & Gao, H. (2018). A Generic Data Analytics System for Manufacturing Production. Big Data Mining and Analytics 1(2), 160-171.

DECISION SCIENCES INSTITUTE

The Dark Side of Bank CEO Risk-taking Incentives: Evidence from Bank Lending Decisions

Linh Thi Thuy Tran
Yuan Ze University
Taoyuan, Taiwan
E-mail address: linh28689@gmail.com

Chien-Ling Lo
Yuan Ze University
Taoyuan, Taiwan
E-mail: cllo@saturn.yzu.edu.tw

Po-Hsin Ho
National Central University
Taoyuan, Taiwan
E-mail address: phho@mail.ntpu.edu.tw

Chih-Yung Lin
National Chiao-Tung University
Hsinchu, Taiwan
E-mail: d95723009@ntu.edu.tw

ABSTRACT

This paper investigates how bank CEO risk-taking incentives influence bank lending decisions. Consistent with the existing CEO incentive literature, we find that CEOs with higher risk-taking incentives (vega) tend to relax their lending standards in bank loan contracts to pursue higher compensation. We find that banks with a high vega tend to charge a significantly lower loan spread, demand fewer loan covenants, and have lower probability to seek collateral. Results become weaker when banks have strong corporate governance mechanisms, supporting the proposition that high CEO risk-taking incentives may create an agency problem between a bank manager and shareholders.

KEYWORDS: CEO incentives, Bank loan contracts, Corporate governance, Agency problem.

INTRODUCTION

Managerial risk-taking behavior in both financial and non-financial firms has been an attractive focus for the lens of many researchers (Hubbard and Palia, 1995; Houston and James, 1995; Knopf, Nam and Thornton, 2002; Coles, Daniel, and Naveen, 2006; Chen, Steiner and Whyte, 2006; Acharya and Naqvi, 2012). Excessive CEO risk-taking in the financial sector especially has been blamed for playing a crucial role in the build up to the 2008-2009 financial crisis. Acharya and Naqvi (2012) develop a theoretical model to show that bank over-lending may result from managers' desire to receive higher compensation in the presence of an agency problem between a bank manager and shareholders.

Other studies have revealed a positive correlation between option compensation and risk-taking incentives, thus increasing bank risk taking and bank-specific default risk (Jeitschko and Jeung, 2005; Mehran and Rosenberg, 2007; Balachandran, Kogut and Harnal, 2010; Bebchuk, Cohen and Spamann, 2010; Fahlenbrach and Stulz, 2011; Hagendorff and Vallascas, 2011). For example, Coles, Daniel and Naveen (2006) suggest that the higher vega gives executives incentive to implement more aggressive debt policy and invest more in riskier assets (e.g. R&D). Similarly, DeYoung, Peng, and Yan (2013) show that banks in which CEOs have high risk-taking incentives (high-vega banks) exhibit substantially larger amounts of both systematic and idiosyncratic risk. Although recent studies have confirmed that CEO risk-taking incentives increase bank risk exposure, whether such exposure affects bank lending decisions has not to date been examined. Specifically, in this paper we investigate the effects of bank CEOs' risk-taking incentives (vega) on bank loan contracting.

Our paper focuses on private debt financing for several reasons. First, the bank loan market has become a predominant source of external funding in the US (Chava, Livdan and Purnanandam, 2009; Graham, Li and Qiu, 2008). Second, a large number of scholars have discussed the determinants of bank loan contracts, such as borrower reputation (Sufi, 2007), borrower accounting quality (Bharath, Sunder and Sunder, 2008), anti-takeover provisions (Klock, Mansi and Maxwell, 2005), and board quality (Fields, Fraser and Subrahmanyam, 2012). Third, the capabilities and incentives of banks to collect information and monitor borrowers are significantly better than those for public bondholders or investors (e.g., Fama, 1985; Diamond, 1984; Roberts and Sufi, 2009; Roberts, 2015).

We attempt to answer the following four questions regarding the vega effects on bank loan contracts: (i) Do banks with higher vega charge lower interest rates on loans?; (ii) Do vega effects on bank loan contracts also exist in non-price terms?; (iii) Are vega effects eliminated by strong corporate governance mechanisms?; and (iv) Do vega effects still hold after adjusting for other CEO compensation schemes and CEO characteristics?

We evaluate these questions by using a sample of 20,502 loans to 5,102 U.S. firms between 1992 and 2014. We obtain all accounting variables and stock prices from the Compustat database and the Center for Research in Security Prices (CRSP). Corporate-governance and CEO-compensation-related variables come from RiskMetrics and ExecuComp. We collect bank loan data from the DealScan database. Empirical results indicate that CEO risk-taking incentives in term of vega are significantly and negatively correlated with bank loan spread. That is, with high CEO vega incentives, banks charge lower interest rates on loans after controlling for lender and borrower characteristics, macroeconomic factors, and other loan conditions. The evidence remains strong when we control for industry and year fixed effects, loan purpose and type, as well as for CEO characteristics. For example, banks with higher vega will charge a 1.5642 lower basis-point ($e^{0.0199} \times 1.5334 = 1.5642$) loan spread per each one-standard-deviation increase in vega. As the average loan spread in our sample is 142.4940 basis points ($e^{4.9593}$), a one-standard-deviation increase in vega lowers a firm's bank loan interest rate an average of 1.1 percent ($1.5642 \div 142.4940 = 1.1\%$). Furthermore, a lower 1.5642 basis point of loan spread is around 36 percent of the estimated coefficients for the effects of corporate tax avoidance and of social capital in Hasan et al. (2014, 2017). Accordingly, this effect of bank CEO risk-taking incentives on loan spread is not only statistically significant, but is also economically important.

We further find evidence that bank managers with high vega tend to impose fewer covenants on loans (especially general covenants) and are less likely to require collateral. All our main results are statistically and economically significant and are robust to specifications with different sets of explanatory variables (namely, lender, borrower and loan characteristics, and macroeconomics).

In further testing, we examine in greater detail whether the effect of CEO vega derives from the agency cost channel. We argue that, if CEO vega is viewed as a representation of agency cost, then its effect would be outweighed when bank corporate governance quality is weak. Our results provide supportive evidence for this governance channel, confirming that our main findings of a vega incentive effect are driven primarily by banks with weak corporate governance. Hence, our results suggest that the influence of CEO vega derives from an agency problem between bank managers and shareholders.

We further investigate whether banks with high vega would react differently to borrower creditworthiness assessments when a firm takes greater risks. We conduct a subsample analysis by partitioning firms by risk measure (market beta, idiosyncratic and cash-flow volatilities) into high and low groups. The results show that the negative association between vega and loan spread occurs primarily in firms with higher risk-taking strategies.

As a robustness check of our findings, we consider alternative measures of compensation and the effect of CEO characteristics on bank loan contracts and find similar results. In addition, we find that the negative effect of vega on loan spreads becomes stronger for small banks. Finally, we use a change regression to control for the omitted-variable bias and the results remain supportive of our hypothesis.

Our work contributes to the literature in two ways. First, we add to the literature on executive-compensation contracting and bank-lending behaviors. A few studies have recently focused on the CEO risk-taking incentives inherent in bank compensation that affect bank performance, on riskier bank investment choices and the business policies of U.S. banking companies in the short period before and during the financial crisis of 2008 (Acharya and Naqvi, 2012; DeYoung, Peng, and Yan, 2013; Bhagat and Bolton, 2014; Cheng, Hong and Scheinkman, 2015; Chesney, Stromberg and Wagner, 2016; Gande and Kalpathy, 2017). Different from previous studies, we provide evidence of the relationship between vega and loan contracting from a lender's point of view, i.e., we follow the studies of Coles, Daniel, and Naveen (2006, 2013) to calculate CEO risk-taking incentives (vega) and find that banks with high vega compensation may reduce lending interest rates and relax lending standards in the interests of their own wealth. Results become weaker, however, when banks have strong corporate governance mechanisms, supporting the view that high CEO risk-taking incentives may create an agency problem between a bank manager and shareholders. Thus, our results provide a reference for scholars, policy makers, and market investors for assessing the significance of the influence of CEO risk-taking incentive on bank lending decisions.

Second, this paper provides an explanation of the determinants of bank loan contracts from a new perspective: the supply side. The determinants of bank loan contracts from the demand side have already been studied extensively (Strahan, 1999; Sufi, 2007; Bharath, Sunder and Sunder 2008; Graham, Li and Qiu, 2008; Lin et al., 2011), and recently include political connections (Houston, Jiang, Lin, and Ma, 2014), tax avoidance (Hasan, Hoi, Wu, and Zhang, 2014), and social capital (Hasan, Hoi, Wu, and Zhang, 2017). Different from these previous studies, we propose that a key factor in bank loan contracts are CEO risk-taking incentives from the supply side. Based on this view, this paper also adds to the literature on bank loan contracts.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Conflict of interest between shareholders and managers is a key agency problem for firms (Jensen and Meckling, 1976; Jensen, 1986). Firms can use equity-based compensation schemes to motivate CEOs' efforts to take on risky projects, thus ameliorating the principal-agent conflict (Jensen and Meckling, 1976; Myers, 1977; Guay, 1999; Smith and Stulz, 1985; Bolton, Mehran, and Shapiro, 2015). During the financial crisis of 2008, the U.S. economy experienced its worst recession since the Great Depression of the 1930s with the collapse of many of its largest banking and financial institutions. As a result, bank executives faced widespread criticism for driving the excessive risk taking that ignited the disastrous financial crisis. A growing number of academic researchers discuss the impacts of equity-based compensation contracts and whether government should do more to limit executive compensation (Fahlenbrach and Stulz, 2011; Balachandran, Kogut and Harnal. 2010; Hagendorff and Vallascas, 2011; DeYoung, Peng and Yan, 2013; Chesney, Stromberg and Wagner, 2016).

In early studies of deregulation of the banking industry, many authors have made efforts to examine the problem of equity-based compensation and bank performance. For example, Houston and James (1995) and Hubbard and Palia (1995) suggest that compensation contracts create incentives for CEOs to take risks. Equity incentives embedded in CEO compensation contracts are strongly associated with bank performance and this relationship is stronger in deregulated interstate banking markets. Moreover, the deregulation of the banking industry during the 1990s provides a natural experiment for investigating how firms adjust their executive compensation contracts as the environment in which they operate becomes relatively more competitive (Brewer, Hunter and Jackson, 2004). They show evidence that riskier banks had significantly higher levels of equity-based compensation in the post-deregulation period, as did banks with greater investment opportunities.

Other recent studies examine the impact of CEO compensation structure on bank risk taking. Fahlenbrach and Stulz (2011) use US banking data to uncover the existence of a relationship between bank performance, stock returns and returns on assets with equity incentives. They find that banks led by CEOs whose have better manager-shareholder alignment was associated with significantly worse performance during the crisis. In other words, financial strategies such as stock holding or option compensation may encourage CEOs to engage in more risky projects to increase their future equity value. Some empirical studies find that the sensitivity of CEO equity and option holdings to stock volatility has an impact on bank risk taking (e.g., Chen et al., 2006; Mehran and Rosenberg, 2007; Hagendorff and Vallascas, 2011; Boyallian and Ruiz-Verdú, 2015).

Acharya and Naqvi (2012) develop a theoretical model to show that bank over-lending may result from bank managers' desire to receive higher compensations. Acharya and Naqvi (2016) further show that, if a bank is awash with deposits from investors, the manager will be more likely to undertake high-risk projects to pursue his/her own self-interest and may sanction excessive loans by lowering lending standards, which can lead to asset price bubbles and sow seeds for future bank failure. Both of these papers report that over-lending may result from bank managers' desire to receive higher compensation, which is an agency problem between managers and shareholders.

To the best of our knowledge, however, no empirical research has to date investigated the specific relationship between vega and bank risk taking in loan contracts. Our research contributes to the literature in that CEOs with higher vega tend to take greater risks by setting up mechanisms to decrease the price of loans and relax lending standards on bank loan contracts. Therefore, we propose the following hypotheses:

Hypothesis 1: Banks with higher CEO vega are more likely to charge lower loan spreads.

In addition to the price of a loan, other non-price terms can also be imposed if banks consider control benefits to be a harmful agency problem. Some of these terms are found in Graham, Li, and Qiu (2008) and Rahaman and Zaman (2013) with covenants and collateral requirements. Therefore, our second hypothesis is as follows:

Hypothesis 2: Banks with higher CEO vega tend to impose more favorable non-price terms on loans.

DATA AND METHODOLOGY

Vega Measure

Following Core and Guay (2002), Coles, Daniel and Naveen (2006), we first define vega as the dollar amount change in the CEO stock and option portfolio per one-percent change in the standard deviation of the annualized stock return (stock-return volatility). Then, to generate annual estimates of *Vega*, for fiscal years 1992-2005 we use the "one-year approximation" method, which explains about 99 percent of the variation in option portfolio values and sensitivities that one would obtain from having full information about option portfolios. For fiscal year 2007 and later, however, it is unnecessary to use approximation (OA) in Core and Guay (2002), because all firms on Execucomp disclose information with complete portfolios of exercised and unexercised stock options in the outstanding equity tables using the new-format rules for DEF14A filings. For 2006, 84 percent of firms' report using the new format, while the rest report under the old format.

For calculation of vega under the new reporting format, Execucomp provides option data in the current year and previously granted options, which include the number of options granted, option exercise price, time to maturity, the firm's expected dividend yield, the firm's expected stock return volatility, and the risk-free interest rate for that year. All the above variables are inputs to estimate option values and sensitivities from the formulas in the Black–Scholes option-valuation model (1973), modified by Merton (1973) to account for dividends:

$$\text{Option value} = [\mathrm{S}e^{-\mathrm{dt}}\mathrm{N}(\mathrm{Z}) - \mathrm{X}e^{-\mathrm{rt}}\mathrm{N}(\mathrm{Z} - \sigma \mathrm{T}^{(1/2)})] \qquad (1)$$

where Z is $[\ln(S/X)+T(r-d+\sigma^2/2)]\sigma T^{(1/2)}$; N is the cumulative probability function for normal distribution; S is stock price; X is the option-exercise price; σ is the expected stock-return volatility; r is the natural logarithm of risk-free interest rate; T is option time to maturity, in years; d is the natural logarithm of dividend yield.

The sensitivity of option value with respect to a 0.01 change in stock-return volatility (where N' is the normal density function) is:

$$[\partial(\text{option value})/\ \partial(\text{stock return volatility})] * 0.01 = e^{-\mathrm{dt}}\mathrm{N}'(\mathrm{Z})\mathrm{S}\mathrm{T}^{\left(\frac{1}{2}\right)} \times 0.01 \qquad (2)$$

The vega of all vested and unvested options awards are summed for each executive-year to obtain our vega of option portfolio.

To calculate vega with pre-2006 data, since Execucomp only provides firms' detailed reports for the current year's option grants, we therefore use the one-year approximation method in Core and Guay (2002) to estimate data on previously granted options. We consider three option portfolios: the current year's option grants, the portfolio of unvested options from previously granted awards, and the portfolio of vested options. CEOs' incentives are given by the summation of the incentives from these three portfolios.

For the current year's option grants, we obtain the number of options granted during that year, the stated exercise price, and maturity (based on expiration date), along with the firm's expected dividend yield, the firm's expected stock-return volatility, and the risk-free interest rate for that year, all from Execucomp, are used to estimate option values and sensitivities from formulas described earlier.

To calculate the portfolio of previously granted unvested options, we make assumptions based on exercise price and times to maturity as in Core and Guay (2002). First, we estimate the total number of options in the portfolio and the average exercise price of each option. We net off the total number of options granted in the current year from the number of unvested options to estimate the number of previously granted unvested options. We then calculate the difference between the reported intrinsic value of all unvested options and the intrinsic value of the current year's grants. Finally, we subtract the average intrinsic value of each option in the portfolio from the stock price to obtain the average exercise price of each previously granted unvested option. If a firm grants options in the current year, the time to maturity of previously granted unvested options is set to the actual maturity of current year option grants minus one. If no option grants are made in the current year, we assume that the average maturity of previously granted unvested options is nine years.

For vested options, we calculate the average exercise price based on the realizable value and number of vested options. The maturity of vested options equals the maturity of unvested options minus three.

The vega is the sum of the vega of the current year options plus previously granted options (both vested and unvested).

Data and Variables

We start with all companies with standard industrial classification (SIC) codes between 6000 and 6300 that intersect in the Standard & Poor's Compustat and ExecuComp databases. Following Fahlenbrach and Stulz (2011), we exclude firms not in the lending business nor those in the traditional banking industry, such as online brokerages, payment processors, or investment advisors (SIC 6282). Our sample includes both depository institutions (commercial banks and savings institutions) and investment banks. Given the availability of data from the ExecuComp database, we choose 1991 as the first year of the sample. As we need to calculate the rate of change in our main variables at year *t* based on the data at year *t* – 1, our sample is further restricted to include firms with at least two years of data. Thus, the effective results period begins with 1992. The final sample contains 20,502 bank-year observations from 1992 through 2014.

Loan data are collected from the DealScan database, which consists of information about *loan spread*, *loan maturity*, *size*, and other non-price terms. The main dependent variable used in this research is *Spread*, which is defined as the natural logarithm of all-in spread in DealScan (spread over LIBOR or equivalent). Variables in loan characteristics include *Maturity* (natural log of loan maturity in months), *Loan size* (natural log of amount of loan in US$ million), *Performance* (a dummy variable that equals one if the loan facility uses performance pricing), *Collateral* (a dummy variable that takes a value of one if a loan is secured, and zero otherwise), *GenCov* (number of general covenants), and *FinCov* (number of financial covenants).

The firm characteristics are from Compustat: *Assets* (natural log of the total assets of the firm), *Leverage* (long-term debt plus debt in current liabilities divided by total assets), *Tangibility* (net property, plant, and equipment divided by total assets), *Profitability* (earnings before interest, taxes, depreciation, and amortization (EBITDA) divided by total assets), *MB* ([total assets – book value of equity + price × common shares outstanding]/total assets), *CF-volatility* (standard deviation of quarterly cash flows from operations over the four fiscal years prior to the loan initiation year scaled by the total debt), *Z-score* (modified Altman's Z-score [1.2 × working capital + 1.4 × retained earnings + 3.3 × EBIT + 0.999 × sales]/total assets).

Lender characteristics are from Compustat Bank: *L_Asset* (log of total assets in billions of dollars), *L_ Leverage* (ratio of assets to book value of equity), and *L_loandep* (ratio of average balance of loans to average balance of deposits). Macroeconomic factors are *Credit*

Spread (the difference between the AAA corporate bond yield and BAA corporate bond yield) and *Term Spread* (the difference between the 10-year Treasury yield and the 2-year Treasury yield).

CEO compensation and characteristics include *Delta* (Dollar change in wealth associated with a one-percent change in the firm's stock price in US$ million), *CEO inside debt ratio* (the value of inside debt divided by the total value of shares and options owned). We define a CEO's inside debt as the sum of the balance in the CEO's pension fund and nonqualified deferred compensation), *TDC1* (this variable includes salary, bonus, stock awards, option awards, long-term incentive plans, and other annual compensation such as perquisites and other personal benefits in $million), *CEO age* (the CEO's age when the company signs the bank loan contract), *CEO gender* (The gender of the executive officer), and *CEO tenure* (the number of years the CEO held his/her position in the company before signing the bank loan contract).

Bank governance variables include, *Board size* (Number of board directors), *Independent directors* (the percentage of outside directors), *Female* (the proportion of female directors on the Board), *Duality* (a dummy variable for when the CEO also holds the position of Chairman of the Board) and *Busy board* (boards in which the majority of the outside directors are labeled "busy," i.e., if he/she holds three or more directorships).
The measures for credit risk of borrowers include *Beta high* (a dummy variable that equals one if the beta of borrowers is larger than the sample median and zero otherwise), *Idiorisk high (*a dummy variable that equals one if the idiosyncratic risk of borrowers is larger than the sample median and zero otherwise) and *CF volatility high (*a dummy variable that equals one if the cash flow volatility of borrowers is larger than the sample median and zero otherwise). Appendix summarizes all variable definitions.

Summary Statistics

Table 1: Descriptive Statistics

	Mean	Q25	Median	Q75	Standard deviation
Vega	5.5677	4.4080	5.9953	6.9332	1.5334
Spread	4.9593	4.4716	5.1648	5.5413	0.8248
L_assets	12.3910	11.1540	12.4858	13.9203	1.6026
L_leverage	12.5954	11.1441	12.3611	13.5527	2.6031
L_loandep	0.8357	0.7141	0.8652	0.9571	0.1704
CEO age	4.0422	3.9890	4.0431	4.0943	0.0728
Maturity	3.5975	3.1781	3.8712	4.0943	0.7276
Loan size	4.7611	3.8586	5.0080	5.8579	1.6233
Performance	0.3887	0.0000	0.0000	1.0000	0.4875
Collateral	0.4658	0.0000	0.0000	1.0000	0.4988
Fincov	0.5862	0.0000	0.6931	1.0986	0.6181
Gencov	1.0187	0.0000	1.0986	1.7918	0.8426
Assets	7.0936	5.7812	7.0817	8.3575	1.9935
Tangibility	0.3000	0.0906	0.2288	0.4664	0.2517
Leverage	0.3494	0.1795	0.3282	0.4779	0.2447

Profitability	0.1139	0.0791	0.1184	0.1658	0.1373
MB	0.6991	0.6244	0.7319	0.8233	0.2064
Z_score	1.4804	0.7903	1.5605	2.4105	2.8178
CF volatility	1.8469	0.0332	0.0769	0.1886	38.6948
Credit spread	-0.8962	-1.0200	-0.8300	-0.6900	0.2883
Term spread	-3.0729	-5.8400	-2.6400	-0.5500	2.4514

The mean, median values of bank loan spread in Table 1 are 4.9593 and 5.1648. The average spread is 4.9593 ($e^{4.9593}$ = 142.4940 basis points), with a standard deviation of 0.8248. The value of *Vega* from 1992 to 2014 is positive in all firms under consideration, which is reasonable and consistent with the previous literature on quantifying methods of vega (Coles, Daniel and Naveen, 2006; DeYoung, Peng and Yan, 2013). The mean value of *Vega* is 5.5677 with the standard deviation of 1.5334 respectively.

The average of *L_Asset*, *L_ Leverage*, and *L_loandep* are 12.3910, 12.5954, and 0.8357, respectively. Regarding loan characteristics, the average of *Maturity, Loan size, Performance, Collateral, FinCov, GenCov* are 3.5975, 4.7611, 0.3887, 0.4658, 0.5862 and 1.0187, respectively. In other words, the loan contracts under consideration have a mean maturity of 37 months (exponential of 3.5975). The mean value of loan size is US$ 117 million (4.7611 in natural logarithm). A loan contract in the sample has, on average, 2.7695 general covenants (maximum of 10 covenants) and 1.7971 financial covenants (maximum of 7 covenants) for the borrowers. For the characteristics of borrowing firms, the average *Assets, Tangibility, Leverage, Profitability, MB, Z-score and CF volatility* values are 7.0936, 0.3000, 0.3494, 0.1139, 0.6991, 1.4804 and 1.8469, respectively. The last two rows provide the summary statistics for the macroeconomics variables. The average of *Credit spread* and *Term spread* are -0.8962 and -3.0729, respectively.

In Table 2, we divide the sample into high- and low-vega banks based on the median of *Vega.*

Table 2: Loan, borrower, and lender characteristics for banks with high and low vega

This table presents the mean comparison of loan, borrower, and lender characteristics between firms with high and low vega. *t*-test deployed to compare the mean. *, ** and*** denote significance at 10%, 5% and 1% levels, respectively.

Variables	High vega firms	Low vega firms	Differences	*t*-statistics
Spread	4.9667	5.0451	-0.0784***	-4.58
L_assets	13.6704	12.2265	1.4439***	46.89
L_leverage	12.5259	11.5169	1.0090***	21.15
L_loandep	0.7758	0.8764	-0.1005***	-28.36
Maturity	3.6559	3.5730	0.0829***	5.36
Loan size	5.3441	4.6386	0.7055***	20.81
Performance	0.4429	0.4852	-0.0423***	-3.89
Collateral	0.4565	0.5425	-0.0860***	-7.91
Gencov	1.0998	1.2919	-0.1920***	-10.85
Fincov	0.6479	0.7937	-0.1458***	-11.12
Assets	7.6915	6.7667	0.9247***	24.08
Leverage	0.3405	0.3145	0.0260***	5.42
Tangibility	0.3471	0.3010	0.0461***	8.68
Profitability	0.1315	0.1018	0.0297***	10.73
MB	0.6803	0.7042	-0.0238***	-5.52
Z-Score	1.5654	1.2037	0.3617***	4.93
CF volatility	1.4369	4.5633	-3.1263***	-2.72

CEO age	4.0301	4.0327	-0.0026***	-1.70

We use a *t*-test to determine the significance of the difference in the means of all variables between the two groups and report the results for this preliminary univariate test. Most of the mean comparisons show a significant difference at the one-percent level. Banks with high vega tend to charge a significantly higher *Spread* than those with lower vega. Banks in the high-vega group also ask for less collateral as well as impose a lower number of covenants (both general and financial) compared to those in the lower-vega group.

We also report the Pearson's correlation coefficient matrix in Table 3

Table 3: Correlation matrix

This table presents the Pearson correlation matrix between all the variables. The sample period is from 1992 to 2014. * and ** denote significance level at 5% and 1%, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Vega	1.00										
(2)Spread	-0.16**	1.00									
(3) Maturity	0.09**	0.15**	1.00								
(4) Loansize	0.28**	-0.42**	0.15**	1.00							
(5)Performance	0.02*	-0.08**	0.18**	0.15**	1.00						
(6) Collateral	-0.13**	0.51**	0.18**	-0.31**	0.11**	1.00					
(7) Gencov	-0.08**	0.16**	0.18**	0.03**	0.61**	0.39**	1.00				
(8) Fincov	-0.06**	0.18**	0.16**	-0.08**	0.57**	0.36**	0.77**	1.00			
(9) Assets	0.28**	-0.47**	0.01	0.78**	0.01**	-0.40**	-0.13**	-0.22**	1.00		
(10) Leverage	0.11**	0.19**	0.13**	0.13**	-0.03**	0.13**	0.04**	0.02*	0.08**	1.00	
(11) Tangibility	0.10**	-0.08**	0.03**	0.12**	0.01*	-0.08**	-0.03**	-0.05**	0.08**	0.15**	1.00
(12) Profitability	0.23**	-0.23**	0.13**	0.23**	0.12**	-0.15**	0.02*	0.04**	0.16**	0.02*	0.11**
(13) MB	-0.02*	0.10**	0.10**	-0.05**	0.06**	0.06**	0.07**	0.12**	-0.18**	0.27**	0.07**
(14) Z_score	0.11**	-0.18**	0.05**	0.12**	0.07**	-0.13**	0.00	0.01	0.10**	-0.20**	-0.05**
(15) CF volatility	-0.07**	0.01	0.00	-0.02**	0.00	0.01	0.00	0.00	-0.03**	-0.07**	-0.03**
(16) Credit spread	0.22**	-0.19**	0.07**	-0.07**	-0.02*	-0.01	-0.05**	-0.03**	-0.11**	0.06**	0.03**
(17) Term spread	0.19**	0.20**	0.08**	0.21**	0.00	0.01	0.05**	0.01	0.24**	-0.06**	-0.03**
(18) L_leverage	0.23**	-0.07**	-0.07**	-0.04**	-0.05**	-0.01*	-0.09**	-0.06**	-0.02**	0.11**	0.07**
(19) L_loandep	-0.16**	0.02*	0.01	-0.13**	0.09**	0.02**	0.08**	0.12**	-0.17**	0.04**	0.02*
(20) L_assets	0.47**	0.04**	0.16**	0.38**	0.02**	-0.07**	0.01	-0.02**	0.38**	0.05**	0.03**
(21)CEO age	0.06**	-0.04**	0.10**	0.01	0.01*	0.01	0.01	-0.02*	-0.01	0.04**	0.02*
Variable	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	
(12) Profitability	1.00										
(13) MB	0.11**	1.00									
(14) Z_score	0.47**	0.12**	1.00								
(15) CF volatility	-0.02*	-0.02*	0.00	1.00							
(16) Credit spread	0.03**	0.02**	0.04**	-0.03**	1.00						
(17) Term spread	0.03**	-0.03**	-0.03**	0.01	-0.61**	1.00					
(18) L_leverage	0.04**	-0.01	0.06**	-0.01	0.22**	-0.34**	1.00				

(19) L_loandep	0.15**	0.11**	0.14**	0.01	-0.11**	-0.20**	-0.05**	1.00			
(20) L_assets	0.20**	-0.01	0.13**	0.00	-0.37**	0.60**	-0.06**	-0.03**	1.00		
(21)*CEO age*	0.07**	0.00	0.06**	0.00	0.16**	-0.10**	-0.04**	0.19**	0.00	1.00	

As expected, we find a negative and significant correlation between *Spread* and *Vega*. The correlation coefficients between bank compensation regarding vega and other loan terms such as *collateral* requirements and *general* and *financial covenants* are also negative and significant at the one-percent level. The correlation coefficient between *Spread* and other control variables such as *L_leverage* and *CEO age* are significantly negative, indicating that banks with high leverage in their capital structure and those with older CEOs are also associated with a significantly lower spread in their loans.

Methodology

Following the studies of Coles, Daniel, Naveen (2006), DeYoung, Peng, and Yan (2013) and Hasan, Hoi, Wu, and Zhang (2014, 2017), we use an ordinary least squares (OLS) regression to investigate the effects of CEO risk-taking incentives in terms of *Vega* on bank loan spread:

$$\mathrm{Spread}_{\mathrm{i,t}} = \alpha_0 + \alpha_1 \mathrm{Vega}_{\mathrm{i,t-1}} + \beta' \mathrm{F}_{\mathrm{i,t-1}} + \theta' \mathrm{Z}_{\mathrm{i,t}} + \gamma_{\mathrm{i}} + \mu_{\mathrm{t}} + \varepsilon_{\mathrm{i,t}} \quad (3)$$

where the dependent variable is *Spread* and the main explanatory variable is *Vega*. Specifically, $Spread_{i,t}$ is the natural logarithm of the bank loan spread for loan *i* in year *t*; $Vega_{i,t-1}$ represents CEO risk-taking incentives, captures the change in the dollar value of CEO wealth for a 0.01-unit change in stock return volatility for bank *i* in year *t-1* (Coles, Daniel and Naveen, 2006, 2013). $F_{i,t-1}$ is a vector of control variables for lenders and borrowers *i* in year *t-1*, including lender and borrower characteristics. $Z_{i,t}$ is the vector of the control variables for loan and macroeconomic factor *i* in year *t*. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\boldsymbol{\varepsilon_{i,t}}$ represents the error term of the regression. We also control for loan primary purpose and loan type. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level robust standard errors (White, 1980).

To test the effect of CEO risk-taking incentives on other non-price loan terms, we use various regression methods. The Poisson regression is used to test the influence of *Vega* on the total number of covenants and on general and financial covenants individually, as follows:

$$\mathrm{Ln}\left(\mathrm{E}\left(\mathrm{Y}\middle|\mathrm{Vega}_{\mathrm{i,t-1}}, \mathrm{F}_{\mathrm{i,t-1}}, \mathrm{Z}_{\mathrm{i,t}}\right)\right)_{\mathrm{i,t}} = \alpha_0 + \alpha_1 \mathrm{Vega}_{\mathrm{i,t-1}} + \beta' \mathrm{F}_{\mathrm{i,t-1}} + \theta' \mathrm{Z}_{\mathrm{i,t}} + \gamma_{\mathrm{i}} + \mu_{\mathrm{t}} + \varepsilon_{\mathrm{i,t}} \quad (4)$$

The dependent variables used in this form, $Y_{i,t}$ are different non-price terms of a bank loan contract for firm *i* in year *t*, namely, *Totalcov*: the total number of covenants in a loan contract; *Gencov*: the number of general covenants; and *Fincov*: the number of financial covenants. Since the number of covenants in a loan contract are countable data, the Poisson regression is a reasonable method to analyze the effect of private benefits of control. This methodology is also widely used in the bank-loan literature, such as Graham, Li, and Qiu (2008) and Hasan, Hoi, Wu, and Zhang (2014).

Regarding loan terms, we also deploy a probit regression model using *Collateral* as the dependent variable. *Collateral* is a dummy that equals one if a bank requires collateral on a loan provided to the company, and zero otherwise. The equation of the probit model is as follows:

$$\Pr\left(\mathrm{Collateral}_{\mathrm{i,t}} = 1 \middle| \mathrm{Vega}_{\mathrm{i,t-1}}, \mathrm{Z}_{\mathrm{i,t}}\right) = \Phi(\alpha_0 + \alpha_1 \mathrm{Vega}_{\mathrm{i,t-1}} + \beta' \mathrm{F}_{\mathrm{i,t-1}} + \theta' \mathrm{Z}_{\mathrm{i,t}} + \gamma_{\mathrm{i}} + \mu_{\mathrm{t}}) \quad (5)$$

EMPIRICAL RESULTS

Bank CEO Risk-taking Incentives and Loan Spread

In this section, we examine the relationship between bank CEO risk-taking incentives and bank loan spread. Table 4 shows the regression results based on Equation (3).

Table 4: Bank's CEO risk-taking incentives and Bank loan spread

This table present the ordinary least squares (OLS) regression results of bank CEO risk-taking incentives on bank loan spread.

$$Spread_{i,t} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{it}$$

where the dependent variable $Spread_{i,t}$ is the natural logarithm of bank loan spread for loan *i* in year *t*; $Vega_{i,t-1}$represents the incentives for the CEO of bank i in year *t-1*; $F_{i,t-1}$is a vector of control variables for borrowers and lenders *i* in year *t-1,* including borrower and lender characteristics; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factors *i* in year *t.* γ_i and μ_t represent the fixed effect of industry and year, respectively. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level-robust standard errors (White, 1980). The sample period is 1992-2014. All the variables are defined in Appendix A *, ** and *** denote the significance level of 10%, 5% and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)
Constant	5.0890***	4.7309***	5.8102***	6.2742***	5.6337***
	(23.13)	(20.82)	(26.83)	(19.52)	(11.69)
Vega	-0.0605***	-0.0572***	-0.0199***	-0.0175***	-0.0156**
	(-12.87)	(-13.18)	(-3.59)	(-2.77)	(-2.47)
L_assets			-0.0528***	-0.0009	-0.0023
			(-9.17)	(-0.13)	(-0.33)
L_leverage			-0.0132***	0.0004	0.0023
			(-3.91)	(0.11)	(0.60)
L_loandep			0.1136***	-0.0226	-0.0751*
			(3.11)	(-0.57)	(-1.88)
Assets				-0.1751***	-0.1151***
				(-36.14)	(-16.46)
Leverage				0.8203***	0.8300***
				(23.76)	(24.27)
Tangibility				-0.2373***	-0.2220***
				(-5.77)	(-5.42)
Profitability				-1.0365***	-0.9560***
				(-6.35)	(-5.97)
MB				-0.2143***	-0.2087***
				(-5.34)	(-5.13)
Z_score				0.0106**	0.0100***
				(2.40)	(2.94)
CF volatility				0.0000	0.0001
				(0.36)	(0.81)
CEO age					0.2328***
					(2.63)
Maturity					-0.0377**
					(-2.34)
Loansize					-0.0773***
					(-10.26)
Performance					-0.0799***
					(-5.49)
Fincov					0.0404**
					(2.39)
Gencov					0.0891***

					(7.22)
Credit spread					-0.2163
					(-1.60)
Term spread					0.0751***
					(3.44)
Control for					
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Loan purpose	No	Yes	Yes	Yes	Yes
Loan type	No	Yes	Yes	Yes	Yes
Obs.	19,821	19,821	12,320	11,866	8,362
Adj R^2	0.2472	0.4446	0.6205	0.6420	0.6456

Our focus is the coefficient of Vega on bank loan spread. For the robustness of our results, we conduct five specifications in the regression setting. The first controls for industry and year fixed effects only; the second adds controls for loan purpose and type; the third for lender characteristics; the fourth for borrower characteristics; and the fifth for CEO age, loan characteristics and macroeconomic factors.

Consistent with our expectation, negative and significant coefficients are observed in all specifications, even though we have controlled for all potential factors, indicating that banks with higher vega compensation schemes will charge lower spreads for bank loans. Thus, the results support our hypothesis 1 that banks with higher CEO vega are more likely to charge lower loan spreads.

Specifically, the coefficients of Vega are significant at the one-percent level from a -0.0156 to -0.0605 decrease in log loan spread. For example, in Model (3), the coefficient of Vega is -0.0199, which shows that, on average, banks with higher vega will charge lower loan spreads. Thus, banks with higher vega will charge a 1.5642 lower basis-point ($\mathbf{e}^{\mathbf{0.0199}} \times 1.5334 = 1.5642$) loan spread per each one-standard-deviation increase in vega. As the average loan spread in our sample is 142.4940 basis points ($e^{4.9593}$), a one-standard-deviation increase in Vega lowers a firm's bank loan interest rate an average of 1.1 percent ($1.5642 \div 142.4940 = 1.1\%$). Furthermore, the lower 1.5642 basis-point loan spread is around 36 percent of the estimated coefficients for the effect of corporate tax avoidance and for the effect of social capital in Hasan, Hoi, Wu, and Zhang (2014, 2017). Accordingly, the effect of bank CEO incentives on loan spread is not only statistically significant but is also economically important.

Concerning lender characteristics control variables, L_assets, L_leverage, and L_loandep, the first two is significantly negative in relation to spread and the last is significantly positive in Model (3). However, when we add the control variables of borrower characteristic in Model (4), the coefficient between lender control variables and spread all become insignificant. In Model (5) we include all control variables so that, while L_assets and L_leverage remain insignificant, the coefficient of L_loandep becomes negatively significant at 10 percent. However, our Hypothesis 1 is more strongly supported since significant negative impacts of Vega to spread are evident in all specifications.

Concerning borrower characteristics, loan spread is significantly negatively associated with Asset and Profitability, which is consistent with the notion that firms can reduce their financing costs when they are larger in size and better performing (Graham, Li, and Qiu, 2008). Conversely, firms that use greater debt in their capital structure are charged more when they seek new loans from banks, consistent with the findings of Chava and Roberts (2008) and Hasan, Hoi, Wu, and Zhang (2014). Regarding loan characteristics, the results indicate that loans with shorter maturities, smaller amounts, and without performance pricing have higher loan spreads, while the existence of collateral requirements and general covenants in the previous period seems to increase the spread for current bank loans. Macroeconomic factors, such as term spreads, show positive and significant effects on financing costs. Overall, the effects of the control variables exhibit a similar pattern as reported in the existing literature.

Bank CEO Risk-taking Incentives and Non-Price Terms

We next examine whether CEO vega compensation drives banks to relax their lending standards in non-price terms for borrowers. Following Graham, Li, and Qiu (2008) and Hasan, Hoi, Wu, and Zhang (2014), we use different non-price terms as dependent variables in Table 5. In Models (1) to (3), following Equation (3), we use the Poisson regression with total covenants (TotalCov), general covenants (Gencov), and financial covenants (Fincov) as dependent variables. Model (4) employs a probit model of the probability of a firm being required to provide collateral for a loan (following Equation (5)).
Models (1) to (3) relate to covenants in loan contracts. The vega is negatively correlated with the number of covenants at a one-percent significance level. However, when we consider both general and financial covenants, it seems that Vega significantly affects only the number of general covenants. The coefficient of Vega for financial covenants in Model (3) is negative, but not significant. Finally, in Model (4), the probit model indicates that high-vega banks are less likely to seek collateral for their loans. The results in Table 5 are all consistent with our hypothesis 2 that banks with higher CEO vega tend to impose more favorable non-price terms on loans.

Table 5. Bank CEO risk-taking incentives: Non-price terms

This table presents the results of bank CEO risk-taking incentives on non-price terms. Specifications (1), (2), and (3) follow the equation forms:

$$Ln(E(Y|Vega_{i,t-1}, F_{i,t-1}, Z_{i,t}))_{it} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{it}$$

where the dependent variables, $Y_{i,t}$ are different terms of a bank loan for loan *i* in year *t*. For Poisson regression: *Totalcov* is the total number of covenants of a loan contract; *Gencov* is the number of general covenants, and *Fincov* is the number of financial covenants.
Specification (4) follows a probit regression model of the probability that collateral is required when firm *i* takes a loan in year *t*:

$$Pr(Collateral_{i,t} = 1|Vega_{i,t-1}, F_{i,t-1}, Z_{i,t}) = \emptyset(\alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t)$$

where $Vega_{i,t-1}$ represents the private benefits of control for firm *i* in year *t-1*; $F_{i,t-1}$ is a vector of control variables for borrowers and lenders *i* in year *t-1*; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factor *i* in year *t.* γ_i and μ_t represent the fixed effect of industry and year, respectively. In all specifications, the *t-statistics reported* are based on heteroscedasticity at firm-level-robust standard errors (Hubert-White, 1980). The sample period is 1992-2014. All the variables are defined in Appendix A. *, ** and *** denote the significance level of 10%, 5% and 1%, respectively

	(1)	(2)	(3)	(4)
	Poisson regression			Probit
Dependent variables	*TotalCov*	*GenCov*	*FinCov*	*Collateral*
Constant	0.7305	0.1977	-0.1979	-1.3207
	(1.58)	(0.42)	(-0.35)	(-0.86)
Vega	-0.0141**	-0.0190***	-0.0058	-0.0781***
	(-2.19)	(-3.13)	(-0.66)	(-3.82)

L_assets	-0.0035	0.0018	-0.0138	0.0110
	(-0.51)	(0.27)	(-1.56)	(0.48)
L_leverage	-0.0107**	-0.0121***	-0.0080	-0.0001
	(-2.49)	(-2.83)	(-1.50)	(-0.01)
L_loandep	0.1668***	0.1175***	0.2504***	-0.4018***
	(3.87)	(2.77)	(4.48)	(-2.65)
CEO age	-0.1726*	-0.1864*	-0.1581	0.6707**
	(-1.69)	(-1.82)	(-1.28)	(2.00)
Assets	-0.0666***	-0.0601***	-0.0779***	-0.2424***
	(-8.34)	(-7.53)	(-7.96)	(-9.73)
Leverage	0.2225***	0.2408***	0.1955***	1.5974***
	(6.53)	(6.90)	(4.83)	(11.55)
Tangibility	-0.0210	-0.0528	0.0242	-0.3979**
	(-0.46)	(-1.10)	(0.45)	(-2.56)
Profitability	-0.0862	-0.2054***	0.1111	-2.0760***
	(-1.63)	(-3.97)	(1.25)	(-5.67)
MB	0.0879*	0.0337	0.1834***	-0.4663***
	(1.81)	(0.69)	(3.01)	(-2.86)
Z score	0.0033***	0.0036***	0.0028***	-0.0330
	(3.48)	(3.09)	(2.85)	(-0.93)
CF volatility	-0.0001	-0.0000	-0.0003**	0.0005
	(-1.31)	(-0.44)	(-2.44)	(1.09)
Maturity	-0.0385**	-0.0293*	-0.0563***	0.0722
	(-2.20)	(-1.66)	(-2.62)	(1.30)
Loansize	0.0108	0.0200**	-0.0034	-0.1403***
	(1.23)	(2.28)	(-0.32)	(-5.35)
Performance	0.8240***	0.7753***	0.9093***	-0.1610***
	(46.08)	(45.13)	(40.38)	(-2.97)
Gencov				0.7250***
				(15.54)
Fincov				0.2488***
				(4.52)
Credit spread	-0.5171**	-0.4912**	-0.6276**	0.0215
	(-2.16)	(-2.14)	(-2.02)	(0.04)
Term spread	-0.0128	0.0421*	-0.0857***	-0.0828
	(-0.62)	(1.96)	(-3.38)	(-0.94)
Control for				
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes
Obs	8,362	8,362	8,362	8,066
Pseudo R^2	0.1557	0.1144	0.1314	0.4825

The CEO Risk-taking-incentive Effect: Corporate-Governance Channels

The literature on corporate governance suggests that a good governance mechanism is a fundamental instrument to reduce agency problems. For example, Adams and Mehran (2012) argue for a positive relation between bank performance and board size. In other studies, a higher number of independent directors is associated with lower bond yields and higher credit ratings (Bhojraj and Sengupta, 2003); outside directors have greater incentives to monitor management because of reputation considerations (Fama and Jensen, 1983; Weisbach, 1988); and the probability of financial statement fraud is lower for firms with independent boards (Beasley, 1996). Regarding the gender issue, Adams and Ferreira (2009) suggest that female directors have better attendance records than male directors, and gender-diverse boards allocate greater effort to monitoring. Moreover, Dayton (1984) argues that board duality reduces efficiency monitoring by directors due to the excessive concentration of power in one person's hands. Fich and Shivdasani (2006) show that firms with busy boards tend to be detrimental to corporate governance.

This study proposes that the effect of CEO risk-taking incentives derives from the corporate governance channel. The agency problem that stems from compensation could be the reason why banks charge lower loan spreads or provide more favorable non-price terms to borrowers. Therefore, in banks with diverse corporate governance, the vega incentive may have a different effect on bank loan contracts. Based on prior studies, we suggest that the agency problem of CEO risk-taking incentives could be increased (reduced) when banks have weak (strong) governance practices. Thus, in this section, we further examine this issue by conducting an analysis of sub-samples representing weak and strong governance practices. Following Hoechle, Schmid, Walter, and Yermack (2012), we use several different variables to measure corporate governance structure. We divide the sample into subsamples based on banks' governance measures (*board size, independent director ratio, female directors, CEO duality*, *busy board*) and run the baseline regressions for each subsample according to Equation (3). If CEO risk-taking incentives are viewed as an agency problem, then we would expect the effect of *Vega* to be more (less) significant when firms are associated with weak (strong) corporate governance quality.

Panels A and B of Table 6 present the results of subsamples for firms with weak and strong governance practices.

Table 6: CEO risk-taking incentives and bank loan spread: Evidence from bank governance channels

This table presents OLS regression results for CEO risk-taking incentive on bank loan spread by considering different levels of bank governance. First, we divide the sample into subsamples based on banks governance's measures (Board size, Independent board, Female, Duality, and Busy board). Second, we run the regressions for each subsample. The empirical model is:

$$Spread_{i,t} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{it}$$

where $Spread_{i,t}$ represents the natural logarithm of bank loan spread for loan *i* in year *t*; $\mathrm{Vega}_{i,t-1}$ is the change in the dollar value of CEO wealth for a 1% change in stock return volatility that a bank grants to its CEO in $million; $F_{i,t-1}$ is a vector of control variables for borrowers and lenders *i* in year *t* − 1; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factors *i* in year *t;* γ_i and μ_t represent the fixed effect of industry and year respectively; and $\varepsilon_{i,t}$ is the random error. In parentheses are *t*-statistics based on standard errors adjusted for heteroskedasticity (White, 1980). Superscripts *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. To save space, we do not report the coefficients for industry and year dummies.

Panel A: Weak Bank Governance					
	Board size ≤ Median	Independent board ≤Median	Female ≤ Median	Duality=1	Busy board ≥ Median
	(1)	(2)	(3)	(4)	(5)

Constant	4.1918***	1.2339	5.3083***	3.9252***	5.3788***
	(4.70)	(0.41)	(5.63)	(5.50)	(4.86)
Vega	-0.0370***	-0.0773**	-0.0557***	-0.0176**	-0.0527***
	(-3.23)	(-2.51)	(-3.49)	(-2.06)	(-3.52)
L_assets	0.0059	0.0581**	-0.0035	0.0153*	0.0387***
	(0.45)	(2.05)	(-0.23)	(1.78)	(2.82)
L_leverage	-0.0076	-0.0199	-0.0018	0.0023	-0.0241**
	(-0.85)	(-0.79)	(-0.22)	(0.38)	(-2.42)
L_loandep	0.0962	0.1051	-0.0115	0.0375	0.1556
	(1.45)	(0.62)	(-0.18)	(0.68)	(1.38)
CEO age	0.6608***	0.8413	0.8107***	0.5214***	0.2671
	(3.09)	(1.58)	(3.59)	(3.42)	(1.00)
Assets	-0.1255***	-0.1238***	-0.1486***	-0.1303***	-0.1221***
	(-10.37)	(-7.00)	(-10.50)	(-14.24)	(-10.75)
Leverage	0.7886***	0.6730***	0.7489***	0.8168***	0.8527***
	(13.77)	(6.49)	(10.55)	(18.13)	(14.96)
Tangibility	-0.2181**	-0.0342	-0.3469***	-0.1337**	-0.0737
	(-2.55)	(-0.34)	(-3.60)	(-2.53)	(-1.18)
Profitability	-0.8722***	-1.0346***	-0.6690***	-1.3451***	-1.3670***
	(-8.29)	(-6.42)	(-7.26)	(-10.23)	(-9.73)
MB	-0.1340	-0.2711***	-0.0351	-0.2211***	-0.2739***
	(-1.56)	(-3.04)	(-0.39)	(-3.51)	(-3.78)
Z score	0.0058	-0.0084	0.0115***	0.0135***	0.0141***
	(1.26)	(-0.72)	(7.86)	(7.71)	(7.03)
CF volatility	0.0000	-0.0006**	0.0005	-0.0000	0.0000
	(0.17)	(-2.26)	(1.48)	(-0.14)	(0.09)
Maturity	-0.0416	-0.0134	-0.0638**	-0.0284	-0.0339
	(-1.42)	(-0.35)	(-2.13)	(-1.34)	(-1.33)
Loansize	-0.0753***	-0.0422**	-0.0650***	-0.0582***	-0.0549***
	(-5.95)	(-2.26)	(-4.43)	(-6.06)	(-4.63)
Performance	-0.0646**	-0.0546	-0.0889***	-0.0977***	-0.0983***
	(-2.37)	(-1.35)	(-2.93)	(-5.00)	(-3.95)
Gencov	0.0951***	0.1204***	0.1540***	0.1170***	0.1055***
	(4.04)	(3.68)	(5.85)	(7.33)	(5.53)
Fincov	0.0348	-0.0028	0.0306	0.0318	0.0040
	(1.20)	(-0.07)	(0.94)	(1.53)	(0.16)
Credit spread	-0.0952	-2.0094**	0.6690***	-0.1787**	-0.2118**
	(-0.98)	(-2.23)	(3.08)	(-1.98)	(-2.04)
Term spread	0.1087***	-0.0254	0.1894***	0.0898***	0.0729***
	(7.47)	(-0.45)	(9.34)	(8.28)	(5.17)
Control For					
Loan purpose	Yes	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	2,383	1,527	2,154	5,015	3,368
Adj R^2	0.6756	0.6288	0.6852	0.6354	0.6224
Panel B: Strong Bank Governance					
	Board size >Median	Independent board >Median	Female > Median	Duality=0	Busy board < Median

	(6)	(7)	(8)	(9)	(10)
Constant	7.5536***	5.0646***	4.4796***	8.9917*	4.1242***
	(3.18)	(6.75)	(4.02)	(1.77)	(3.86)
Vega	-0.0176	-0.0117	-0.0174	-0.0452	-0.0165
	(-1.33)	(-0.93)	(-1.58)	(-0.62)	(-1.19)
L_assets	-0.0123	-0.0213	-0.0107	-0.1733***	-0.0274*
	(-0.82)	(-1.43)	(-0.80)	(-2.84)	(-1.78)
L_leverage	0.0109	-0.0017	-0.0086	-0.0684	-0.0006
	(1.17)	(-0.19)	(-0.88)	(-1.47)	(-0.07)
L_loandep	-0.2617**	-0.0236	0.1945	0.8880	-0.0338
	(-2.10)	(-0.20)	(1.47)	(1.60)	(-0.47)
CEO age	0.0592	0.3137*	0.3373	0.5828	0.7612***
	(0.20)	(1.80)	(1.32)	(0.40)	(3.17)
Assets	-0.1207***	-0.1295***	-0.1156***	-0.0540**	-0.1239***
	(-9.27)	(-11.64)	(-9.47)	(-2.01)	(-8.36)
Leverage	0.7921***	0.9447***	0.8281***	0.0990	0.6976***
	(12.28)	(18.15)	(13.48)	(0.63)	(8.75)
Tangibility	-0.1224*	-0.1582**	-0.0232	-0.1489	-0.2886***
	(-1.72)	(-2.39)	(-0.36)	(-0.50)	(-2.84)
Profitability	-1.2655***	-1.2827***	-1.4724***	-0.2077	-0.7483***
	(-8.02)	(-9.56)	(-7.25)	(-1.48)	(-5.56)
MB	-0.2250***	-0.2517***	-0.2392***	-0.1196	-0.1068
	(-2.70)	(-2.91)	(-2.92)	(-0.73)	(-0.99)
Z score	0.0136***	0.0163***	-0.0190	-0.0028	0.0028
	(10.25)	(11.57)	(-1.59)	(-0.73)	(0.32)
CF volatility	-0.0000	-0.0000	-0.0001*	-0.0007	-0.0002
	(-0.40)	(-0.23)	(-1.90)	(-0.72)	(-0.67)
Maturity	-0.0293	-0.0317	0.0034	-0.0630	-0.0186
	(-1.09)	(-1.17)	(0.13)	(-1.23)	(-0.58)
Loansize	-0.0643***	-0.0656***	-0.0588***	-0.0568*	-0.0961***
	(-4.84)	(-5.60)	(-4.87)	(-1.83)	(-6.27)
Performance	-0.0702**	-0.0972***	-0.0707***	0.2327***	-0.1108***
	(-2.46)	(-3.96)	(-2.74)	(2.72)	(-3.25)
Gencov	0.1206***	0.1268***	0.0846***	-0.1077	0.1488***
	(5.44)	(6.14)	(4.20)	(-1.47)	(5.33)
Fincov	-0.0061	0.0108	0.0005	-0.0984*	0.0429
	(-0.21)	(0.40)	(0.02)	(-1.75)	(1.17)
Credit spread	0.8564	-0.1310	-0.1809*	0.4947	0.0947
	(0.55)	(-1.53)	(-1.83)	(0.70)	(0.43)
Term spread	0.1938***	0.0292	0.1269***	0.1271*	0.1511***
	(2.70)	(0.92)	(4.80)	(1.68)	(6.40)
Control For					
Loan purpose	Yes	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	2,911	2,839	3,140	279	1,926
Adj R^2	0.6384	0.6797	0.6362	0.6776	0.6787

Weak corporate governance proxies in Panel A are defined as follows: Banks with (1) board size higher than the median; (2) a ratio of independent directors greater than median; (3) more female directors than the median; (4) CEO duality equal to zero; (5) busier boards than the median. Across all models, in banks with weak corporate governance, the coefficients of *Vega* are significantly negatively related to loan spread. For example, in Model (1), the coefficient is –0.0370, indicating that banks with higher vega compensation tend to charge lower loan spreads than banks with low board size. However, in banks with strong corporate governance (Panel B), the coefficients of *Vega* are all insignificant. These results add further support to our hypothesis 1, i.e., banks with greater vega tend to charge a lower bank loan spread, especially when these banks have weak corporate governance. More importantly, this result implies that high CEO risk-taking incentives create an agency problem between bank manager and shareholders.

Bank CEO risk-taking Incentives and Riskiness of Borrowers

In this section, we further examine the relationship between CEO risk-taking incentives and bank loan spread by considering borrower credit risk. Riskier borrowers (such as small borrowers, borrowers with less cash, and borrowers that are harder for outside investors to value) must pay more for their loans and rely more on bank financing (Machauer and Weber, 1998; Strahan, 1999). As a normal fact of business, high risk may result in high profit, since the higher loan-rate premium encourages a bank to take the deal. If our hypothesis is correct, we should find that our results mainly come from high-risk borrowers as a result of bank over-lending behaviors. We thus conduct a subsample analysis based on Equation (3). We use three variables to measure firm riskiness: First, *Beta* is the equity beta from regressing daily excess returns on market excess returns over the previous two years, and is a well-known regression measure of systematic risk (Blume, 1971), especially when applying robust methods in the statistical literature (Chan and Lakonishok, 1992). Second, *Idiosyncratic risk* (*Idiorisk*) is the standard deviation of residuals obtained from regressing daily stock excess returns on market excess returns over the previous two years. Measured from a market model regression, high idiosyncratic uncertainty will negatively affect firm-level investment (Gilchrist, Sim and Zakrajsek, 2014). Another study shows that idiosyncratic risk is important in explaining changes in firm financial soundness (Atkeson, Eisfeldt and Weill, 2014). Finally, *Cash-flow volatility* (*CFvol*) is the volatility of daily returns over the previous two years. High cash-flow volatility results in a pattern of underinvestment or undervaluing of the firm by outsiders (Scordis, Barrese and Wang, 2008; Minton, Schrand and Walther, 2002). Higher volatility is related to lower investment (Minton and Schrand, 1999). Thus, lack of investment may encourage high cash-flow volatility in a firm that uses bank loans as a capital-mobilization channel. In this study, firms are defined as high- (low-) risk borrowers if *Beta*, *Idiorisk* and *CFvol* are above or equal to (below) the sample median.

Table 7 presents the regression results for the subsample analysis. The relation between *Vega* on loan spreads is stronger for borrowers with high risk. The coefficient of *Vega* for firms with high idiosyncratic volatility is –0.0272, which is significant at the one-percent level. In contrast, the coefficient of *Vega* for firms with low idiosyncratic volatility is –0.0072, which is not statistically significant. Similarly, the coefficients of *Vega* for firms with high beta and cash-flow volatility are significantly negative at -0.0144 and -0.0160, whereas the coefficients for firms with low beta and low cash-flow volatility are -0.0129 and -0.0153, which are not significant.

Table 7: CEO risk-taking incentives and bank loan spread: Credit risk of borrowers

This table presents OLS regression results for CEO risk-taking incentives on bank loan spread by considering credit risk of borrowers. First, we divide the sample into subsamples based on borrower risk (beta, idiosyncratic risk and cash-flow volatility). Second, we run the regressions for each subsample with bank loan spread as the dependent variable. The empirical model is:

$$Spread_{i,t} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{it}$$

where $Spread_{i,t}$ represents the natural logarithm bank loan spread for loan *i* in year *t*; $Vega_{i,t-1}$ is the change in the dollar value of CEO wealth for a 1% change in stock return volatility which a bank grants to its CEO in $million; $F_{i,t-1}$ is a vector of control variables for borrowers and lenders *i* in year *t* − 1; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factors *i* in year *t;* γ_i and μ_t represent the fixed effect of industry and year respectively; and $\varepsilon_{i,t}$ is the random error. In parentheses are *t*-statistics based on standard errors adjusted for heteroskedasticity (White, 1980). Superscripts *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. To save space, we do not report the coefficients for industry and year dummies.

	High risk borrowers			Low risk borrowers		
	Beta ≥Median	*Idiorisk ≥Median*	*CFvol ≥Median*	*Beta <Median*	*Idiorisk <Median*	*CFvol <Median*
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	5.2395***	5.9038***	6.1372***	5.3158***	6.3659***	5.3481***
	(7.17)	(10.91)	(10.23)	(8.61)	(8.37)	(8.50)
Vega	-0.0144*	-0.0272***	-0.0160*	-0.0129	-0.0072	-0.0153
	(-1.64)	(-3.21)	(-1.86)	(-1.06)	(-0.69)	(-1.53)
L_assets	-0.0042	-0.0111	-0.0105	-0.0069	0.0131	0.0245**
	(-0.39)	(-1.28)	(-1.12)	(-0.61)	(1.01)	(2.20)
L_leverage	0.0027	0.0031	-0.0008	0.0028	-0.0015	0.0034
	(0.39)	(0.65)	(-0.15)	(0.49)	(-0.19)	(0.56)
L_loandep	-0.1573**	-0.1989***	-0.0415	0.0075	-0.0446	-0.0335
	(-2.26)	(-3.80)	(-0.81)	(0.13)	(-0.61)	(-0.49)
CEO age	0.3221**	0.3229***	0.2024*	0.3492***	0.0986	0.2474*
	(2.04)	(2.82)	(1.69)	(2.84)	(0.67)	(1.69)
Assets	-0.1280***	-0.0717***	-0.1353***	-0.1389***	-0.1704***	-0.1034***
	(-11.63)	(-7.26)	(-13.44)	(-11.60)	(-14.51)	(-10.16)
Leverage	0.8322***	0.7249***	0.8276***	1.0389***	1.0673***	0.7742***
	(13.74)	(13.34)	(15.11)	(13.75)	(13.69)	(14.38)
Tangibility	-0.1538**	-0.1134**	-0.1672**	-0.3165***	-0.2261***	-0.2056***
	(-2.46)	(-2.06)	(-2.46)	(-4.48)	(-3.02)	(-3.69)
Profitability	-0.6302***	-0.5362***	-0.7156***	-1.3152***	-2.4424***	-2.0178***
	(-2.85)	(-4.13)	(-4.60)	(-9.00)	(-11.31)	(-11.80)
MB	-0.2751***	-0.2586***	-0.1605**	-0.2593***	-0.2593***	-0.2314***
	(-3.41)	(-4.08)	(-2.35)	(-3.14)	(-2.62)	(-4.26)
Z score	0.0022	0.0087***	0.0119***	0.0159***	-0.0614***	-0.0188
	(0.46)	(3.68)	(5.69)	(6.55)	(-3.01)	(-1.57)
CF volatility	0.0001	-0.0001	0.0000	0.0002	-0.0001	0.0003**
	(1.01)	(-0.47)	(0.61)	(0.54)	(-1.02)	(2.14)
Maturity	-0.0397	-0.0599***	-0.0302	-0.0172	0.0201	-0.0364*
	(-1.55)	(-3.33)	(-1.34)	(-0.72)	(0.59)	(-1.65)
Loansize	-0.0628***	-0.0619***	-0.0815***	-0.0869***	-0.0622***	-0.0595***
	(-5.72)	(-6.18)	(-7.63)	(-6.85)	(-4.94)	(-5.63)
Performance	-0.0685***	-0.1223***	-0.0920***	-0.0613***	-0.0389*	-0.0702***
	(-3.15)	(-5.85)	(-4.15)	(-2.60)	(-1.67)	(-3.65)
Gencov	0.0707***	0.1115***	0.1200***	0.1106***	0.0556***	0.0567***
	(3.67)	(6.99)	(6.31)	(5.25)	(2.61)	(3.40)

Fincov	0.0401	-0.0025	0.0110	0.0352	0.0896***	0.0433*
	(1.57)	(-0.12)	(0.46)	(1.23)	(3.08)	(1.83)
Credit spread	-0.1087	0.0828	-0.1117	-0.8739**	-0.5219***	-0.0809
	(-0.71)	(0.27)	(-0.51)	(-2.05)	(-2.99)	(-0.46)
Term spread	0.0532*	0.0644***	0.0769***	0.0844**	0.1074***	-0.0850***
	(1.93)	(2.89)	(3.39)	(2.49)	(7.82)	(-3.41)
Control For						
Loan purpose	Yes	Yes	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	3,556	3,545	4,291	3,248	3,259	3,825
Pseudo R^2	0.6366	0.5435	0.6743	0.7127	0.6979	0.6414

ROBUSTNESS CHECKS

Control for CEO Characteristics and Other Compensation Schemes

CEO characteristics and other forms of CEO compensation may also affect considerations of loan spread when banks negotiate contracts with creditors (Donelson, Jennings and McInnis, 2014; Kabir, Li, and Veld-Merkoulova, 2013). In this study, we claim that CEO vega represents a form of executive compensation, although not a compulsory regular form such as salary, bonus or stock options. However, the economic intuition of vega, in terms of either agency cost proxy or the complementary incentive-alignment effect, may be more meaningful in comparison with other forms of compensation. Therefore, to confirm that our results are not biased by other CEO characteristics and compensation, such as *CEO tenure*, *CEO gender*, *TDC1*, *Delta*, and *CEO inside debt ratio,* we additionally include the latter in our regressions.

Table 8 presents the regression results of bank CEO risk-taking incentives on bank loan spread by considering CEO characteristics and other CEO compensation. Across all specifications, the coefficients of *Vega* are significantly negatively related to loan spread, even when we control for CEO characteristics and other compensation. For example, in Model (4), the coefficient of *Vega* is -0.0258, which is statistically significant with t-statistics of -2.94. Thus, the evidence still supports our hypothesis 1.

Table 8: Robustness check (I): Control for CEO characteristics and other CEO compensation
This table presents ordinary least squares (OLS) regression results for the influences of CEO vega incentives on bank loan spread by considering CEO characteristics and other CEO compensation as control variables. The empirical model is:

$$Spread_{i,t} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{it}$$

where $Spread_{i,t}$ represents the natural logarithm of bank loan spread for loan *i* in year *t*; $Vega_{i,t-1}$is the change in the dollar value of CEO wealth for 1% change in stock-return volatility that a bank grants to its CEO in \$million; $F_{i,t-1}$ is a vector of control variables for borrowers and lenders *i* in year $t-1$; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factors *i* in year *t;* γ_i and μ_t represent the fixed effect of industry and year respectively; and $\varepsilon_{i,t}$ is the random error. In parentheses are *t*-statistics based on standard errors adjusted for heteroskedasticity (White, 1980). Superscripts *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. To save space, we do not report the coefficients for industry and year dummies.

Bank CEO Risk-taking Incentives

	(1)	(2)	(3)	(4)	(5)
Constant	4.9349***	5.4908***	5.6616***	5.5959***	3.8564***
	(6.65)	(11.32)	(11.62)	(11.28)	(4.93)
Vega	-0.0259**	-0.0162**	-0.0144**	-0.0258***	-0.0159**
	(-2.49)	(-2.56)	(-1.99)	(-2.94)	(-1.96)
CEO tenure	0.0387				
	(1.25)				
CEO gender		0.1516**			
		(2.56)			
TDC1			-0.0040		
			(-0.34)		
Delta				0.0234**	
				(2.19)	
CEO inside debt ratio					-0.0110
					(-1.53)
L_assets	0.0082	-0.0040	-0.0010	-0.0081	-0.0174
	(0.91)	(-0.57)	(-0.13)	(-1.12)	(-1.41)
L_leverage	0.0016	0.0024	0.0021	0.0026	0.0161
	(0.33)	(0.64)	(0.55)	(0.68)	(1.23)
L_loandep	0.0156	-0.0766*	-0.0777*	-0.0443	0.1263
	(0.32)	(-1.92)	(-1.91)	(-1.04)	(1.17)
CEO age	0.3135*	0.2392***	0.2315***	0.2285**	0.3907**
	(1.95)	(2.70)	(2.61)	(2.48)	(2.49)
Assets	-0.1199***	-0.1151***	-0.1150***	-0.1149***	-0.1004***
	(-13.82)	(-16.46)	(-16.45)	(-16.26)	(-9.64)
Leverage	0.8191***	0.8314***	0.8302***	0.8359***	0.8336***
	(20.34)	(24.30)	(24.28)	(24.14)	(15.02)
Tangibility	-0.1749***	-0.2198***	-0.2218***	-0.2293***	-0.0660
	(-3.38)	(-5.37)	(-5.41)	(-5.55)	(-1.09)
Profitability	-1.3110***	-0.9533***	-0.9555***	-0.9976***	-1.4729***
	(-12.21)	(-5.97)	(-5.96)	(-5.76)	(-7.74)
MB	-0.2282***	-0.2078***	-0.2085***	-0.2165***	-0.2182***
	(-4.65)	(-5.11)	(-5.12)	(-5.26)	(-2.76)
Z score	0.0121***	0.0101***	0.0100***	0.0092**	-0.0255
	(4.22)	(3.01)	(2.94)	(2.22)	(-1.51)
CF volatility	0.0000	0.0001	0.0001	0.0001	-0.0001
	(0.53)	(0.80)	(0.81)	(0.76)	(-0.97)
Maturity	-0.0424**	-0.0372**	-0.0377**	-0.0379**	0.0476*
	(-2.19)	(-2.30)	(-2.33)	(-2.31)	(1.83)
Loansize	-0.0767***	-0.0770***	-0.0774***	-0.0779***	-0.0540***
	(-8.28)	(-10.22)	(-10.26)	(-10.21)	(-5.18)
Performance	-0.1007***	-0.0798***	-0.0800***	-0.0816***	-0.0374*
	(-5.35)	(-5.48)	(-5.49)	(-5.57)	(-1.94)
Gencov	0.1124***	0.0890***	0.0892***	0.0886***	0.0277

	(7.26)	(7.22)	(7.23)	(7.18)	(1.57)
Fincov	0.0348*	0.0404**	0.0403**	0.0415**	0.0432*
	(1.71)	(2.39)	(2.38)	(2.44)	(1.77)
Credit spread	-0.0958	-0.2085	-0.2061	-0.2302*	-0.7948***
	(-0.41)	(-1.54)	(-1.49)	(-1.70)	(-5.48)
Term spread	-0.0184	0.0780***	0.0749***	0.0825***	0.3867***
	(-1.33)	(3.55)	(3.43)	(3.71)	(16.39)
Control for					
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Loan purpose	Yes	Yes	Yes	Yes	Yes
Loan type	Yes	Yes	Yes	Yes	Yes
Obs.	5,733	8,362	8,362	8,247	3,179
Adj R^2	0.6561	0.6458	0.6456	0.6473	0.6282

Bank Size Analysis

Bank size can be also considered an important factor in the determination of bank lending decisions (Berger and Bouwman, 2009 and 2013). To some extent, bigger banks are associated with strong corporate governance, more market power, good reputation and trust, since they can easily mobilize deposits (even at lower rates) and attract higher loan demands (even at higher rates) compared to small banks. However, if downside risks were revealed, shareholders may suffer greater losses and, thus, risk-taking incentives should be less preferred in more profitable firms (Keeley, 1990). Demsetz, Mark and Philip (1996) suggest that high franchise value banks operate more safely than those less highly ranked. Meanwhile, John and Qian (2003) show that pay-performance sensitivity declines with bank size, which supports the idea that managers of larger banks will choose to maintain current performance levels rather than opt for a risky project. In other words, according to traditional theory, bigger or more profitable banks may not willing to take risks in order to avoid unexpected effects to their business. Therefore, we expect that our results will be stronger in the small-bank subsample. We therefore run a baseline regression according to Equation (3) that includes the variable *Small bank* and its interaction term *Vega*. *Small bank* is a dummy variable where bank size is lower than the sample median.

Model (1) of Table 9 displays the results of the bank size test. The interaction term of *Vega* and *Small bank* is negative and significant, showing that the effect of CEO risk-taking incentives on loan spreads is stronger for small banks. For example, the effect of *Vega* in the small-bank group is –0.0436 (–0.0337 – 0.0099 = –0.0436), whereas the effect in the large-bank group is –0.0099.

Table 9: Robustness check (II): Bank size analysis and the change regression

This table presents the results of the effect of bank CEO vega incentives on different bank sizes and presents the change regression results of the effect of bank CEO risk-taking incentives (*Vega*). Model (1) follows the below equation form:

$$Spread_{i,t} = \alpha_0 + \alpha_1 Vega_{i,t-1} + \alpha_2 Vega_{i,t-1} \times Small\ bank_{i,t-1} + \alpha_3 Small\ bank_{i,t-1} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t}$$

where $Spread_{i,t}$ represents the natural logarithm of bank loan spread for loan *i* in year *t*; $Vega_{i,t-1}$is the change in the dollar value of CEO wealth for a 1% change in stock-return volatility that a bank grants to its CEO in $million; $\mathrm{Small\ bank}_{i,t-1}$ is dummy variable that equals

1 if the bank size is smaller than the sample median and 0 otherwise; $F_{i,t-1}$ is a vector of control variables for borrowers and lenders *i* in year *t* − 1; $Z_{i,t}$ is a vector of control variables for loans and macroeconomic factors *i* in year *t*. γ_i and μ_t represent the fixed effect of industry and year respectively; and $\varepsilon_{i,t}$ is the random error.

Model (2) follows the below equation form:

$$\Delta Spread_{i,t} = \alpha_0 + \alpha_1 \Delta Vega_{i,t-1} + \beta' \Delta Firm_{i,t-1} + \theta' Z_{i,t} + \varepsilon_{i,t}$$

where the dependent variable $\Delta Spread_{i,t}$ is the change in loan spread for firm *i* in year *t* from year *t-1*. $\Delta Vega_{i,t-1}$ represents the change in bank CEO risk-taking incentives, *Vega i* in year *t*-1 from year *t*-2. $\Delta Firm_{i,t-1}$ represents the change in firm characteristics for lenders and borrowers *i* in year *t*-1 from year *t*-2. $Z_{i,t}$ is the vector of the control variables for loan and macroeconomic factors *i* in year *t*.

In all specifications, the *t*-statistics reported are based on heteroscedasticity (White, 1980). The sample period is 1992-2014. All the variables are defined in Appendix A. *, ** and *** denote significance levels of 10%, 5% and 1%, respectively.

	(1)	(2)
Constant	5.1862***	1.9273**
	(9.99)	(2.32)
Vega	-0.0099	
	(-1.38)	
Vega×Small bank	-0.0337**	
	(-2.39)	
Small bank	0.2581**	
	(2.52)	
ΔVega		-0.0525**
		(-2.18)
Assets	-0.1151***	
	(-16.47)	
Leverage	0.8313***	
	(24.26)	
Tangibility	-0.2197***	
	(-5.35)	
Profitability	-0.9530***	
	(-5.89)	
MB	-0.2101***	
	(-5.15)	
Z score	0.0100***	
	(2.96)	
CF volatility	0.0001	
	(0.86)	
CEO age	0.2494***	0.0031
	(2.78)	(0.02)
Maturity	-0.0379**	-0.0323
	(-2.35)	(-0.72)
Loansize	-0.0770***	0.0033
	(-10.23)	(0.27)
Performance	-0.0798***	-0.0331
	(-5.48)	(-0.83)

Bank CEO Risk-taking Incentives

Gencov	0.0888***	-0.0064
	(7.20)	(-0.19)
Fincov	0.0413**	0.0322
	(2.44)	(0.67)
Credit spread	-0.2294*	0.6825**
	(-1.69)	(2.32)
Term spread	0.0753***	-0.0243
	(3.47)	(-0.59)
L_assets	0.0213*	
	(1.77)	
L_leverage	0.0016	
	(0.39)	
L_loandep	-0.0839**	
	(-2.08)	
ΔAssets		-0.0223
		(-0.36)
ΔLeverage		0.6743***
		(3.95)
ΔTangibility		0.0787
		(0.33)
ΔProfitability		-0.3450
		(-1.44)
ΔMB		-0.0776
		(-0.67)
ΔZ-score		0.0010
		(0.06)
ΔCF-volatility		-0.0006***
		(-3.67)
Δ L_assets		-0.0218
		(-0.12)
ΔL_leverage		-0.0288**
		(-2.43)
ΔL_loandep		-0.1611
		(-0.67)
Control for		
Industry FE	Yes	Yes
Year FE	Yes	Yes
Loan purpose	Yes	Yes
Loan type	Yes	Yes
Obs.	8,362	1,556
Adj R^2	0.6458	0.1966

Change Regression

Following the study of Chava, Livdan, and Purnanandam (2009), we utilize a change-regression analysis to determine whether the loan spread for a given bank is adjusted after a significant increase in their vega. If our agency-cost view is correct, we should observe that banks charge lower spreads when their vega incentives increase. The regression model is as follows:

$$\Delta Spread_{i,t} = \alpha_0 + \alpha_1 \Delta Vega_{i,t-1} + \beta' \Delta F_{i,t-1} + \theta' Z_{i,t-1} + \varepsilon_{i,t} \qquad (6)$$

We compute the change in the main dependent, explanatory, and control variables. In the above formula, $\Delta Spread_{i,t,}$ is the change in loan spread in year *t* from year *t-1*. The main explanatory variable, $\Delta Vega_{i,t-1}$, is the change in vega from year *t-2* to year *t-1*. $\Delta F_{i,t-1}$ is a vector of change in lender and borrower characteristics for firm *i* in year *t-1* from year *t-2*. $Z_{i,t-1}$ is a vector of control variables for firm *i* in year *t-1* from year *t-2*.

Model (2) of Table 9 above shows the results of the change regressions. The change in vega ($\Delta Vega$) significantly and negatively affects the change in loan spread. Thus, the results support our agency-cost view that banks charge lower spreads when their CEO risk-taking incentives increase, which is consistent with our main expectation and shows that our results are not driven by any omitted-variable bias.

CONCLUSION

Banks commonly use equity-based compensation schemes to motivate CEO efforts to take on risky projects, then ameliorate the principal-agent conflict. However, bank executives have faced widespread criticism for driving excessive risk taking that ignited the disastrous 2008 financial crisis. Acharya and Naqvi (2012) develop a theoretical model to show that bank over-lending may result from managers' desire to receive higher compensation. Furthermore, Acharya and Naqvi (2016) show that managers are more likely to undertake high-risk projects to pursue their own self-interest and to sanction excessive loans by lowering lending rates and loosening lending standards, all of which leads to asset price bubbles and sows seeds of future bank failure. Based on their model implications, this study provides the first empirical evidence of how CEO risk-taking incentives (vega) influence bank lending decisions.

We find that vega is negatively and significantly related to loan spread and less stringent non-price terms such as fewer financial and general covenants and a lower probability of collateral requirements. Our results are robust even we control for CEO characteristics and other compensation schemes. Thus, we find that banks with high CEO risk-taking incentives may reduce lending interest rates and relax their lending standards for their own wealth concerns. This effect becomes weaker when banks have strong corporate governance mechanisms, supporting the notion that high vega may create an agency problem between a bank manager and shareholders. Thus, our results provide a reference for scholars, policy makers, and market investors for assessing the significant impact of vega in bank lending decisions.

APPENDIX

Variables	Definition	Data source
Panel A: CEO risk-taking incentive variable		
Vega	Nature log of the dollar amount change of CEO stock and option portfolio per 1 percent change in standard deviation of the annualized stock return (US$ million)	Execucomp
Panel B: Loan characteristics		
Spread	Natural logarithm of loan spread.	DealScan
Maturity	Natural logarithm of loan maturity in months.	DealScan
Loan size	Natural logarithm of loan amount US$ millions.	DealScan
Performance	Dummy variable, equal to one if a loan facility uses performance pricing, and zero otherwise.	DealScan

Collateral	Dummy variable, equal to one if a loan is secured, and zero otherwise.	DealScan
Gencov	Number of general covenants.	DealScan
Fincov	Number of financial covenants.	DealScan
Loan type	Dummy variable for loan types, including term loan, revolver greater than one year, revolver less than 1 year, and 364-day facility.	DealScan
Loan purpose	Dummy variable for loan purposes, including corporate purposes, debt repayment, working capital, takeover, etc.	DealScan
Panel C: Firm characteristics		
Assets	Natural logarithm of total assets in US$ millions.	Compustat
Leverage	Long-term debt and debt in current liabilities divided by total assets.	Compustat
Tangibility	Net property, plant, and equipment divided by total assets.	Compustat
Profitability	Earnings before interest, taxes, depreciation, and amortization (EBITDA), divided by total assets.	Compustat
MB	Market value of net assets to book value of net assets ratio.	Compustat and CRSP
Z-score	Modified Altman's Z-score (1.2 × working capital + 1.4 × retained earnings + 3.3×EBIT + 0.999×sales)/Total assets.	Compustat
CF-volatility	Standard deviation of quarterly cash-flows from operations over the four fiscal years prior to the loan initiation year, scaled by total assets.	Compustat
Panel D: Lender characteristics		
L_Asset	Log of total assets (billions of us dollars).	Compustat Bank (CB)
L_ Leverage	Ratio of assets to book value of equity.	Compustat Bank (CB)
L_loandep	Ratio of average balance of loans to average balance of deposits.	Compustat Bank (CB)
Panel E: Macroeconomic factors		
Credit spread	Difference between AAA corporate bond yield and BAA corporate bond yield.	Datastream
Term spread	Difference between 10-year and 2-year Treasury yields.	Datastream
Panel F: CEO compensation and characteristics		
CEO age	Natural logarithm of Age of CEO. Age of the CEO when the company signs the bank loan contract.	Execucomp
CEO gender	The gender of the executive officer	Execucomp
CEO tenure	Natural logarithm of CEO tenure. Number of years the CEO held his/her position in a company before he/she signed the bank loan contract.	Execucomp
TDC1	We follow the definition of TDC1 in ExecuComp to calculate this variable. TDC1 includes salary, bonus, stock awards, option awards, long-term incentive plans, and other annual compensation such as perquisites and other personal benefits (including termination or change-in-control payments, 401K plans, etc). ($millions)	Execucomp

CEO inside debt ratio	Natural logarithm of CEO D/E (the value of inside debt divided by the total value of shares and options owned. We define the inside debt of the CEO as the sum of the balance in the CEO's pension fund and nonqualified deferred compensation	Execucomp and SEC Edgar DEF 14A Filings
Delta	Natural logarithm of dollar change in wealth associated with a 1% change in the firm's stock price (in US$ million)	Execucomp
Panel G: Bank governance mechanisms		
Busy board	Boards in which the majority of the outside directors are labeled "busy." A director is labeled "busy" if he/she holds three or more directorships	Risk Metrics
Board size	Number of board directors	Risk Metrics
Independent Directors	The percentage of outside directors	Risk Metrics
Duality	A dummy variable for when the CEO also holds the position of Chairman of the Board.	Risk Metrics
Female	The proportion of female directors in the board	Risk Metrics
Panel H: Credit risk of borrowers		
Beta	Firm's equity beta from a market model of daily returns in excess of three-month T-bills using previous two-year data, where the market is represented by the value-weighted CRSP index	CRSP
Idiosyncratic volatility (Idiorisk)	The standard deviation of the residuals obtained from a market model of daily returns in excess of three-month T-bills using previous two-year data, where the market is represented by the value-weighted CRSP index	CRSP
CF volatility	Firm's cash flow volatility using the previous two-year daily data	CRSP

REFERENCES

Acharya, V., Naqvi, H., 2012. The seeds of a crisis: A theory of bank liquidity and risk taking over the business cycle. Journal of Financial Economics. 106, 349-366.

Acharya, V., Naqvi, H., 2016. On reaching for yield and the coexistence of bubbles and negative bubbles. Unpublished working paper. New York University.

Adams, R., Ferreira, D., 2009. Women in the Boardroom and their impact on Governance and Performance. Journal of Financial Economics. 94, 291–309.

Adams, R., Mehran, H., 2012. Bank Board Structure and Performance: Evidence for Large Bank Holding Companies. Journal of Financial Intermediation. 21, 243-267.

Atkeson, A.G., Eisfeldt, A.L., Weill, P.O., 2014. Measuring the financial soundness of U.S. Firms 1926-2012. Working paper, University of California Los Angeles, NBER, and Federal Reserve Bank of Minneapolis

Balachandran, S., Kogut, B., Harnal, H., 2010. The probability of default, excessive risk, and Executive compensation: a study of financial services firms from 1995 to 2008. Unpublished working paper. Columbia University.

Beasley, M.S., 1996. An empirical analysis of the relation between board of director composition and financial statement fraud. Accounting Review. 71, 443-465.
Bebchuk, L. A., Cohen, A., Spamann, H., 2010. The Wages of Failure: Executive Compensation at Bear Stearns and Lehman 2000-2008. Yale Journal on Regulation. 27, 257-282.

Berger, A.N., Bouwman, C.H.S., 2009. Bank Liquidity Creation. The Review of Financial Studies 22, 3779–3837.

Berger, A.N., Bouwman, C.H.S., 2013. How does capital affect bank performance during financial crises? Journal of Financial Economics 109, 146–176.

Bhagat, S., Bolton, B., 2014. Financial crisis and bank executive incentive compensation. Journal of Corporate Finance. 25, 313-341.

Bharath, S.T., Sunder, J., Sunder, S.V., 2008. Accounting quality and debt contracting. Accounting Review. 23, 1–28.

Bhojraj, S., Sengupta, P., 2003. Effects of corporate governance on bond ratings and yields: The role of institutional investors and outside directors. Journal of Business. 76, 455-476.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy. 81, 637-654.

Blume, M. E., 1971. On the assesment of risk. Journal of Finance . 26, 1-10.
Bolton, P., Mehran, H., Shapiro, J., 2015. Executive compensation and risk taking. Review of Finance. 19, 2139-2181.

Boyallian, P., Ruiz-Verdú, P., 2015. CEO risk-taking incentives and bank failure during the 2007-2010 financial crisis. Unpublished working paper. Carlos III University.

Brewer, E., Hunter, W.C., Jackson, W.E., 2004. Investment opportunity set, product mix, and the relationship between bank CEO compensation and risk-taking. Working Paper. Federal Reserve Bank of Atlanta.

Chan, L. K. C ., Lakonishok, J., 1992. Robust Measurement of Beta Risk. Journal of Financial and Quantitative Analysis. 27, 265-282

Chava, S., Livdan, D., Purnanandam, A., 2009. Do Shareholder Rights Affect the Cost of Bank Loans? Review of Financial Studies. 22, 2973–3004.

Chava, S., Roberts, M.R., 2008. How does financing impact investment? The role of debt covenants. Journal of Finance. 63, 2085-2121.

Chen, C. R., Steiner, T.L., Whyte, A.M., 2006. Does Stock Option-Based Executive Compensation Induce Risk-Taking? An Analysis of the Banking Industry. Journal of Banking and Finance. 30, 915-945.
Cheng, I., Hong, H., Scheinkman, J., 2015. Yesterday's Heroes: Compensation and Risk at Financial Firms. Journal of Finance 70, 839-879.

Chesney, M., Stromberg, J., Wagner, A., 2016. Managerial Incentives to Take Asset Risk. Unpublished working paper. University of Zurich.

Coles, J. L.; Daniel, N. D., Naveen, L., 2006. Managerial Incentives and Risk-Taking. Journal of Financial Economics. 79, 431–468.

Coles, J. L., Daniel, N. D., Naveen, L., 2013. Calculation of compensation incentives and firm-related wealth using execucomp: Data, program, and explanation. Working paper. Arizona State University.

Core, J., Guay, W., 2002. Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility. Journal of Accounting Research. 40, 613–630.

Dayton, N., 1984. Corporate governance: The other side of the coin. Harvard Business Review. 62, 34-37.

Demsetz, R.S., Mark, R.S., Philip, E.S., 1996. Banks with Something to Lose: The Disciplinary Role of the Franchise Value. *Federal Reserve Bank of New York Economic Policy Review*. 2, 1–14.

DeYoung, R., Peng, E., Yan, M., 2013. Executive Compensation and Business Policy Choices at U.S. Commercial Banks. Journal of Financial and Quantitative Analysis. 48, 165-196.
Diamond, D., 1984. Financial Intermediation and Delegated Monitoring. Review of Economic Studies. 51, 393–414.

Donelson, D.C., Jennings, R., McInnis, J.M., 2014. Debt contracting and financial reporting: Evidence from a survey of commercial lenders. Working paper. University of Texas.

Fama, E.F., 1985. What’s Different about Banks? Journal of Monetary Economics. 15, 29 - 39.

Fama, E.F., Jensen, M.C., 1983. Separation of ownership and control. Journal of Law and Economics. 26, 301–325.

Fich, E.M., Shivdasani, A., 2006. Are Busy Boards Effective Monitors? Journal of Finance. 61, 689-724.

Fields, L.P., Fraser, D.R., Subrahmanyam, A., 2012. Board quality and the cost of debt capital: The case of bank loans. Journal of Banking and Finance. 36, 1536–1547.

Gande, A., Kalpathy, S.L., 2017. CEO compensation and risk-taking at financial firms: Evidence from U.S. federal loan assistance. Journal of Corporate Finance. 47, 131-150.

Gilchrist, S., Sim, J.W., Zakrajsek, E., 2014. Uncertainty, financial frictions, and investment dynamics. Federal Reserve Board. FEDS working paper.

Graham, J. R., Li, S., Qiu, J., 2008. Corporate misreporting and bank loan contracting. Journal of Financial Economics. 88, 44–61.

Guay, W. R., 1999. The Sensitivity of CEO Wealth to Equity Risk: An Analysis of the Magnitude and Determinants. Journal of Financial Economics. 53, 43–71.

Hagendorff, J., Vallascas, F., 2011. CEO Pay Incentives & Risk-Taking: Evidence from Bank Acquisitions. Journal of Corporate Finance. 17, 1078–1095.

Hasan, I., Hoi, S., Wu, Q., Zhang, H., 2014. Beauty is in the eye of the beholder: The effect of corporate tax avoidance on the cost of bank loans. Journal of Financial Economics. 113, 109–130.

Hasan, I., Hoi, C.K., Wu, Q., Zhang, H., 2017. Social capital & debt contracting: Evidence from bank loans and public bonds. Journal of Financial and Quantitative Analysis. 52, 1017-1047.
Hoechle, D., Schmid, M., Walter, I., Yermack, D., 2012. How much of the diversification discount can be explained by poor corporate governance? Journal of Financial Economics. 103, 41–60.

Houston, J. F., James, C., 1995. CEO Compensation and Bank Risk: Is Compensation in Banking Structured to Promote Risk Taking? Journal of Monetary Economics. 36, 405–431.
Houston, J., Jiang, L., Lin, C., Ma, Y., 2014. Political connections and the cost of bank loans. Journal of Accounting Research. 52, 193–243.

Hubbard, R. G., Palia, D., 1995. Executive Pay and Performance: Evidence from the U.S. Banking Industry. Journal of Financial Economics. 39, 105–130.

Jeitschko, T. D., Jeung, S. D., 2005. Incentives for risk-taking in banking: a unified approach. Journal of Banking and Finance. 29, 759–777.

Jensen, M. C., Meckling, W. H., 1976. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. Journal of Financial Economics. 3, 305–360.

Jensen, M.C., 1986. Agency costs of Free cash flow, Corporate Finance and Takeovers. American Economic Review. 76, 323–329.

John, K., Qian, Y.., 2003. Incentive Features in CEO Compensation in the Banking Industry. Economic Policy Review. 9, 109-121.

Kabir, R., Li, H., Veld-Merkoulova, Y.V., 2013. Executive compensation and the cost of debt. Journal of Banking and Finance. 37, 2893–2907.

Keeley, M.C., 1990. Deposit Insurance, Risk, and Market Power in Banking. *American Economic Review*. 80, 1183-1200.

Klock, M.S., Mansi, S.A., Maxwell, W.F., 2005. Does Corporate Governance Matter to Bondholders? Journal of Financial and Quantitative Analysis. 40, 693–719.

Knopf, J. D., Nam, J., Thornton, J.H., 2002. The Volatility and Price Sensitivities of Managerial Stock Option Portfolios and Corporate Hedging. Journal of Finance. 57, 801–813.

Lin, C., Ma, Y., Malatesta, P., Xuan, Y., 2011. Ownership structure and the cost of corporate borrowing. Journal of Financial Economics. 100, 1–23.

Machauer, A., Weber, M., 1998. Bank behavior based on internal credit ratings of borrowers. Journal of Banking & Finance. 20, 1355-1383.

Mehran, H., Rosenberg, J., 2007. The effect of CEO stock options on bank investment choice, borrowing, and capital. Unpublished working paper. Federal Reserve Bank of New York.
Merton, R. C., 1973. Theory of rational option pricing. Bell Journal of Economics and Management Science. 4, 141-183.

Minton, B.A., Schrand, C.M., Walther, B.R., 2002. The Role of Volatility in Forecasting. Review of Accounting Studies. 7, 195-215.

Minton, B.A., Schrand, C.M., 1999. The Impact of Cash Flow Volatility on Discretionary Investment and the Cost of Debt and Equity Financing. Journal of Financial Economics. 54, 423-460.

Myers, S., 1977. Determinants of corporate borrowing. Journal of Financial Economics. 5, 147–175.

Rahaman, M.M., Zaman, A.A., 2013. Management quality and the cost of debt: Does management matter to lenders? Journal of Banking and Finance. 37, 854-874.

Roberts, M.R., 2015. The role of dynamic renegotiation and asymmetric information in financial contracting. Journal of Financial Economics. 116, 61–81.

Roberts, M.R., Sufi, A., 2009. Renegotiation of financial contracts: Evidence from private credit agreements. Journal of Financial Economics. 93, 159–184.

Scordis, N., Barrese, J., Wang, P., 2008. The Impact of Cash Flow Volatility on Systematic Risk. Journal of Insurance Issues. 31, 43-71.

Smith, C. W., Stulz, R.M., 1985. The Determinants of Firms' Hedging Policies. Journal of Financial and Quantitative Analysis. 20, 391–405.

Strahan, P.E., 1999. Borrower risk and the price and non-price terms of bank loans. Unpublished working paper, Boston College.

Sufi, A., 2007. Information asymmetry and financing arrangements: Evidence from syndicated loans. Journal of Finance. 62, 629–668.

Weisbach, M.S., 1988. Outside Directors and CEO Turnover. Journal of Financial Economics. 20, 431–460.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica. 48, 817–838

DECISION SCIENCES INSTITUTE

Independent Directors' Connection and Firm Operating Performance

Valentine Siagian
Yuan Ze University, Taiwan
Email: s1049430@mail.yzu.edu.tw

Chia-Wei Huang
National Taiwan Normal University, Taiwan
Email: huangcw@ntnu.edu.tw

Huai-Chun Lo
Yuan Ze University,Taiwan
Email: hclo@saturn.yzu.edu.tw

ABSTRACT

This study explores the connection of independent directors to operating performance. Independent directors' connection is measured by their primary companies' lobbying activities. We find that firms with independent directors from a lobbied company have significantly positive impact to the operating performance. Reduction in the firm's financing costs can be a channel through which outside directors' connection affect firm operating performance. Our evidence suggests that in addition to monitoring and advising functions, independent directors can contribute their connection through political influence to the operating performance of the appointed firms.

KEYWORDS: Independent Directors, Lobbying Activities, Operating Performance

INTRODUCTION

Board existence play an important role on advising and monitoring (Coles et al, 2008; Adams and Ferreira, 2007; Adams and Mehran, 2003; Agrawal and Knoeber, 2001). Each firms represented by many directors. Board of directors consist of inside directors and outside directors (often called independent directors). Independent directors have no right on removing firms' assets. After the SOX regulation on 2002, firms are required to have independent directors of more than 50% of the total board members. These independent directors might work as an executives on other companies (Agrawal and Knoeber, 2001), even more as CEO. Firms where they work as executives called primary company.

Existing literature has shown that contribution of independent directors to firm performance is insignificant (MacAvoy, Cantor, Dana, and Peck, 1983; Bhagat and Black, 1999, 2002; Hermalin and Weisbach, 1991; Klein, 1998) or even negative (Agrawal and Knoeber, 1996). There are exceptions Rosenstein and Wyatt (1990) who show that stock price reacts positively to the nomination of independent directors to the board. Holthausen, and Larcker (1999) find a positive relationship between the fraction of independent directors and market-to-book ratio.

Connection often related to political connection. Prior literature presents only inconclusive evidence on the effect of corporate political activity on firm operating performance (Faccio, 2006; Goldman et al., 2009; Cooper et al., 2010; Hill et al., 2013; Chen et al., 2015; Ansolabehere et al., 2004; Hadani and Schuler, 2013). When it comes to lobbying activities, Hill et al. (2013) and Chen et al. (2015) both find a positive link between firm performance and lobbying expenditures. Other researches find either no link (Ansolabehere et al., 2004) or a negative link (Hadani and Schuler, 2013) between corporate lobbying and firm performance. Lobbying activities found to be positively related to financial performance (Chen et al 2015; Borisov et al 2015). The study of Duchin et al (2010) suggests that independent directors might affect appointed firm positively on their operating performance.

Independent directors which primary company executes lobbying activities called ID-PCL. We aim to study whether appointed firm with ID-PCL have better or lower firm performance.

After the SOX regulation on 2002, firms are required to have independent directors of more than 50% of the total board members. With major rules of board of directors which are advising and monitoring (Coles et al, 2008; Adams and Ferreira, 2007; Adams and Mehran, 2003; Agrawal and Knoeber, 2001), appointed firms with ID-PCL would likely to have a better firm performance. Daily experience of independent directors from their primary company might help on how they monitoring and advising appointed firms.

In this study, we wanted to see how independent directors affect the performance of their appointed firms through their connection which is lobbying activities of their primary companies. Previous studies mention the regulatory definition of an independent directors (e.g., Duchin, Matsusaka, and Ozbas, 2010; Cohen, Frazzini, and Malloy, 2012). Baysinger and Hoskisson (1990) mention that board is an important instrument of internal control. Larger boards potentially bring more experience and knowledge and offer better advice (Dalton, Daily, Johnson, and Ellstrand, 1999), complex firms should have larger boards size, which open an opportunity to welcome independent directors from various primary companies. This phenomenon supported when appointed firm bring CEO's of other firms to join the board members as independent directors shows that each firm at least have independent directors with primary company.

Most of the papers in the literature examine board characteristics such as the fraction of independent directors or specific types of independent directors and their relationship to corporate performance and policies.

Agrawal and Knoeber (2001) suggest that independent directors are senior managers in other firms and these directors play a political role. Furthermore, they focus on the independent directors that can aid dealings of a firm by using their skills to predict or affect government and prospective business collaboration, because these directors have friendship with decision makers and experience to cut deal with government.

The question of whether the appointed firms with ID-PCL have a better operating performance has been difficult to test due to simultaneity between lobbying activities of the independent director's primary company and firm performance. To address this issue we run natural experiment using modified difference-in-difference following Coles et al (2014). Also, we run two-stage least square to help us as our lobbying activities might be exogenous. To overcome endogeneity issue, we adopt the pre and post SOX event in 2002 and see the changes of the independent directors' composition between 2002 and 2003 while adjusting to the new regulation. We find that appointed firms with ID-PCL have a better operating performance.

We document a positive relation between independent directors (measured by lobbying activities of the independent director's primary company) and operating performance (measured by return on assets, return on equity and net profit margin). These findings supported by previous literatures showing independent directors tend to have wider networking and experience and thus affect areas such as tax rates (Richter et al., 2009), regulatory oversight

(Bonardi et al., 2006; Lux et al., 2011), tariffs (Mayda et al., 2010), and government contracts (Goldman et al., 2013).
We next explore possible channel on independent directors and operating performance. Study of Anderson et al (2004) suggests that cost of debt are related to board independence. Hovakimian et al, (2001) explain on how firm move based on the tradeoff between cost and benefit of debt to adjust capital structure. Hackbarth & Mauer (2012) study on how firm predict on priority structure including financial leverage.
Furthermore, we wanted to test whether the existence ID-PCL could lower cost of debt. To do so we regress ID-PCL with cost of debt. We find that firms with ID-PCL have lower cost of debt. This results shown that the appointed firms with ID-PCL have a better operating performance which shown through our channels that firms with ID-PCL would enjoy lower cost of debt.
The main contribution of this paper is to enrich literature for the function of the board. Compared with previous studies of board functions as monitoring and advising (Coles et al, 2008; Adams and Ferreira, 2007; Adams and Mehran, 2003; Agrawal and Knoeber, 2001), we provide a new angle to argue that independent directors could also bring connection through beneficial social networking into the appointed firm. This kind of networking is profitable for the company, shown by lower cost of debt through the existence of ID-PCL. By the existence of ID-PCL with networking and political influence, could lower the assymetric information problem and open wider opportunities for the firm.
The rest of the paper is organized as follows. Section 2 contains the literature review and hypothesis development, while Section 3 discusses the data and methodology. Section 4 presents and analyzes the empirical results. Section 5 figures the endogeneity issue. Section 6 displays robustness test. Section 7 concludes. Variable definitions are explained in Appendix A.

LITERATURE REVIEW

Independent Directors' Connection and Firm Operating Performance

NYSE and Nasdaq regulations approved by the SEC in 2002 go beyond SOX and require a majority of directors on the board to be independent. Firms responded to the new regulations by significantly increasing the representation of independent directors on their boards and committees over time. Fama and Jensen (1983) theorize that the board of directors is the highest internal control mechanism responsible for monitoring the actions of top management. There are lots of theoretical and empirical result which shown the function of board of directors as monitoring and advising the firm, or both. Particularly, the outside independent directors are into that roles inside the board. Masulis and Mobbs (2014) suggest that independent directors are attracted by available board seats. Outside independent directors are expected to avoid potential conflict of interest (Ryan and Wiggins, 2004) that can reduce their monitoring capacity. While Fracassi and Tate (2012) mention that firm with powerful CEO are appointing directors that have a background story or ties to them. Agrawal and Knoeber (2001) discuss whether outside director with political background are valuable to firms and the results shown that outside directors are valuable.
Nguyen and Nielsen (2010) mention that the after effect of the death of independent directors, firm value decreased. Fahlenbrach et al (2016) suggest that departure of independent directors impact stock and operating performance of the firms negatively. Independent director and inside director have their own roles on affecting firm operating performance. Fama and Jensen (1983) find that inside directors are more likely to possess superior information that, together with their experience, allows them to contribute to firm value. When board allows both independent directors and inside directors to perform their roles optimally then they are valuable to the firm. Thus, firms with ID-PCL would have a better operating performance.
This reasoning leads to our hypotheses:

H*₁*:* *Firms with ID-PCL are associated with better operating performance

Independent Directors' Connection and Cost of Debt

Larger firms are more likely to have external contracting relationships (Booth and Deli, 1996) and, thus, require larger boards (Pfeffer, 1972). Knyazeva, Knyazeva, and Masulis (2013) find that larger firms tend to draw experience director candidates even from distant location. Helleand & Sykuta (2004) find that the number of political directors increases as firms shift from market to political competition.

Firms most likely scout CEO from other company as their independent directors (Fahlenbrach et al., 2010). If the CEO is actively involved in lobbying activities from their primary company, he will most likely agree to the decision made by the CEO's as independent directors in the appointed firms. Hwang and Kim (2009) mentioned that 87% boards are precisely independent but only 25% are socially independent. This means that independent directors are more likely ties to the firms. In this paper, we wanted to show empirically that the existence of independent directors which primary company execute lobbying activities have positively strengthened appointed firm operating performance. Hill et al (2013) shares that shareholders value the lobbying activities pursued by management on their behalf. Ryan et al (2004) mentioned that independent directors have bargaining power over CEO.

Lobbying activity was a critical component in the rise and fall of the firms. Corporate lobbying activities are applied most likely from in-house lobbyist or lobbying firms. Corporation can also spend on gifts, meals and free trips for legislators. In general, since corporate lobbying activities are associated with a firm's business strategy and culture and bring multitude of advantages, lobbying activities could bring more connections to the firm. Lobbying activities are formed to influence other parties who are decision-maker [Lobbying are more likely connected to political activity: Agrawal and Knoeber (2001), de Figueiredo and Tiller (2006), Fisman (2001), Ansolabehere et al. (2004), Hillman et al. (2004), Fan et al. (2007), Kim (2008), Alexander et al. (2009), Goldman et al. (2009), Igan et al. (2009), Cooper et al. (2010), Yu and Yu (2011), and Hill et al. (2013).] Lobbying activities also open to a wider networking and thus affect areas such as tax rates (Richter et al., 2009), regulatory oversight (Bonardi et al., 2006; Lux et al., 2011), tariffs (Mayda et al., 2010), and government contracts (Goldman et al., 2013).

This reasoning leads to our hypotheses:

H*₂*:* *Firms with ID-PCL are associated with lower cost of debt.

DATA, METHODOLOGY AND KEY VARIABLES

Data

Directors data was taken from RiskMetrics database for the years 1996 to 2015. We focus on directorships within RiskMetrics database. RiskMetrics contains individual director information for the S&P 1500 firm variables resulting in 190,441 director-firm-year observations. We filtered the "board affiliation" variable, then we exclude executive directors and only retrieve independent directors' data.

In 1995 there was Lobbying Disclosure Act which requires a registration and reporting on lobbying activities. Registrant must file reports detailing the issue of lobby and the total amount spent. Brasher and Lowery (2006) develop a different test on new model of corporate political activity with Hoover's Handbook, while others test on the corporation provided by Fortune 500 and 1000. Our lobbying expenditure data taken from Center for Responsive Politics (CRP) website [https://www.opensecrets.org/lobby/] that provides lobbying expenditure of registered firms that covers public firm, private firms and institutions not limited only to US data. The CRP data includes lobbying, PAC contributions, and soft money categories. The CRP used the amount reported by the organization (including both in-house lobbying and external lobbyist filings) as the total lobbying expenditure for the period. Even if an organization does not report, the CRP sums lobbying expenditures reported by its external lobbyist. Lobbying data are available from 1998. Since the CRP does not use company identifiers (e.g., CUSIP, PERMNO, etc) we manually verified the names of the public firms to ensure the matching of RiskMetrics and CRP lobbying data. Lobbying expenditure data was recorded one by one as one firms might have a lot of independent director coming from different primary company. Accounting variables are obtained from Compustat database and stock return are generated from CRSP. Institutional ownership data are obtained from Thomson 13F report. To construct data from firm-year-director data (primary company) into firm-year data we create our own Ticker for director names, since director identifier is not available for every director on the list. To match the data of RiskMetrics and Compustat we use CUSIP identifier. To match the data of RiskMetrics and CRSP we use GVKEY.

Measuring ID-PCL (Independent Director with Primary Company That Lobby)
Our measurement for ID-PCL (independent directors with primary company that lobby) is the lobbying expenditure spent by independent director's primary company. RiskMetrics contains individual director information for the S&P 1500 firm variables resulting in 190,441 director-firm-year observation. Applying the "board affiliation" variable, we exclude executive directors and only retrieve independent director data. From Risk Metrics data, there are 141,474 independent directors-years observations. Then filtered by "primary company" variable we exclude independent directors without primary company and resulting 97,366 independent directors-firm-year observation with primary company. Here we generate our own ticker for each director [Names of the directors are many. Some with tittle, some with space. We generate our own ticker for directors to match repetition of director-primary company easier.]
Filtering the year from 1998 – 2015 give us 97,366 independent directors with primary company, there are 19,318 primary company which spent lobbying expenditure. We exclude all the primary company with no record on CRP lobbying data. We sum up the amount based on the appointed firms and put it as lobbying expenditure for the appointed firms. We sum up the total lobbying expenditure of the ID-PCL from the same firm which leads us to the final 12,318 firm-year observation. We collect accounting variables and other variables needed based on 1,743 unique appointed firms.

Measuring Firm Operating Performance
Most of the literatures measure firm performance with return on assets (ROA), return on equity (ROE), and net profit margin (NPM). Following paper Gompers (2010) we measure firm operating performance as ROA, ROE and NPM. [Appendix A shows variable explanations used in this paper].

Possible Channel
To extend our evidence on the effect of ID-PCL to firm operating performance, we run regression of ID-PCL to cost of debt. Previous research (construct cost of debt as total interest expense plus capitalized interest divided by total assets. We construct cost of debt measured by total interest expense for long debt divided by total debt, since we wanted to see the effect of the long-term cost of debt specifically.

Control Variables
Consistent with prior literature, control variables include controls for firm characteristics: firm size (log sales), dividend, leverage, and return; director characteristics: independence (percentage of independent director) female on board (dummy variable with condition of 1 if at least one female on board and 0 otherwise), board size (total number of the board); and

governance control: takeover index (construct by Cain et al), institutional ownership (percentage of ownership of the shares), e-index (the sum of cboard, labylw, lachtr, supermajor, golden parachute, ppill) following Gompers (2010) paper, and herfindahl index (the sum of squared market share).

Summary Statistics

Table 1 reports descriptive statistics for the sample of S&P 1500 appointed firms with independent director which primary company execute lobbying activities. The minimum of lobbying activities of $0 are reported on CRP lobbying data. Some primary company have been active on lobbying activities, however there are some years that they report $0 lobbying expenditure. Previous lobbying studies use the minimum reporting amount of lobbying expenditure which is $10,000 [see paper Borisov et al 2015 for further details on the lobbying expenditure reported]. We include those $0 lobbying activities as long as they are recorded in CRP lobbying data. The average of $840 million per year is the lobbying activities of the primary company. Since one firm will have a combination of several independent directors from different primary companies, the highest dollar amount of ID-PCL is $1,3 billion. Firms with ID-PCL have an average of 5% return on assets, average of 10% of return on equity and 5% of net profit margin. The size of the appointed firms with ID-PCL is included in the top rank firms that needs larger board with wider networking.

Table 1: Summary Statistics

	N	Mean	Std dev	Min	Max
Log ID-PCL	12318	6.18	6.68	0.00	16.07
Donation (US$ million)	12318	0.84	2.51	0.00	13.2
ROA	11947	0.05	0.09	-0.38	0.25
ROE	11947	0.10	0.25	-1.22	1.07
NPM	11947	0.05	0.14	-0.85	0.34
Cost of Debt	12318	0.06	4.40	0.00	456.39
Firm Characteristics					
Firm Size	12317	6971	23096	0.00	483521
Dividend	10377	0.01	0.02	0.00	0.11
Leverage	10377	0.21	0.16	0.00	0.65
Return	10575	0.08	0.35	-0.96	8.43
Director Characteristics					
Independence	10575	0.75	0.15	0.09	1.00
Director Age	10478	62.11	4.46	31.00	81.75
Board Tenure	9921	9.48	3.51	1.00	31.00
Female Dir	10575	0.46	0.50	0.00	1.00
Governance controls					
Takeover	12318	0.14	0.11	0.00	0.82
Inst. Own	10575	0.00	0.00	0.00	0.03
e_Index	12318	2.96	1.11	0.00	6.00
HHI	10571	0.02	0.01	0.00	0.18

Table 2: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1)ID-PCL	1																
(2)ROA	**0.06**	1															
(3)ROE	**0.11**	**0.83**	1														
(4)NPM	**0.08**	**0.83**	**0.70**	1													
(5)COD	-0.01	0.01	0.02	0.02	1												
(6)SIZE	**0.17**	**0.04**	**0.07**	**0.03**	0.07	1											
(7)DIV	**0.12**	**0.26**	**0.28**	**0.17**	0.01	**0.08**	1										
(8)LEV	**0.10**	**-0.14**	**0.02**	**-0.02**	-0.01	**0.07**	0.01	1									
(9)RET	0.01	**0.14**	**0.13**	**0.10**	-0.01	-0.01	0.01	0.01	1								
(10)IND	**0.25**	-0.01	**0.04**	**0.03**	0.07	**0.07**	**0.06**	**0.11**	0.01	1							
(11)AGE	**-0.09**	**0.02**	0.02	**0.05**	0.08	**0.04**	**0.08**	**0.03**	0.01	**0.05**	1						
(12)TENU	**-0.07**	**0.08**	**0.05**	**0.07**	-0.01	**-0.05**	**0.11**	**-0.03**	**0.03**	**-0.07**	**0.40**	1					
(13)FEM	**0.06**	0.05	0.01	-0.01	-0.01	-0.01	**0.05**	**0.08**	0.01	**0.04**	-0.01	**-0.02**	1				
(14)TOI	**0.09**	**0.07**	**0.07**	**0.05**	-0.01	**0.13**	**0.12**	**-0.03**	**0.07**	-0.01	**0.05**	**0.13**	**0.10**	1			
(15)IO	**-0.17**	**-0.16**	**-0.17**	**-0.15**	-0.01	**-0.22**	**-0.16**	**-0.15**	**-0.06**	**-0.03**	-0.02*	-0.02	**-0.09**	**-0.05**	1		
(16)eIndex	**0.04**	-0.07	-0.01	0.02	-0.01	**-0.07**	-0.01	0.01	**-0.04**	**0.19**	**0.09**	**0.07**	-0.01	**-0.02**	**0.15**	1	
(17)HHI	-0.02	**0.05**	**0.056**	0.01	0.01	**0.03**	**0.04**	**0.07**	0.02	**-0.16**	-0.01	0.01	**0.11**	**0.19**	**-0.02**	**-0.03**	1

Notes: This table presents Pearson's correlations of variables. Bolded values indicate correlation significance at 5 percent level or better. All variables are defined in the appendix.

EMPIRICAL RESULTS

To test whether firms with ID-PCL have a better firm operating performance, we run ordinary least square (OLS) regression for all our dependent variables: return on assets, return on equity, net profit margin, and cost of debt.

ID-PCL and Firm Operating Performance

We study the ID-PCL effect on firm operating performance in order to find the relationship by using return on assets (ROA), return on equity (ROE) and net profit margin (NPM) as the measurement of operating performance and use following models for estimation:

$$ROA_{i,t} = \beta_0 + \beta_1 \log(ID\text{-}PCL)_{(t-1)} \ \ \beta_{2-5} Firm\ Characteristics_{i,(t-1)} + \beta_{6-9} Director\ Characteristics_{i,(t-1)} + \beta_{10-13} Governance\ \ Controls_{i,(t-1)} + Year\ Dummies_{i,t} + Industry\ Dummies_{i,t} + \varepsilon \quad (1)$$

$$ROE_{i,t} = \beta_0 + \beta_1 \log(ID\text{-}PCL)_{(t-1)} \ \ \beta_{2-5} Firm\ Characteristics_{i,(t-1)} + \beta_{6-9} Director\ Characteristics_{i,(t-1)} + \beta_{10-13} Governance\ \ Controls_{i,(t-1)} + Year\ Dummies_{i,t} + Industry\ Dummies_{i,t} + \varepsilon \quad (2)$$

$$NPM_{i,t} = \beta_0 + \beta_1 \log(ID\text{-}PCL)_{(t-1)} \ \ \beta_{2-5} Firm\ Characteristics_{i,(t-1)} + \beta_{6-9} Director\ Characteristics_{i,(t-1)} + \beta_{10-13} Governance\ \ Controls_{i,(t-1)} + Year\ Dummies_{i,t} + Industry\ Dummies_{i,t} + \varepsilon \quad (3)$$

Where ROA is the ratio of total assets divided by net income. ROE is the ratio of total equity divided by net income and NPM is the ratio of sales divided by net income. ID-PCL is the log of the sum of lobbying expenditure of primary company's independent directors of firm *i* in year *t*. For controlling unobserved factors of a firm, we use a fixed effect model to control the time effect by using year dummies to capture factors which affect all firms at the same time and use 48 Fama industry classification to control industry effect to capture all firms in the same industry. All variables are winsorized at 1% and 99%.

Table 3 reports the results of ID-PCL and firm operating performance. We examine each dependent variable with single regression in model (1), model (3) and model (5), and including all control variables in model (2), model (4) and model (6). Coefficient of ID-PCL are positive and significant in 1% level, respectively, indicating that firms with ID-PCL have a better operating performance of 0.06% return on assets, 0.2% of return on equity and 0.12% of profit margin and statistically significant in 1% level after controlling firm characteristics, director characteristics and governance control.

Table 3: Effect of ID-PCL on Firm Operating Performance

	Dependent Variables = ROA, ROE, NPM					
	Return on Assets		Return on Equity		Net Profit Margin	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0906***	0.0489	0.198***	-0.00260	0.102***	0.00498
	(6.20)	(1.83)	(4.73)	(-0.03)	(4.25)	(0.11)
ID-PCL	0.0007***	0.0006***	0.0033***	0.002***	0.0016***	0.0012***
	(6.50)	(4.47)	(9.61)	(4.71)	(7.98)	(4.93)
$Q_{(t-1)}$		0.0001*		-0.0003*		-0.0001
		(2.40)		(-2.57)		(-0.40)
$SIZE_{(t-1)}$		1.29		0.0001*		-1.04
		(0.32)		(2.11)		(-0.02)
$DIV_{t-1)}$		1.018***		3.009***		0.977***
		(20.95)		(21.42)		(12.04)
$LEV_{(t-1)}$		-0.0674***		0.0296		-0.0384***
		(-10.34)		(1.57)		(-3.52)
$RET_{(t-1)}$		0.0293***		0.0735***		0.0287***
		(10.98)		(9.52)		(6.44)
$IND_{(t-1)}$		-0.0176*		0.0282		-0.0153
		(-2.31)		(1.28)		(-1.20)
$AGE_{(t-1)}$		-0.0001		-0.0002		0.0002
		(-0.46)		(-0.28)		(0.59)
$TENU_{(t-1)}$		0.0015***		0.0026**		0.0018***
		(4.97)		(3.09)		(3.69)
$FEM_{(t-1)}$		-0.0026		-0.0036		-0.0083*
		(-1.23)		(-0.58)		(-2.29)
$TOI_{(t-1)}$		-0.0343**		-0.0420		0.0177
		(-3.06)		(-1.30)		(0.95)
$IO_{(t-1)}$		-4.889***		-12.17***		-5.763***
		(-14.31)		(-12.32)		(-10.10)
$eIndex_{(t-1)}$		0.003***		0.0082**		0.003
		(3.30)		(3.06)		(1.96)
$HHI_{t-1)}$		0.361**		0.996**		0.509*
		(2.81)		(2.68)		(2.38)
Industry FE		Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes
N	11947	9921	11947	9921	11947	9921
Adj. R.Sq	0.067	0.183	0.071	0.172	0.055	0.059

This table reports OLS estimates of the relationship between ID-PCL and operating performance. Firm operating performance is measured by: ROA= net income divided by total assets in year t. ROE= net income divided by total equity in year t.; NPM=net income divided by total sales in year t. Independent variable is ID-PCL= lobbying expenditure executed by primary company of independent director in year t. Control variables for firm characteristics, director characteristics and governance characteristics. The samples are from S&P 1500 with coverage of CRSP and COMPUSTAT files annually, for the years 1998-2015. All regressions include Fama and French 48 industry fixed effects and year fixed effects. t statistics in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001 indicate the regression coefficient significantly different from zero at the 5%, 1% and 0,1% levels, respectively.

ID-PCL and Cost of Debt

We want to extend our analysis and give more evidence through some channels on ID-PCL and firm operating performance. We test leverage and cost of debt as the channels. To test the channel of independent director with primary company lobbying expenditure, we study the ID-PCL to book leverage (BL), market leverage (ML) and cost of debt with these following models:

$$COD_{i,t} = \beta_0 + \beta_1 \log(ID\text{-}PCL)_{(t-1)} + \beta_{2\text{-}5}\,Firm\ Characteristics_{i,(t-1)} + \beta_{6\text{-}9}\,Director\ Characteristics_{i,(t-1)} + \beta_{10\text{-}13}\,Governance\ Controls_{i,(t-1)} + Year\ Dummies_{i,t} + Industry\ Dummies_{i,t} + \varepsilon \quad (4)$$

Where cost of debt measured by interest expense of long-term debt divided by total debt. ID-PCL is log of the sum of lobbying expenditure of primary company's independent directors of firm *i* in year *t*. For controlling unobserved factors of a firm, we use a fixed effect model to control the time effect by using year dummies to capture factors which affect all firms at the same time and use 48 Fama industry classification to control industry effect to capture all firms in the same industry. All variables are winsorized at 1% and 99%.

Table 4 report the results of ID-PCL and cost of debt. Coefficient of ID-PCL are negative and significant in 1% level, respectively, indicating that firms with more independent director with primary company that lobby (ID-PCL) have a lower cost of debt of 2%. Economically significant when the firms with ID-PCL would enjoy a lower cost of debt. In addition to Gamba & Triantis (2008) that shows the cost of external financings is on of key factor to financing flexibility that determine firm's growth.

ENDOGENEITY ISSUE

Difference in difference

Endogeneity is a common issue in the literature on boards of directors (Hermalin and Weisbach, 2003). Endogeneity is an important issue for corporate governance study (Coles et al 2011). Specifically, there is possibility that ID-PCL and firm operating performance variables are high due to an unobserved variable. However, it is not easy to find an instrument that is related to ID-PCL but is not related to any operating performance variables. Following paper Coles et al (2014) we run a natural experiment to help us address endogeneity concerns.

We use the SOX rules in 2002 which require all listed firms to have a majority of independent director on their board. Because the rules were adopted after the passage of SOX, we refer to 2003 as the starting year of Post-SOX period. Pre-SOX noncompliant firms were required to have majority of independent director (Engel et al, 2007; Linck et al, 2009). With this, an exogenous increase was found in ID-PCL.

Following Coles et al (2014 we modify the Bertrand and Mullainathan (2003) difference-in-difference (DID) methodology. The difference will be reflected on the possibility that SOX have a direct effect on operating performance and the channel.

We modify the typical DID setup to see the effect of ID-PCL (with the term called "clean" effect). Commonly, DID setup for Q, for example, would be the regression of Q on three dummy variables: *Post-SOX, Noncompliant*, and interaction term *Post-SOX x Noncompliant*, where *Post-SOX* is an indicator variable that equals one if the year is 2003 or later and equals 0 otherwise, and *Noncompliant* is an indicator variable that equals one if the firm was not in compliance in 2002, and equals 0 otherwise. ID-PCL is not included in the above condition, and only focus on the coefficient of *Post-SOX x Noncompliant*. However, this coefficient will capture the effect that we want to isolate and the direct effect of SOX.

Table 4: Effect of ID-PCL on Cost of Debt

	Dependent Variable = COD				
	Cost of Debt				
	(1)	(2)	(3)	(4)	(5)
Constant	-0.0047	-0.0044	0.0221*	-0.0032	0.0192
	(-0.95)	(-0.76)	(2.52)	(-0.39)	(1.71)
ID-PCL	-0.0002***	-0.0001**	-0.0002***	-0.0002**	-0.0002***
	(-3.55)	(-2.88)	(-3.65)	(-3.13)	(-3.58)
Tobin'sQ_1$_{(t-1)}$		-0.0001			-0.0001
		(-0.25)			(-0.20)
ROA_1$_{(t-1)}$		-0.0009			0.0001
		(-0.20)			(0.02)
SIZE$_{(t-1)}$		-9.13			-6.13
		(-0.67)			(-0.39)
DIV$_{t-1)}$		0.003			0.0024
		(0.14)			(0.10)
LEV$_{(t-1)}$		-0.0017			
		(-0.68)			
RET$_{(t-1)}$		-0.001			-0.0013
		(-0.90)			(-1.12)
IND$_{(t-1)}$			0.0022		0.0023
			(0.75)		(0.72)
AGE$_{(t-1)}$			-0.0003***		-0.00027*
			(-3.42)		(-2.49)
TENU$_{(t-1)}$			0.0001		-0.0001
			(0.51)		(-0.16)
FEM$_{(t-1)}$			0.0004		0.0002
			(0.48)		(0.27)
TOI$_{(t-1)}$				0.0018	0.0027
				(0.46)	(0.58)
IO$_{(t-1)}$				-0.0145	-0.0603
				(-0.11)	(-0.40)
eIndex$_{(t-1)}$				0.0009*	0.001*
				(2.37)	(2.32)
HHI$_{t-1)}$				-0.0239	-0.0416
				(-0.50)	(-0.77)
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	12318	10377	9921	10571	9921
Adj. R. Sq	0.026	0.022	0.022	0.003	0.003

This table reports OLS estimates of the relationship between ID-PCL and cost of debt. Cost of debt is measured by: cod = total interest expense for long debt divided by total debt in year t. Independent variable is ID-PCL= lobbying expenditure in year t-1. Control variables for firm characteristics, director characteristics and governance characteristics. The samples are from S&P 1500 with coverage of CRSP and COMPUSTAT files annually, for the years 1996-2015. All regressions include Fama and French 48 industry fixed effects and year fixed effects. *t* statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ indicate the regression coefficient significantly different from zero at the 10%, 5% and 1% levels, respectively

We take ROA as an example in the equation and run this regression to all our dependent variables. The controls in the regression include independent variables used in the *ROA, ROE and NPM* regression in Table 3 and the individual dummies, and the interactions of the independent variable with the three key dummy variables: *Post-SOX, Noncompliant*, and *Post-SOX x Noncompliant*. Then we calculate joint test for $\beta_1 + \beta_3 + \beta_4 = 0$ for each dependent variables to find the "clean" estimate.

To estimate the effect of ID-PCL, we run the modified regression which includes ID-PCL and the interaction of ID-PCL with the three dummy variables:

$$\begin{aligned} \textit{Return on Assets}_{i,t} = {} & \beta_0 + \beta_1 \log(\textit{ID-PCL})_{(t-1)} + \beta_2\, \textit{POST-SOX} \times \log \textit{ID-PCL}_{(t-1)} \\ & + \beta_3\, \textit{Non Compliance} \times \log \textit{ID-PCL}_{(t-1)} \\ & + \beta_4\, \textit{POST-SOX} \times \log \textit{ID-PCL}_{(t-1)} \times \textit{Non Compliance} \\ & + \beta_5\, \textit{POST-SOX} + \beta_6\, \textit{Non Compliance} \\ & + \beta_{7\text{-}10}\, \textit{Firm Characteristics}_{i,(t-1)} + \beta_{11\text{-}14}\, \textit{Director Characteristics}_{i,(t-1)} \\ & + \beta_{15\text{-}18}\, \textit{Governance Controls}_{i,(t-1)} + \textit{Year Dummies}_{i,t} + \textit{Industry Dummies}_{i,t} + \varepsilon \end{aligned} \quad (5)$$

Table 5 provides the result of the clean estimates. Briefly, we present only the clean estimates for the total estimates of ID-PCL on each of the four variables of interest including the channel (return on assets, return on equity, net profit margin, and cost of debt). To make it easier to compare, we report the base case regression (model 2, 4, 6 of Table 3, and model 5 of Table 4).

The results in Table 5 shows that parallel to the base case, the estimates based on an exogenous shock to ID-PCL have the same sign and statistically significant in 1% level.

Table 5: Clean Estimates

Coefficient Estimate	Result from base case	"Clean" Estimate
Table 4: Model 2 (ROA)	0.000665*** 4.47	0.0263** 2.22
Table 4: Model 4 (ROE)	0.00203*** 4.71	0.0416** 2.03
Table 4: Model 6 (NPM)	0.00122*** 4.93	0.015** 2.43
Table 6: Model 5 (COD)	-0.000229*** -3.58	-0.0329** -2.13

Instrumented Variables

We run two-stage least square (2sls) with instrumented variables. We use regulatory changes to adopt more independent boards following Sarbanes-Oxley Act (SOX) regulation, which required firms to have a majority of independent directors (for a similar approach, see Duchin, Matsusaka, and Oguzhan (2010)). At the first stage we regress the exogenous variable changes

in the composition of independent director of appointed firm using the changes in regulation of SOX on 2003 as the exogenous variable and at the second stage we replace the fitted variable from the first stage.
Panel a, b, c and d of Table 6 presents the results of the test for the omitted variable problem. Each table runs five models to regress ID-PCL with each dependent variables: model (1) single

Table 6: Instrumented Variables

	First stage ID-PCL	Second Stage *ROA*	Second Stage *ROE*	Second Stage *NPM*	Second Stage *COD*
Constant	9.987***	0.0370	0.137	0.339	0.0214
	(8.91)	(0.51)	(0.66)	(1.54)	(0.91)
IV	0.0053**				
	(2.05)				
$\widehat{\mathrm{ID}-\mathrm{PCL}}$		0.0038*	0.0001*	0.0349*	-0.0022*
		(1.95)	(2.01)	(1.77)	(-1.83)
Tobin'sQ_1*(t-1)*	0.0069*	0.0001	0.0001***	0.0001	-0.0001
	(1.91)	(1.01)	(2.59)	(0.98)	(-0.59)
SIZE(t-1)	0.0001***	-9.49	0.0001	0.0001	-2.73
	(10.44)	(-0.43)	(0.35)	(1.52)	(-0.04)
DIVt-1)	29.08***	0.959***	3.359***	2.076***	0.0065
	(8.60)	(4.33)	(5.35)	(3.12)	(0.07)
LEV(t-1)	2.272***	-0.0826***	0.0086	0.0389	-0.0032
	(5.05)	(-4.42)	(0.16)	(0.69)	(-0.85)
RET(t-1)	-0.0529	0.0295***	0.0733***	0.0267***	-0.001
	(-0.27)	(10.55)	(9.25)	(3.18)	(-0.89)
IND(t-1)	9.975***	-0.0508	0.0365	0.346	0.0006
	(18.60)	(-0.69)	(0.17)	(1.56)	(0.03)
AGE(t-1)	-0.189***	0.0001	-0.0019	-0.0067	-0.0002
	(-10.94)	(0.10)	(-0.47)	(-1.55)	(-0.66)
TENU(t-1)	-0.0735***	0.0018***	0.0027	-0.0007	-0.0001
	(-3.44)	(2.95)	(1.49)	(-0.40)	(-0.27)
FEM(t-1)	0.423***	-0.0038	-0.0073	0.0101	-0.0006
	(2.93)	(-0.93)	(-0.62)	(0.81)	(-0.40)
TOI(t-1)	6.545***	-0.0566	-0.0422	0.240	0.0034
	(8.43)	(-1.13)	(-0.30)	(1.59)	(0.21)
IO(t-1)	-241.7***	-3.901**	-12.16**	-14.70***	-0.103
	(-10.06)	(-2.15)	(-2.36)	(-2.69)	(-0.22)
eIndex(t-1)	0.266***	0.002	0.0066	0.0128**	0.0007
	(3.98)	(0.88)	(1.07)	(1.96)	(1.24)
HHIt-1)	-9.312*	0.225***	0.605**	-0.0163	-0.05
	(-1.75)	(2.68)	(2.54)	(-0.06)	(-1.61)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
N	9921	9921	9921	9921	9921

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01

regression, model (2) include only control of firm characteristics, model (3) include only control of director characteristics, model (4) include only control of governance and model (5) include all control variables. Comparing to the result of OLS regression, firms that changed from less to more independent directors against firms that already had independent boards, we find that ID-PCL is significantly and positively associated with operating performance and the channels: return on assets, return on equity, net profit margin, cost of debt. The results are consistent with Table 3, and Table 4.

ROBUSTNESS TESTS

For robustness test we wanted to test whether the result would be the same when we change the measurement of operating performance. Following Chen et.al (2015) we change measurement of operating performance into financial performances: cash from operation, income before extra items, and net income.

We run the OLS regression with the following models similar to equation 1 and 2 and include all control variables:

$$\begin{aligned} \textit{Net Income}_{i,t} = {} & \beta_0 + \beta_1 \textit{ log ID-PCL}_{(t-1)} + \beta_{2\text{-}5}\textit{ Firm Characteristics}_{i,(t-1)} \\ & + \beta_{6\text{-}9}\textit{ Director Characteristics}_{i,(t-1)} + \beta_{10\text{-}13}\textit{ Governance Controls}_{i,(t-1)} \\ & + \textit{Year Dummies}_{i,t} + \textit{Industry Dummies}_{i,t} + \varepsilon \end{aligned} \tag{6}$$

$$\begin{aligned} \textit{IBEI}_{i,t} = {} & \beta_0 + \beta_1 \textit{ log ID-PCL}_{(t-1)} + \beta_{2\text{-}5}\textit{ Firm Characteristics}_{i,(t-1)} \\ & + \beta_{6\text{-}9}\textit{ Director Characteristics}_{i,(t-1)} + \beta_{10\text{-}13}\textit{ Governance Controls}_{i,(t-1)} \\ & + \textit{Year Dummies}_{i,t} + \textit{Industry Dummies}_{i,t} + \varepsilon \end{aligned} \tag{7}$$

$$\begin{aligned} \textit{CFO}_{i,t} = {} & \beta_0 + \beta_1 \textit{ log ID-PCL}_{(t-1)} + \beta_{2\text{-}5}\textit{ Firm Characteristics}_{i,(t-1)} \\ & + \beta_{6\text{-}9}\textit{ Director Characteristics}_{i,(t-1)} + \beta_{10\text{-}13}\textit{ Governance Controls}_{i,(t-1)} \\ & + \textit{Year Dummies}_{i,t} + \textit{Industry Dummies}_{i,t} + \varepsilon \end{aligned} \tag{8}$$

ID-PCL is log of the sum of lobbying expenditure of primary company's independent directors of firm *i* in year *t*. For controlling unobserved factors of a firm, we use a fixed effect model to control the time effect by using year dummies to capture factors which affect all firms at the same time and use 48 Fama industry group to control industry effect to capture all firms in the same industry. All variables are winsorized at 1% and 99% level.
Table 7 reports the results of OLS regression between ID-PCL and three dependent variables of financial performance for robustness test. Each model represents each dependent variable with all control variables. Coefficient of ID-PCL are positive and significant in each model, respectively, indicating that firms with more independent director with primary company that engage in lobbying activities (ID-PCL) have a better financial performance. This is in line with the finding of Chhaochharia and Grinstein (2007) that shows the announcement of SOX rules in 2002 which to fix auditing of US public company (Coates and John (2007) is significantly affect firm value.

Table 7: Robustness Test

Dependent Variables = Cash, Income Before Extra Items, Net Income			
	(1)	(2)	(3)
	CASH	*IBEI*	*NET INCOME*
Constant	2298.3***	-130.9	-137.8
	(2.97)	(-0.28)	(-0.29)
ID-PCL	13.58***	5.791**	6.450**
	(3.23)	(2.29)	(2.48)
$ROA\_1_{(t-1)}$	382.2	2113.3***	2143.2***
	(1.12)	(10.25)	(10.11)
$SIZE_{(t-1)}$	0.0603***	0.0618***	0.0617***
	(51.70)	(89.10)	(86.96)
$DIV_{t-1)}$	-1162.7	6841.0***	7329.5***
	(-0.81)	(7.88)	(8.21)
$LEV_{(t-1)}$	611.6***	-154.8	-183.9
	(3.34)	(-1.41)	(-1.63)
$RET_{(t-1)}$	-39.37	186.6***	197.7***
	(-0.51)	(4.00)	(4.13)
$IND_{(t-1)}$	-52.14	152.3	165.3
	(-0.24)	(1.16)	(1.23)
$AGE_{(t-1)}$	-12.43*	-2.640	-2.510
	(-1.79)	(-0.63)	(-0.58)
$TENU_{(t-1)}$	6.759	5.918	4.746
	(0.80)	(1.17)	(0.92)
$FEM_{(t-1)}$	-87.09	-49.04	-42.77
	(-1.43)	(-1.35)	(-1.14)
$TOI_{(t-1)}$	1939***	387.5**	403.0**
	(6.35)	(2.12)	(2.12)
$IO_{(t-1)}$	-34404***	-702.9	-1765.5
	(-3.51)	(-0.12)	(-0.29)
$eIndex_{(t-1)}$	-139.9***	-89.00***	-91.16***
	(-5.31)	(-5.62)	(-5.61)
$HHI_{t-1)}$	-11283.7***	2725.8	2970.3
	(-2.97)	(1.19)	(1.27)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	9546	9660	9560
Adj R^2	0.3579	0.5271	0.5194

CONCLUSSIONS

This study presents empirical evidence on the firms with the existence of ID-PCL (independent director with primary company engaged in lobbying activities) have a better firm operating performance. Previous literature stated that the existence of independent directors is expected to have roles in monitoring and advising management, we find that the firms with ID-PCL have a better operating performance through their social networking. Furthermore, using cost of debt as the channel, we find that firms with ID-PCL do have lower cost of debt. Existence of ID-PCL in the board formation is the reflection of wider networking and more experience of the independent director on influencing decision makers. This brings bigger opportunities to the appointed firms by having independent directors coming from public listed firms, private company and institutions as well. The results are robust with an alternative measurement of operating performance which are cash from operation, income before extra items and net income.

This paper contributes a different angle of board function as networking source. Networking is reflection of connection and advantageous for the firms while generating financing decisions. Moreover, we show lobbying activities from different point of views. Most of the studies on lobbying activities find direct effect of the lobbying activities itself to other dependent variables while our paper use lobbying activities of the primary company of independent director and see the effect on appointed firm as networking function. This is in line with previous study of Balsmeier et al (2014) that mention independent director from primary company which is innovative firms increase patenting activities. Masulis and Mobbs (2011) mention that foreign independent directors make better cross-border acquisition. Our paper is the first paper on measuring independent director from their primary company's activities that provide economically significant evidence.

REFERENCES

Adams, R. B., & Ferreira, D. (2007). A theory of friendly boards. *The Journal of Finance*,*62*(1), 217-250.

Adams, R.,& Mehran, H. (2003). Board structure and banking firm performance. Federal Reserve Bank of New York Economic Policy Review 9, 123–142.

Agrawal, A., & Knoeber, C. R. (1996). Firm performance and mechanisms to control agency problems between managers and shareholders. *Journal of Financial and Quantitative Analysis*, *31*(3), 377-397.

Agrawal, A., & Knoeber, C. (2001). Do some independent directors play a political role? *Journal of Law and Economics*,14, 179–198.

Ansolabehere, S., Snyder, J., & Ueda, M. (2004), Did Firms Profit from Soft Money?. *Election Law Journal*, Vol. 3, No. 2, pp. 193–8.

Balsmeier, B., Buchwald, A., & Stiebale, J. (2014). Independent directors on the board and innovative firm performance. *Research Policy*, *43*(10), 1800-1815.

Baskin, J. (1989). An empirical investigation of the pecking order hypothesis. *Financial Management*, 26-35.

Baysinger, B., & Hoskisson, R. E. (1990). The composition of boards of directors and strategic control: Effects on corporate strategy. *Academy of Management Review*, *15*(1), 72-87.

Bhagat, S., & Black, B. (1999). The uncertain relationship between board composition and firm performance. *The Business Lawyer*, 921-963.

Bhagat, S., & Black, B. S. (2002). The non-correlation between board independence and long-term firm performance. *As published in Journal of Corporation Law*, *27*, 231-273.

Bonardi, J.-P., Holburn, G. L. F., & Vanden Bergh, R. G. (2006), 'Nonmarket Strategy Performance: Evidence from US Electric Utilities', *Academy of Management Journal*, Vol. 49, No. 6, pp. 1209–28.

Booth, J., & Deli, D. (1996). Factors affecting the number of independent directorships held by CEOs. *Journal of Financial Economics* 40, 81–104.

Borisov, A., Goldman, E., & Gupta, N. (2015). The corporate value of (corrupt) lobbying. *The Review of Financial Studies*, *29*(4), 1039-1071.

Bradley, M., Jarrell, G. A., & Kim, E. H. (1984). On the existence of an optimal capital structure: Theory and evidence. *The Journal of Finance*, *39*(3), 857-878.

Brasher, H. & Lowery, D. (2006), The Corporate Context of Lobbying activities, *Business and Politics*, Vol. 8, No. 1, pp. 1–25.

Cain, M. D., McKeon, S. B., & Solomon, S. D. (2017). Do takeover laws matter? Evidence from five decades of hostile takeovers. *Journal of Financial Economics*, *124*(3), 464-485.

Chen, H., Parsley, D., & Yang, Y. (2015). Corporate lobbying and financial performance. *Journal of Business, Finance & Accounting* 42 (3&4), 444–481.

Chhaochharia, V., & Grinstein, Y. (2007). Corporate governance and firm value: The impact of the 2002 governance rules. *The Journal of Finance* 62 (4): 1789–1825.

Chittenden, F., Hall, G., & Hutchinson, P. (1996). Small firm growth, access to capital markets and financial structure: Review of issues and an empirical investigation. *Small Business Economics*, *8*(1), 59-67.

Chung, K. H. (1993). Asset characteristics and corporate debt policy: an empirical test. *Journal of Business Finance & Accounting*, *20*(1), 83-98.

Coates, I. V., & John, C. (2007). The goals and promise of the Sarbanes-Oxley Act. *Journal of Economic Perspectives*, *21*(1), 91-116.

Cohen, L., Frazzini, A., & Malloy, C. J. (2012). Hiring cheerleaders: Board appointments of "independent" directors. *Management Science*, *58*(6), 1039-1058.

Coles, J. L., Daniel, N. D., & Naveen, L. (2008). Boards: Does one size fit all?. *Journal of Financial Economics*, *87*(2), 329-356.

Cooper, M.J., Gulen, H., Ovtchinnikov, A.V. (2010). Corporate political contributions and stock returns. *The Journal of Finance* 65 (2), 687–724.

Core, J. E., Holthausen, R. W., & Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance1. *Journal of Financial Economics*, *51*(3), 371-406.

de Figueiredo, J. & Silverman, B. (2006). Academic Earmarks and the Returns to Lobbying', *Journal of Law and Economics*, Vol. 49, No. 2, pp. 597–626.

de Figueiredo, J. & Tiller, E. (2001), 'The Structure and Conduct of Lobbying: an Empirical Analysis of Corporate Lobbying at the Federal Communications Commission', *Journal of Economics and Management Strategy*, Vol. 10, No.1, pp. 91–122.

Duchin, R., Matsusaka, J. G., & Ozbas, O. (2010). When are independent directors effective? *Journal of Financial Economics*, *96*(2), 195-214.

Duchin, R., Matsusaka, J.G., & Ozbas, O. (2010). When are independent directors effective? *Journal of Financial Economics* 96, 195–214.

Engel, E., R. Hayes, and X. Wang. (2007). The sarbanes–oxley act and firms' going-private decisions. *Journal of Accounting and Economics* 44 (1-2): 116–145.

Faccio, M. (2006). Politically connected firms. *American Economic Review* 96 (1), 369–386.

Fahlenbrach, R., Low, A., & Stulz, R. M. (2010). Why do firms appoint CEOs as independent directors? *Journal of Financial Economics*, *97*(1), 12-32.

Fama, E. F. & French, K. R. (1997). Industry Costs of Equity, *Journal of Financial Economics*, Vol. 43, No. 2, pp. 153–93.

Fan, J., Wong, T., & Zhang, T. (2007). Politically-connected CEOs, Corporate Governance and Post-IPO Performance of China's Newly Partially Privatized Firms', *Journal of Financial Economics*, Vol. 84, No. 2, pp. 330–57.

Fisman, R. (2001). Estimating the Value of Political Connections, *American Economic Review*, Vol. 91, No. 4, pp. 1095–102.

Fracassi, C., & Tate, G. (2012). External networking and internal firm governance. *The Journal of Finance*, *67*(1), 153-194.

Friend, I., & Lang, L. H. (1988). An empirical test of the impact of managerial self-interest on corporate capital structure. *The Journal of Finance*, *43*(2), 271-281.

Gamba, A., & Triantis, A. (2008). The value of financial flexibility. *The Journal of Finance* 63, 2263–2296.

Goldman, E., Rocholl, J., & So, J. (2013). Politically Connected Boards of Directors and the Allocation of Procurement Contracts, *Review of Finance*, Vol. 17, No. 5, pp. 1617–48.

Goldman, E., Rocholl, J., & So, J. (2008). Do politically connected boards affect firm value? *The Review of Financial Studies*, *22*(6), 2331-2360.

Gompers, P., Ishii, J., & Metrick, A. (2003), 'Corporate Governance and Equity Prices', *Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 107–56.

Graham, J. R., Leary, M. T., & Roberts, M. R. (2015). A century of capital structure: The leveraging of corporate America. *Journal of Financial Economics*, *118*(3), 658-683.

Griner, E. H., & Gordon, L. A. (1995). Internal cash flow, insider ownership, and capital expenditures: a test of the pecking order and managerial hypotheses. *Journal of Business Finance & Accounting*, *22*(2), 179-199.

Hackbarth, D., & Mauer, D. (2012). Optimal priority structure, capital structure, and investment. *Review of Financial Studies* 25, 747–796.

Hadani, M., & Schuler, D.A. (2013). In search of El Dorado: the elusive financial returns on corporate political investments. *Strategic Management Journal* 34 (2), 165–181.

Hansen, C. (2007). Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics* 140 (2): 670–694.

Helland, E., & Sykuta, M. (2004). Regulation and the evolution of corporate boards: Monitoring, advising, or window dressing? *The Journal of Law and Economics*, *47*(1), 167-193.

Hermalin, B. E., & Weisbach, M. S. (1991). The effects of board composition and direct incentives on firm performance. *Financial Management*, 101-112.

Hill, M.D., Kelly, G.W., Lockhart, G.B., & Van Ness, R.A. (2013). Determinants and effects of corporate lobbying. *Financial Management* 42 (4), 931–957.

Hillman, A., Keim, G., & Schuler, D. (2004), 'Corporate Political Activity: a Review and Research Agenda', *Journal of Management*, Vol. 30, No. 6, pp. 837–57.

Hovakimian, A., Opler, T., & Titman, S. (2001). Debt-equity choice. *Journal of Financial and Quantitative Analysis* 36, 1–24.

Hwang, B. H., & Kim, S. (2009). It pays to have friends. *Journal of Financial Economics*, *93*(1), 138-158.

Kester, W. C. (1986). Capital and ownership structure: A comparison of United States and Japanese manufacturing corporations. *Financial Management*, 5-16.

Kim, J. (2008), Corporate Lobbying Revisited, *Business and Politics*, Vol. 10, No. 2, pp. 1–23.

Klein, A. (1998). Firm performance and board committee structure. *The Journal of Law and Economics*, *41*(1), 275-304.

Knyazeva, A., Knyazeva, D., & Masulis, R. (2013). The supply of corporate directors and board independence. *Review of Financial Studies* 26,1561–1605.

Lang, L., Ofek, E., & Stulz, R. (1996). Leverage, investment, and firm growth. *Journal of Financial Economics*, *40*(1), 3-29.

Litvak, K. (2007). The effect of the sarbanes-oxley act on non-US companies cross-listed in the US. *Journal of Corporate Finance* 13 (2-3): 195–228.

Long, M. S., & Malitz, I. B. (1985). Investment patterns and financial leverage. In *Corporate capital structures in the United States* (pp. 325-352). University of Chicago Press.

MacAvoy, P. W., Cantor, S., Dana, J., & Peck, S. (1983). ALI proposals for increased control of the corporation by the board of directors: An economic analysis. *Statement of the Business Roundtable on the American Law Institute's Proposed "Principles of Corporate Governance and Structure: Restatement and Recommendation." New York: Business Roundtable*.

Marsh, P. (1982). The choice between equity and debt: An empirical study. *The Journal of Finance*, *37*(1), 121-144.

Masulis, R. W., Wang, C., & Xie, F. (2012). Globalizing the boardroom—The effects of foreign directors on corporate governance and firm performance. *Journal of Accounting and Economics*, *53*(3), 527-554.

Masulis, R., & Mobbs, S. (2013). Reputation incentives of independent directors: impacts on board monitoring and adverse corporate actions. Unpublished working paper University of New South Wales, Sydney, Australia.

Mayda, A., Ludema, R., & Mishra, P. (2010), Protection for Free? *The Political Economy of US Tariff Suspensions*, IMF Working Paper 10/211.

Myers, S. C., & Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, *13*(2), 187-221.

Nguyen, B. D., & Nielsen, K. M. (2010). The value of independent directors: Evidence from sudden deaths. *Journal of Financial Economics*, *98*(3), 550-567.

Pfeffer, J. (1972). Size and composition of corporate boards of directors: the organization and its environment. *Administrative Science Quarterly* 17 218–229.

Pinches, G. E., Mingo, K. A., & Caruthers, J. K. (1973). The stability of financial patterns in industrial organizations. *The Journal of Finance*, *28*(2), 389-396.

Rajan, R. G., & Zingales, L. (1995). What do we know about capital structure? Some evidence from international data. *The Journal of Finance*, *50*(5), 1421-1460.

Richter, B. K., K. Samphantharak & J. F. Timmons (2009), Lobbying and Taxes, *American Journal of Political Science*, Vol. 53, No. 4, pp. 893–909.

Roberts, M., & Whited, T. (2013). *Endogeneity in empirical corporate finance1*. In: Handbook of the Economics of Finance, edited by George M, Constantinides M, R.M.S. Chapter 7, 493–572. Elsevier.

Rosenstein, S., & Wyatt, J. G. (1990). Independent directors, board independence, and shareholder wealth. *Journal of Financial Economics*, *26*(2), 175-191.

Ryan, Harley E., Jr., & Wiggins III, R.A. (2004). Who Is in Whose Pocket? Director Compensation, Board Independence, and Barriers to Effective Monitoring. *Journal of Financial Economics*, 73(3): 497–524.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regressions," In: *identification and inference for econometric models: A festschrift in honor of Thomas Rothenberg,* edited by Stock, J., and D. Andrews. Cambridge, MA: Cambridge University Press.

Strebulaev, I. A., & Yang, B. (2013). The mystery of zero-leverage firms. *Journal of Financial Economics*, *109*(1), 1-23.

Titman, S., & Wessels, R. (1988). The determinants of capital structure choice. *The Journal of Finance*, *43*(1), 1-19.

Walsh, E. J., & Ryan, J. (1997). Agency and tax explanations of security issuance decisions. *Journal of Business Finance & Accounting*, *24*(7-8), 943-961.

Wooldridge, J. (2015). *Introductory econometrics: A modern approach.* Mason, OH: Nelson Education

APPENDIX

A. Variable Definition

Variable	Definition	Data Source
ID-PCL	Lobbying expense measured by natural logarithm of dollar amount of lobbying expense of independent director primary company	CRP lobbying data
ROA	Measured by net income divided by total assets (ni/at)	Compustat
ROE	Measured by net income divided by total equity (ni/ceq)	Compustat
NPM	Measured by net income divided by total sales (ni/sale)	Compustat
COD	Measured by interest expense of long term debt/ total debt (intl/tdebt)	Compustat
Control Variables:		
Firm Characteristics		
SIZE	Measured by total sales (sales)	Compustat
DIV	Measured by cash dividend (data item 127)	Compustat
LEV	Measured by long term debt divided by total debt (dlt/tdebt)	Compustat
RET	Measured by buy hold return	CRSP
Director Characteristics		
IND	Measured by number of independent director divided by total number of boards	Risk Metrics
AGE	Age of the independent director	Risk Metrics
TENU	Measured by number of year independent director has been working in the company	Risk Metrics
FEM	Dummy variable which equals to 1 if at least 1 female director listed, 0 otherwise	Risk Metrics
Governance Characteristics		
TOI	Takeover likelihood	Cain et al
IO	Average of institutional ownership	Thomson Reuter 13F
eIndex	Sum of (cboard,labylw,lachtr,supermajor,GoldenParachute,ppill)	Risk Metrics
HHI	Sum of square market share	Compustat

DECISION SCIENCES INSTITUTE

Corporate Social Network and Tax Aggressive

Junmao Chiu
Yuan Ze University
jchiu@saturn.yzu.edu.tw

Le Van Hung
Yuan Ze University
lehungtcnh@gmail.com

Yu-Shan Wei
National Yunlin University of Science and Technology
wus@yuntech.edu.tw

Chih-Liang Liu
National Yunlin University of Science and Technology
chlliu@yuntech.edu.tw

ABSTRACT

This study highlights the role that social networks play in a corporate tax aggressive. We construct a comprehensive social network using data on Taiwan-listed companies that arises from education background, prior experience, and other social activities such as training courses, business forums and meetings of CEO and CFO during the period from 2010 to 2014. We find that the social networks of CEO, CFO help reduce (increase) corporate tax avoidance. Those findings suggesting that the information on a range of tax avoidance strategies is share among firms by the corporate network ties of CEO and CFO.

KEYWORDS: Corporate Governance, Corporate Social Network, Corporate Tax Aggressive

INTRODUCTION

Corporate tax avoidance is an important research topic in both finance and management, attracted more attention of scholars. Prior literatures show the benefits of tax aggressive on firm earning target, reduce tax burden, increase cash flow (Hanlon, 2005), reduce cost of equity capital (Chi et al., 2014). Numerous studies also pointed out the potential penalty costs of implement tax avoidance (Mills et al., 2013), damage to firms and manager reputation(Chen et al., 2010) and how to firms exercise tax evasion by adopting tax strategies, tax planning (Frank et al., 2009).

At the same time, social capital has been a subject of attractive topic not only economic but also in other science fields. Social networks play as an important role in all social and economic transactions (Butler & Gurun, 2012). Social network impact to firm behavior because they promote the diffusion of new ideas and allow network firms to study vicariously from each other's experience (Bizjak et al., 2009; Fracassi & Tate, 2012). The pros and cons of social topic that received significant attention in academic research. One line of research explores the benefits of social network to their firms. From the perspective of *Information and Resource Hypothesis*, social connection ties provided more helpful information that cannot be acquired from other sources to corporate decision making (L. Cohen et al., 2008; Horton et al., 2012;

Renneboog & Zhao, 2011). In addition, *Quality Hypothesis* suggests that CEOs' external social networks can be used to assess CEO quality (Butler & Gurun, 2012). *Reputation Hypothesis* explored that greater external social networks represent individual personal quality, talent, past success and reputable figures (Renneboog & Zhao, 2011).

A numerous of literatures show that CEO networks do not constitute valuable resources, further resulting in several negative outcomes. From the perspective of *Power Hypothesis*, CEOs can use social network in order to gain more power over the board and acquire self-dealing compensation arrangements beneficial to themselves (Belliveau et al., 1996). Subrahmanyam (2008) shows that a CEO's pre-existing network can harm governance practices. Fich and Shivdasani (2006), Malmendier and Tate (2009) suggest that executive social networks result in less time to monitor.

Based on this emerging line of research to elicit the value of CEO and CFO social networks, we explore the role of corporate social networks in corporate tax avoidance. We assume that the relationship between social networks and tax avoidance has not received much studied from scholar, partly due to the limit of data and measurement difficulties. We explore two research questions: (1) whether the external social networks influence corporate tax avoidance; (2) whether the effects of social networks on tax aggressive vary between CEOs and CFOs.

To examine the tax aggressiveness of firms, we adopt on multiple measures of tax aggressiveness drawn from the literature. Specifically, we use two effective tax rate measures and two book-tax difference measures: long-term effective tax rate (defined as total tax expense divided by pre-tax income), long-term cash effective tax rate (defined as total refund income tax divided by pre-tax income), the permanent book-tax difference measure advanced by Lennox, et al. (2013). Firms that are more tax aggressive have lower effective tax rates (ETRs) and higher permanent book-tax differences. We follow Freeman (1978) to construct two main variables to measures the network ties of CEO and CFO by direct connections to assess actual influence (degree) and indirect connection measures to capture information collection ability (closeness). Our sample is constructed a comprehensive social network of Taiwan listed companies that arises from employment, education background, prior experiences, and other social activities such as training courses, business forums, and meetings. These social ties are tracked over span 5 years, from 2010 to 2014. Our final sample firm for the main test contains more 6,700 firm year observation for 1115 firms.

We find that firms with higher CEO social network have a significant higher effective tax rate (ETRs) and lower permanent book tax difference, while our results show the positive relation between CFO and tax avoidance for all three our tax measure. Next, we use the instrumental regression to reduce endogeneity problems. We find statistically and economically significant evidence that firm have lower (higher) tax avoidance if theirs CEOs (CFOs) are social connected, supporting for our main results. In addition, we also perform robustness tests by using alternative measure tax avoidance consist of effective tax rate and effective cash rate, and alternative social network includes Betweenness and Eigenvector, and Tobit regression. Our main results hold under all specifications. Finally, we examine the interaction between CEOs (CFOs) personal traits such as powerful and background and their network ties effect on tax aggressive [See Appendix A for variables definition]. We find the statistic empirical evidence supported for our main results.

Our study adds several findings to the extant literature on the ways in which social network affects tax aggressive. Firstly, unlike (Coff, 2002), who explore the effect of abilities and knowledge of CEO and CFO on firm performance and financial report, our study and the work of (Hanlon & Heitzman, 2010) is closely related, we focus on the directly the ties between tax avoidance and social network. To our best knowledge, we are the first in the literature to document evidence both CEOs and CFOs have impact to corporate tax aggressive by constructing a comprehensive social network from their employment, education and other experience. Second, our research adds more evidence about the impact of social network in

finance and accounting to firm decision, especially on firm's tax related decisions. Finally, we also show the consistent empirical that the interaction between CEO's and CFO's background and their social network significant impact on tax avoidance.
The remainder of this paper is organized as follows. Section 2 reviews related literature and develop our hypothesis. Section 3 provides a description of the data and the research methodology. Section 4 presents and analyzes empirical results. Finally, section 5 offers the conclusion drawn from this study.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Tax Aggressive

Taxes represent a significant cost to the firm and shareholder. Tax savings often increase after-tax income, cash flow and shareholder wealth. Therefore, all firms and shareholders generally expected minimize corporate tax payments through tax aggressive activities. Phillips et al. (2003), Hanlon (2005) find that tax avoidance is an effective income management tool for a firm to achieve their earning target, while also reducing tax burden and increasing cash flow. Tax avoidance is valued by shareholders so that firms are willing to participate in tax planning if it is considered as value-increasing activities. Desai and Dharmapala (2006) pointed out the positive relation between tax avoidance and firm value. Dyreng et al. (2010) recognize that firms adopting various tax strategies (even possibly illegal) to reduce taxes. Further, the reduction in a firm's effective tax rate provided by tax avoidance is potentially a positive signal to investors, thereby reducing the cost of equity capital (Chi et al., 2014; Inger, 2014; McGuire et al., 2014). Firms can reduce their tax burdens through four primary ways: multinationality, tax havens, intangible assets, and transfer pricing aggressiveness. Firms can reduce their tax burdens by "offshore sheltering", it allows firms to syphon their profits to foreign subsidiaries where the tax rate is lower. For example, National Public Radio disclosed that companies such as Pfizer, Microsoft, and Google were taking advantage of offshore tax havens. Companies also may create complex structures involving tax-indifferent related parties to avoid detection by auditors and tax authorities.
Shackelford and Shevlin (2001) find the unbalance between aggressive financial reporting and tax avoidance. Frank et al. (2009) suggested that firms might use tax planning, aggressive tax reporting for their tax avoidance. Phillips et al. (2003) show that adopt compensation policy based on after tax -performance leads to less effective tax rates. Agree with point, Rego and Wilson (2009) pointed that a positive tie between compensation and aggressive tax reporting, they find no relationship between this aggressiveness impact to firm performance in the future or is as a result of inefficient management. Rego and Wilson (2012) point out those companies provide incentive compensation to engage in tax aggressive activities. In addition, Dyreng et al. (2008) finds that the quality of corporate governance, institutional ownership, debt policy and international operations are also related to firm's tax avoidance activities. Minnick and Noga (2010) examine corporate governance's role in a firm's choice of tax strategies. Governance plays important role for choosing long-term tax management. Company with different governance structure may be purse difference type’s tax management. Chen et al. (2010) find that ownership structure can also associated with aggressive tax, family owners are less tax aggressive than no-family firms to avoid other cost and potential penalties.
Although, the most obvious benefit of tax avoidance is maximizing firm value, tax saving is greater but one of the important cost of tax aggressiveness is the potential penalty (Mills et al., 2013), negative publicity (Lisowsky, 2009; Wilson, 2009); be labeled as “poor corporate citizens” (Hanlon & Slemrod, 2009) and potential reputation damage to firms and their managers (Chen et al., 2010; Graham et al., 2014; Rego & Wilson, 2009). According to recently survey by Graham et al. (2014), 69% of tax executives agree that potential harm to their firms’ reputations

when they make a decision about tax planning strategies. Managers implementing tax avoidance strategies are also personally exposed to the risk of penalties and fines, and damage to their reputations. Moreover, Austin and Wilson (2015) argue that another reason firms refrain from tax aggressive is the associate with loss of reputation with customers. They show the evident that reputable and valuable-brand firm names companies have higher effective tax rates. Also, in poorly managed companies, tax avoidance secrets may be exploited to obscure rent extraction by managers, resulting in future negative abnormal profits (Desai & Dharmapala, 2006). In addition, tax avoidance was used to hide the accumulation of bad news by managers, leading to the risk of stock price crash (J.-B. Kim et al., 2011)

Effects of Social Network

Social capital is an important research topic in both finance and management, attracted more attention of scholars. Social capital is the value of reputation, resources, and social actions embedded in relationships among people (Adler, 2001; Adler & Kwon, 2002; Carrie R. Leana & Buren, 1999; Lin, 2001). A social network is the media through which social capital is created, maintained and used. Thus, firms recently pay more attention to invest for social capital to get the benefits from acquiring richer resources. Although, the ties between social capital and business performance has received a numerous previous studies but detailed investigate about the role and influence of the establishment and maintenance of top executives' external networks on managerial effectiveness and strategic leadership is still limited.
External social networks refer to the relationships of executive with individuals outside the firm (Cao et al., 2006). Engelberg et al. (2012) pointed out that social network are established through education background from the same highest ranked programs. Collins and Clark (2003) argued executives' network are created when they have a larger set outside ties, such as external boards and managerial services, customers or competitors, alliance partners, government agencies, and trade associations. External networks play as an important role in all social and economic transactions, are often regarded as a valuable organizational resource and convey important strategic value to the company (Butler & Gurun, 2012; Engelberg et al., 2012). Social network theory suggests that social capital benefits not only through the ties themselves, but also their connection structures (Burt, 1997). Social network impact to corporate behavior because they promote the diffusion of new ideas and allow networked firms to study vicariously from each other's experiences (Bizjak et al., 2009; Fracassi & Tate, 2012). Adler and Kwon (2002) find that the impact of social network on firm behavior may depend on characteristics of both firm's relations to other firms and the kind of knowledge being share across those relations. Next, the benefits and potential drawbacks of external networks are discussed. From a positive perspective, social networks connection as facilitating information flows between individuals or companies. Many network advantages are from information collection. First, social ties help firms access to a broader source of information at a lower cost and improve information quality, correctness, and levels of specific details and timeliness (Adler & Kwon, 2002; Horton et al., 2012). Second, social ties help firms access information and knowledge from external parties thence combine acquired information with their own existing knowledge to explore new learning opportunities and even to formulate new competitive strategies (Grant, 1996; Kogut & Zander, 1992). Third, networks may enable firms to develop more effective corporate strategies from accessing valuable information, firms may formulate and implement useful competitive strategies (Ahuja & Katila, 2004; Athanassiou & Nigh, 2002; Geletkanycz & Hambrick, 1997). In addition, external social connections improve the sharing of information thus reduce the possibility of deception, error, and misconduct and allow networked firms to learn vicariously from each other's experiences (Baker & Faulkner, 2004; Fracassi, 2017; Kilduff & Tsai, 2003) argued that managers are influenced when making investment decision and other corporate finance decisions by their social partner.

On the other side, socially influential people may abuse their reputation or power to hide their misconduct lead to ineffective manage that may be detrimental to the firm and shareholders (Jack Dorminey et al., 2012). Bebchuk and Fried (2006) find an empirical that executives using their power to impact board and compensation to get excessive compensation arrangements are not necessarily related to performance. The social ties of executives and directors impose greater information asymmetry in equity markets. L. Cohen et al. (2008) provide empirical evidence that social networks alter the increasing probability of informed trades in a company's stock. Hwang and Kim (2009) show that companies with board members who are personally connected to the CEO have higher CEO compensation, lower pay-performance sensitivity, and lower turnover-performance sensitivity. Chidambaran et al. (2011) find that the abilities of fraud are higher when CEO have the connections outside.

Recent studies argued the ties among social networks and tax strategies and tax policy decision. Hanlon and Heitzman (2010) studies the impact of directors networks and their social connections to other firms on tax-avoidance behavior. They strongly find empirical evidence that an interlocked board member was influential in spreading the use of an installment sales tax shelter. Brown (2011) finds evidence that network ties via board interlocks influence the discrete tax policy decision at the focal firm adopting the corporate-owned life insurance tax shelter. Larcker et al. (2013) provide evidence that board of director connections have an economic impact on the firm performance; better-connected firms earn significantly higher risk-adjusted returns and experience higher gains in profitability. In addition, Brown and Drake (2014) network ties created by executive director impact on tax avoidance

The Role and Influence of CEOs and CFOs

The role of managers in the firm's activities is topic to debate among researcher. According to (M. D. Cohen et al., 1972), Hannan and Freeman (1977) pointed out that the success of an organization is determined largely by its quality of products, core competency, life cycle, and to some extent even luck, and not by CEO ability. On the contrary, the upper echelon theory assumed that top executives make a significant contribution to firm performance, the organization's results reflect the values and abilities of top managers (Hambrick & Mason, 1982; Harris & Helfat, 1997; Hayes & Schaefer, 2000). Agree with this view, the human capital theory proposes that the knowledge and abilities of top managers can be an important determinant of firm performance ((Amit & Schoemaker, 1993; Becker, 1962; Coff, 2002).

Prior research has focused on analyzing the impact of CEOs on firm performance rather than CFO's. Scholars assume that CFOs are CEO agents (Graham & Harvey, 2001) and a CEO has the power to replace a CFO who does not follow the CEO's preferences. Agree with this view, Mian (2001) assumed that CEOs who occupy higher positions of power in the firms than to CFOs, as a result have the ability to impact the compensation, job security, future career opportunities, and financial reporting decisions of their CFOs (Feng et al., 2011; Krishnan et al., 2011). However, the important role of CFO also emphasized in prior literatures. CFO is considered as the second most important financial manager of the company after CEO in the enterprise hierarchy. CFOs are typically in charge of financial planning, budgeting, internal control and financial reporting (Gore et al., 2007; Kaufman, 2003), which helps them to have a significant understanding of firm weath. CFOs often supervise the process of making financial statements and are considered as supervisory agencies for financial reporting quality. Moreover, compared to other executives, CFOs is in a unique position to perform accounting operations, such as structuring transactions, selecting an inappropriate accounting method or implementing false journal entries. Prior studies have found that CFOs have a significant impact on firm's financial reporting (GE et al., 2011; Geiger & North, 2006). Although CFOs may generate a significant impact on the firm's financial decisions, they do not act alone if without CEO's guidelines and policy support. As the business partner of CEO, CFO creates the decision power to deliver desirable business results by collaborating with CEOs related in companies' financial decisions (Jiang et al., 2010). Agree with this views, Feng et al. (2011) argue that CEOs use their power to promote shareholder goals or put pressure on CFOs to manipulate reporting systems and excessive performance. He also argued that CEO pressure is the main driver of CFO regarding accounting manipulation. A recently survey of 141 CFO by public companies find that that 17% of CFO reports were pressured to misjudge their results by their company's CEOs during the past 5 years [More detail please refer at: The fear of all sums, CFO magazine. URL: ⟨http://www3.cfo.com/article.cfm/3005788⟩ (accessed 10.06.13)]. Dichev et al. (2013) surveyed 169 CFOs, of which 91% reported that pressure as a motivation to manage earnings. Therefore, the Securities and Exchange Commission's (SEC) recent regulations (e.g., Section 906 of the Sarbanes-Oxley Act, effective July 30, 2002; Section 302 of the Sarbanes-Oxley Act, effective August 29, 2002; The SEC June 27 Order in 2002) require both CEOs and CFOs to take responsibility for corporate financial reports, reflecting an understanding of the CEO and CFO collaboration on financial reporting and policy decisions.

Hypothesis Development

Prior scholars most research about CEO specific heterogeneity such as education, personal characteristics, or personality traits, but less study researches about the social network of CEOs, also CFO. The benefit and cons of personal connections are discussed by prior research. Personal connection leads to better analyst performance (L. Cohen et al., 2008), investment allocation efficiency (Duchin & Sosyura, 2013), more efficient mergers and acquisitions (M&A) decisions (Cai & Sevilir, 2012). On the other hand, personal ties also have some cons such as leading to deteriorate corporate governance and monitoring of managers, to lower pay-performance and turnover-performance sensitivities (Fracassi & Tate, 2012; Hwang & Kim, 2009).

We acknowledge that deeply research on the both CFO's and CEO's social network influence to firm performance is less studied. Our research question considers that whether CEOs and CFO who are most important top executive in firms use their personal connect to impact firm's performance through by tax avoidance activities. It tests following hypothesis: CEO's, CFO's social network is positive (negative) associated with corporate tax aggressive.

DATA SOURCE AND RESEARCH METHODS

For exploring the value of CEO's and CFO's corporate social network and its effect on tax aggressive, this study adopts regression models in which the proxies of CEOs' and CFOs' social network are regressors. The following introduces the research methodologies, including dataset, variable candidates, and the regression models.

Data Source and Sample Selection

Our sample selection starts with all Taiwan-listed firm on the Taiwan Economic Journal database from 2010 (tax reforms) to 2014 fiscal years. The initial sample consists of 10,914 firm-year observations. We first conduct screen the data to minimize observations with errors and to obtain the necessary data for empirical analysis. We eliminate foreign registrants and firms designated as real estate investment trusts (REITs) because these firms are subject to different tax rules. We continue delete observations that do not have sufficient data for calculating the at least one of the three measures of tax avoidance: long term effective tax rate (ETR1), long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD). Then, we truncated our long term effective tax rate (ETR1), long term cash effective tax rate (ETR2) to the range from 0 to 1.

To calculate the effects of CEOs', CFOs' corporate social networks, our study also used the Taiwan Economic Journal database which collects detailed information on the identity of all CFO, CEO from proxy statements, annual reports, and corporate websites. These CFO and CEO' profiles include information on: (1) working experiences; (2) educational background; and (3) training courses and social meetings. We further eliminated firm-year observation with any missing values to calculate the CEO and CEO social network measures.
We then retain firm-year observations for which we have data available for our control variables (defined below). Finally, we retain firm-years for which we have data available for our governance and incentives variables. As a result, sample sizes for the ETR1, ETR2, and BTD measures are 5,699, 5,706 and 5,810 firm-year observations, respectively. Note that for some of the tests, there are fewer observations due to additional data limitations.

Measures of Social Network

Social networks are a broad construct, including numerous dimensions, and no single indicator can effectively and completely capture all dimensions of the concept of a social network. We define three ties representing the dimension of social networks as follows:

- *We identify the formal tie if CFO and CEO who share current and past employment positions outside the firms or other industrial associations.*
- *We identify the informal tie if CFO and CEO who went to the same educational background or memberships in the same school alumni association.*
- *Other social network ties are identified if CFO and CEO are required to attend training courses or conference meetings held by the authorities, and these training/meeting records should be disclosed in an annual report.*

We follow the Fracassi and Tate (2012) and used other dimension of social network, including educational platform, previous experience, training course and social meeting, from detailed primary information on CFO and CEO to construct a social network index with the degree centrality of the CEOs, CFOs. Centrality measures the overall embeddedness of a personal company in the network that includes of all firms. In the context of this research, centrality is amount of vertices that an CFO or CEO is tied to. It pointed out the level that a CFO or CEO takes a strategy position in their network (Wasserman & Faust, 1994). We follow Freeman (1978), Renneboog and Zhao (2011) and El-Khatib et al. (2015) to construct different social network measurement consist of closeness, degree, betweenees and eigenvector.

(1) *Degree* is the sum of all links that a CFO or CEO has with other individuals in the network divided by the total number of companies in the network. Brass and Burkhardt (1993) assumed that the greater the degree of centrality is, the more prestige CFO or CEO has within the network, and well-connected executives are likely to be more visible.
(2) *Closeness* is the inverse of the sum of the (shortest) distances between a director or executive and all other individuals in a network. It indicates how efficiently the director or executive can obtain information from other individuals in the network.
(3) *Betweeness* measures how often the director or executive lies on the shortest path between any other two individuals of the network.
(4) *Eigenvector* is a measure of the importance of the director or executive in the network. It takes into account the extent to which the director or executive is connected with external highly connected directors or executives.

In this study, we adopt two corporate social network measures based on degree centrality which measures direct (local) connections to assess actual influence; and closeness which indirect connection measures to capture information collection ability (El-Khatib et al., 2015; Renneboog & Zhao, 2011), we then use *betweeness* and *Eigenvector* for our robustness test. In our dataset, we set CEO's and CFO's social network as an indicator variables coded as one if the value is higher than mean of it by each year, zero otherwise.

Measures of Tax Aggressiveness

We follow Chen et al. (2010), Dyreng et al. (2008) to adopt two tax-avoidance measures based on a long term effective tax rate (ETR1) and a long term cash effective tax rate (ETR2). We use a measurement period of three years to avoid significant year-to-year variations on a single-year tax avoidance measure, and also to alleviate concern about potential survivorship bias associated with the use of horizons longer than three years. ETR is set as missing when the denominator is zero or negative. We truncate ETR to the range [0,1]. As expressed in Eq. (1), the first measure, ETR1, is the ratio of total tax expense over the most recent three years scaled by total pre-tax income net over the same period. It discovers the company's tax evasion through permanent book-tax differences, such as investing in municipal bonds and participating in tax shelters.

$$ETR1_{i,t} = \frac{\sum_{t=1}^{t=t+2} \text{Total income tax expense}_{i,t}}{\sum_{t=1}^{t=t+2} \text{Pre\_tax book income}_{1,t}} \qquad (1)$$

As shown in Eq. (2), our second measure, ETR2, calculates long term cash effective tax rate by using the total refund income tax to chief financial officer over last three years scaled by total pre-tax income over the same period of time. The reason for our selection this tax measure is: firstly, cash effective tax rates take into account the tax benefits of employee stock options, whereas GAAP effective tax rates do not; secondly, cash effective tax rates are not affected by changes in accounting estimates such as the valuation allowance or tax contingency reserve (Dyreng et al., 2008). We acknowledge that this measure is better matching between taxes paid and the income to which these taxes relate, it also considers both permanent and temporary book-tax differences.

$$ETR2_{i,t} = \frac{\sum_{t=1}^{t=t+2} \text{Refund income tax to chief financial officer}_{i,t}}{\sum_{t=1}^{t=t+2} \text{Pre\_tax book income}_{1,t}} \qquad (2)$$

Many prior research also used a variety of book-tax difference measures for tax avoidance, where book-tax difference is defined as the difference between income reported to the capital market and that reported to the tax authorities. Large book-tax differences can also reflect for accrual manipulation (or a combination of tax avoidance and accrual manipulation), or simply represent for the differences in financial and tax accounting rules. Thus, book-tax difference measures are likely to be appropriate for our research question. We follow Frank et al. (2009), Lennox et al. (2013) to adopt other tax-avoidance measures based on permanent book-tax differences (BTD) which is specifically design to capture the effect due to aggressive tax planning activities. As expressed in Eq. (3), the third measure, BTD is define as:

$$BTD_{i,t} = \frac{\text{Pre\_tax book income}_{i,t} - \left(\frac{\text{Total income tax expense}_{i,t}}{Law\ tax\ rate_{i,t}}\right)}{\text{Total income tax expense}_{i,t}/law\ \text{tax rate}_{i,t}} \qquad (3)$$

For robustness check, we follow Lennox et al. (2013), (Chen et al., 2010); Dyreng et al. (2010) to construct effective tax rate (ETR3) and Cash effective tax rate (ETR4). Details can be found in the Appendix A.

Basic Statistics

Table 1 provides the results of our sample composition. Panel A presents for the ETR1 group, ETR2 group in Panel B, and BTD group in Panel C. As shown, the means of there our tax aggressive measure variables are 0.148, 0.127 and 2.593 with a standard derivation of 0.118, 0.121 and 4.840 for ETR1, ETR2 and BTD, respectively. These values are similar to those reported in Dyreng et al. (2008) and Dyreng et al. (2010), suggesting that our sample representativeness is reasonable. The means of CEOs range from 0.529 to 0.638, while the means of CFOs range from 0.836 to 0.888. The means of SIZE, MB, LEVER variables range from 15.018 to 15.04, from 1.991 to 1.988, from 0.482 to 0.484, from 15.018 to 15.042, respective. ESUB, PPE, INTAN and DFI has the mean value approximate 1.013, 0.299, 0.017, 0.475, respectively.

Table 1: Descriptive statistics

Variables	Mean	Median	SD	Min.	Max.
Panel A: DV=ETR1 (N=5699)					
ETR1	0.148	0.156	0.118	0.001	0.650
CEO Closeness	0.638	1.000	0.481	0.000	1.000
CFO Closeness	0.887	1.000	0.317	0.000	1.000
CEO Degree	0.593	1.000	0.491	0.000	1.000
CFO Degree	0.837	1.000	0.370	0.000	1.000
SIZE	15.042	14.878	1.474	10.612	21.730
MB	1.991	1.401	5.406	0.067	295.624
LEVER	0.482	0.424	1.269	0.004	64.473
NOL	0.726	1.000	0.446	0.000	1.000
ΔNOL	-0.001	0.000	0.006	-0.094	0.249
ROA	8.763	8.490	10.510	-76.120	96.450
ESUB	1.013	0.840	0.891	0.000	27.479
PPE	0.299	0.276	0.219	0.000	2.877
INTAN	0.017	0.003	0.050	0.000	1.089
DFI	0.475	0.296	0.655	0.000	15.153
RD	8.262	10.431	5.107	0.000	17.689
Panel B: DV=ETR2 (N=5706)					
ETR2	0.127	0.118	0.121	0.000	0.685
CEO Closeness	0.636	1.000	0.481	0.000	1.000
CFO Closeness	0.887	1.000	0.316	0.000	1.000
CEO Degree	0.591	1.000	0.492	0.000	1.000
CFO Degree	0.836	1.000	0.370	0.000	1.000
SIZE	15.035	14.871	1.476	10.612	21.730
MB	1.998	1.402	5.414	0.067	295.624
LEVER	0.484	0.424	1.281	0.004	64.473
NOL	0.725	1.000	0.446	0.000	1.000
ΔNOL	-0.001	0.000	0.006	-0.094	0.249
ROA	8.772	8.515	10.513	-76.120	96.450
ESUB	1.013	0.839	0.903	0.000	27.479
PPE	0.298	0.276	0.218	0.000	2.877
INTAN	0.017	0.003	0.049	0.000	1.089
DFI	0.474	0.297	0.651	0.000	15.153
RD	8.272	10.435	5.100	0.000	17.689
Panel C: DV=BTD (N=5810)					
BTD	2.593	0.000	4.840	0.000	18.630
CEO Closeness	0.529	1.000	0.499	0.000	1.000
CFO Closeness	0.888	1.000	0.317	0.000	1.000
CEO Degree	0.530	1.000	0.499	0.000	1.000
CFO Degree	0.816	1.000	0.387	0.000	1.000
SIZE	15.018	14.862	1.466	10.346	21.730

MB	1.983	1.392	5.368	0.067	295.624
LEVER	0.483	0.425	1.269	0.004	64.473
NOL	0.728	1.000	0.445	0.000	1.000
ΔNOL	-0.001	0.000	0.006	-0.094	0.249
ROA	8.631	8.285	10.425	-76.120	96.450
ESUB	1.013	0.839	0.900	0.000	27.479
PPE	0.298	0.275	0.218	0.000	2.877
INTAN	0.017	0.003	0.049	0.000	1.089
DFI	0.477	0.299	0.653	0.000	15.153
RD	8.315	10.451	5.079	0.000	17.689

Notes: This table provides the descriptive statistics for the dependent and control variables, with the data covering the period of January 1, 2010 to December 31, 2014. All variables are defined in Appendix A.

Table 2 provides the correlation results. Our results show the positive correlation between CEO's social network (CEO closeness, CEO degree) and two effective tax measures (ETR1, and ETR2), and negative correlation with permanent book tax difference (BTD). These empirical results indicate that higher corporate social network from CEO lead to lower tax aggressive. Table 2 also findings a significantly negative correlation between the CFO's social network (CFO closeness and CFO degree) and ETR1, ETR2, indicating that higher social network from CFO positive impact tax aggressive. Our results showed the significantly negative correlation between ETR1, ETR2 and BTD3. In addition, the correlations among CEO's, CFO's social network variables (CEO closeness, CEO degree, CFO closeness, CFO degree) are all significantly. These results provide us with a first glance at the relationship between CEO's, CFO's social network and tax aggressive prior to the regression analysis.

Table 2: Pearson Correlation Table

Variables	ETR1	ETR2	BTD	CEO Closeness	CFO Closeness	CEO Degree	CFO Degree	SIZE	MB	LEVER	NOL	ΔNOL	ROA	ESUB	PPE	INTAN	DFI	RD
ETR1	1.000																	
ETR2	0.705*	1.000																
BTD	-0.197*	-0.163*	1.000															
CEO Closeness	0.098*	0.107*	-0.054*	1.000														
CFO Closeness	-0.056*	-0.059*	0.028	-0.088*	1.000													
CEO Degree	0.087*	0.095*	-0.022*	0.719*	-0.016*	1.000												
CFO Degree	-0.031*	-0.036*	0.020	0.097*	0.856*	0.093*	1.000											
SIZE	0.117*	0.146*	0.104*	0.286*	-0.062*	0.320*	0.013	1.000										
MB	-0.048	-0.040*	0.064*	-0.005	0.004	-0.009	0.015	0.057*	1.000									
LEVER	0.015	0.002	0.013	0.012	-0.017	0.013	-0.008	0.026	0.001	1.000								
NOL	0.145*	0.118*	-0.050*	0.110*	0.023***	0.076*	0.026***	0.093*	-0.074	-0.008	1.000							
ΔNOL	-0.009	0.024**	0.024*	0.019	-0.034*	0.032**	-0.034***	0.052*	0.023	0.009	0.158*	1.000						
ROA	0.267*	0.231*	0.141*	0.067*	-0.052*	0.060*	-0.039*	0.349*	-0.048	-0.012	0.100*	0.113*	1.000					
ESUB	0.193*	0.157*	-0.037*	0.102	-0.033**	0.077*	-0.018	0.023***	0.002	0.245*	0.075*	0.036*	0.205*	1.000				
PPE	-0.003	-0.002	-0.076*	0.018	-0.043*	0.010	-0.052*	0.116*	-0.048	0.099*	0.025***	0.054*	0.133*	-0.151*	1.000			
INTAN	-0.001	-0.011	-0.024***	0.049*	-0.029**	0.061*	-0.006	0.096*	0.048	0.041*	0.006	0.005	0.009	-0.012	-0.042*	1.000		
DFI	0.116*	0.102*	-0.029**	0.075*	-0.053*	0.056*	-0.029**	0.016	-0.001	0.020	0.033	0.007	0.126*	0.592*	-0.139*	-0.028**	1.000	
RD	0.086*	0.051*	-0.063*	0.162*	0.015	0.161*	0.084*	0.150*	0.001	-0.047*	0.130*	-0.031**	0.062*	0.012	0.026**	0.019	0.214*	1.000

Notes: The correlation statistics are provided for the empirical variables, comprising ETR1, ETR2, BTD, CEO closeness, CFO closeness, CEO degree, CFO degree, SIZE, MB, LEVER, NOL, ΔNOL, ROA, ESUB, PPE, INTAN, DFI, RD. The data covers January 1, 2010 to December 31, 2014. All variables are defined in Appendix A

EMPIRICAL RESULTS

Effects of Corporate Social Network on Tax Aggressive

This section explores how the corporate social network affect tax aggressive, we follow Chen et al. (2010) to adopt a multivariate regression model as following:

$$\begin{aligned} TAXAGG_{i,t} = \alpha_0 &+ \beta_1 NETWORK_{i,t-1} + \beta_2 SIZE_{i;t-1} + \beta_3 MB_{i,t-1} + \beta_4 LEVER_{i,t-1} \\ &+ \beta_5 NOL_{i,t-1} + \beta_6 \Delta NOL_{i,t-1} + \beta_7 ROA_{i,t-1} + \beta_8 ESUB_{i,t-1} + \beta_9 PPE_{i,t-1} \\ &+ \beta_{10} INTAN_{i,t-1} + \beta_{11} DFI_{i,t-1} + \beta_{12} RD_{i,t-1} + YearDummies \\ &+ IndustryDummies + \varepsilon_{i,t} \end{aligned} \quad (4)$$

Where the dependent variable tax aggressive ($TAXAGG_{i,t}$) represents the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. The independent variables include corporate social network ($NETWORK_{i,t-1}$) which represents the CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree), firm size ($SIZE_{i;t-1}$), market to book value ($MB_{i,t-1}$), leverage ratio ($LEVER_{i,t-1}$), net operating loss ($NOL_{i,t-1}$) and change in net operating loss ($\Delta NOL_{i,t-1}$), return on asset ($ROA_{i,t-1}$), equity method ($ESUB_{i,t-1}$), total plant, property and equipment ($PPE_{i,t-1}$), intangible assets ($INTAN_{i,t-1}$), income from foreign operations ($DFI_{i,t-1}$), research and development activities ($RD_{i,t-1}$) [Larger companies and firm complexity access to tax planning strategies that could be exhibit more tax evasion activities (e.g., Chen et al., 2010)]. *YearDummies* and *IndustryDummies* represent the industry and year dummy variables, respectively.

Table 3 presents separate results using closeness (Panel A) and degree (Panel B) to measure the social network of CEOs and CFOs. We begin our empirical analysis by providing a deeper understand of whether corporate social network affect tax aggressive. Table 3 shows the coefficients of CEO's and CFO's social network both are statistically significant almost at the 1% for all empirical results. The coefficients of CEO closeness and CEO degree are significantly positive, ranging from 0.010 to 0.014, significance at the 1% level for all ETR1 and ETR2 sample. An increase in CEO's social network has a clearly negative impact on BTD, ranging from -0.416 to -0.598, significance at the 1% level. These empirical results suggest that firm with a higher external social network of CEO is associate with lower tax aggressive. The coefficients of CFO's social network are negatively significant for two effective tax measures (ETR1 and ETR2), ranging from -0.010 to 0.018 at almost the 1% level, while the coefficients of CFO's external network range from 0.343 to 0.552, positive significant on BTD sample at least at the 5% level for both Panel A and Panel B, implying that CFOs with a higher external social network are more likely to participate tax aggressive activities.

The majority of control variables also show significant coefficients that are mostly consistent with our conjectures and prior literature. The coefficients of MB have significantly negative effect on ETR1 and ETR2, positive effect on BTD at the 1% significance level, respectively. As shown, the coefficients of NOL are statistically significant at the 1% level with an expected sign for our three tax aggressive measurements, indicating that firms with higher net operating loss are less likely to be associated with tax evasion. In the contrary, we also find that increases in change in net operating loss (ΔNOL) are associated with a higher tax avoidance. There is a significantly positive relation between RD and tax aggressive measures ETR1, and significantly negative relation between RD and BTD. These results mean that firms in with greater research and development activities also have a higher level of tax avoidance. Finally, our empirical results finding no correlation between foreign operating firms and tax avoidance. The most of above empirical results are suggested by Chen et al., (2010).

Table 3: Corporate social network and tax aggressive

Dependent variables	ETR1			ETR2			BTD		
Panel A: Social network measured by Closeness									
CEO Closeness	0.010*** (3.21)		0.010*** (2.97)	0.014*** (4.13)		0.013*** (3.87)	-0.598*** (-4.64)		-0.573*** (-4.43)
CFO Closeness		-0.016*** (-3.54)	-0.015*** (-3.31)		-0.018*** (-3.70)	-0.016*** (-3.41)		0.552*** (2.90)	0.486*** (2.55)
SIZE	0.001 (0.82)	0.002 (1.31)	0.001 (0.67)	0.005*** (3.35)	0.006*** (4.08)	0.004*** (3.20)	0.330*** (6.04)	0.284*** (5.31)	0.336*** (6.14)
MB	-0.001*** (-2.60)	-0.001*** (-2.61)	-0.001*** (-2.57)	-0.001*** (-2.79)	-0.001*** (-2.81)	-0.001*** (-2.76)	0.052*** (4.55)	0.052*** (4.53)	0.051*** (4.52)
LEVER	-0.001 (-0.61)	-0.001 (-0.71)	-0.001 (-0.63)	-0.002 (-1.50)	-0.002 (-1.63)	-0.002 (-1.52)	0.115** (2.28)	0.123** (2.42)	0.117** (2.30)
NOL	0.037*** (9.20)	0.038*** (9.50)	0.037*** (9.29)	0.029*** (6.99)	0.030*** (7.34)	0.029*** (7.08)	-0.753*** (-4.59)	-0.801*** (-4.88)	-0.763*** (-4.66)
ΔNOL	-0.889*** (-3.85)	-0.913*** (-3.95)	-0.913*** (-3.96)	-0.191 (-0.80)	-0.214 (-0.89)	-0.215 (-0.89)	10.275 (1.08)	10.907 (1.14)	11.025 (1.16)
ROA	0.003*** (16.05)	0.002*** (15.82)	0.003*** (15.99)	0.002*** (12.31)	0.002*** (11.99)	0.002*** (12.24)	0.069*** (10.55)	0.071*** (10.86)	0.069*** (10.60)
ESUB	0.019*** (8.13)	0.019*** (8.53)	0.019*** (8.18)	0.013*** (5.65)	0.014*** (6.10)	0.013*** (5.70)	-0.383*** (-4.11)	-0.430*** (-4.64)	-0.387*** (-4.15)
PPE	-0.031*** (-4.06)	-0.032*** (-4.18)	-0.032*** (-4.20)	-0.032*** (-4.05)	-0.033*** (-4.14)	-0.033*** (-4.18)	-1.053*** (-3.32)	-1.052*** (-3.31)	-1.024*** (-3.23)
INTAN	-0.036 (-1.20)	-0.036 (-1.20)	-0.038 (-1.26)	-0.078** (-2.44)	-0.077** (-2.41)	-0.080** (-2.50)	-2.246* (-1.79)	-2.408* (-1.91)	-2.216* (-1.76)
DFI	-0.003 (-0.88)	-0.003 (-1.11)	-0.003 (-1.01)	0.003 (1.00)	0.002 (0.77)	0.003 (0.87)	0.047 (0.38)	0.070 (0.57)	0.060 (0.49)
RD	0.001** (2.12)	0.001*** (2.59)	0.001** (2.31)	0.000 (1.01)	0.001 (1.57)	0.000 (1.21)	-0.027* (-1.71)	-0.036** (-2.24)	-0.030* (-1.85)
Intercept	0.051* (1.88)	0.064** (2.32)	0.070** (2.53)	0.048* (1.67)	0.060** (2.05)	0.068** (2.34)	-2.103* (-1.86)	-2.146* (-1.87)	-2.690** (-2.33)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R^2(%)	14.24	14.86	14.39	11.07	11.01	11.23	11.88	12.08	12.16
Obs.	5699	5699	5699	5706	5706	5706	5810	5810	5810
Panel B: Social network measured by Degree									
CEO Degree	0.011*** (3.42)		0.012*** (3.62)	0.012*** (3.62)		0.013*** (3.86)	-0.416*** (-3.22)		-0.447*** (-3.44)
CFO Degree		-0.010** (-2.45)	-0.011*** (-2.73)		-0.011*** (-2.74)	-0.013*** (-3.04)		0.343** (2.20)	0.394** (2.51)
SIZE	0.001 (0.67)	0.002 (1.52)	0.001 (0.60)	0.005*** (3.32)	0.006*** (4.30)	0.004*** (3.24)	0.318*** (5.77)	0.274*** (5.13)	0.319*** (5.80)
MB	-0.001*** (-2.55)	-0.001*** (-2.60)	-0.001** (-2.49)	-0.001*** (-2.74)	-0.001*** (-2.80)	-0.001*** (-2.68)	0.051*** (4.48)	0.052*** (4.53)	0.050*** (4.44)
LEVER	-0.001 (-0.62)	-0.001 (-0.69)	-0.001 (-0.62)	-0.002 (-1.54)	-0.002 (-1.60)	-0.002 (-1.54)	0.119** (2.35)	0.123** (2.41)	0.120** (2.36)
NOL	0.037*** (9.35)	0.038*** (9.47)	0.037*** (9.40)	0.030*** (7.21)	0.030*** (7.32)	0.030*** (7.26)	-0.787*** (-4.80)	-0.799*** (-4.87)	-0.796*** (-4.86)
ΔNOL	-0.904*** (-3.91)	-0.905*** (-3.91)	-0.925*** (-4.00)	-0.208 (-0.87)	-0.208 (-0.87)	-0.232 (-0.96)	10.742 (1.13)	10.517 (1.10)	11.338 (1.19)
ROA	0.003*** (16.09)	0.002*** (15.79)	0.003*** (16.03)	0.002*** (12.29)	0.002*** (11.97)	0.002*** (12.23)	0.069*** (10.60)	0.071*** (10.88)	0.070*** (10.66)
ESUB	0.019*** (8.20)	0.020*** (8.55)	0.019*** (8.23)	0.013*** (5.75)	0.014*** (6.12)	0.014*** (5.79)	-0.405*** (-4.35)	-0.434*** (-4.67)	-0.409*** (-4.40)

PPE	-0.031*** (-4.03)	-0.032*** (-4.15)	-0.032*** (-4.16)	-0.032*** (-4.00)	-0.033*** (-4.11)	-0.033*** (-4.14)	-1.069*** (-3.37)	-1.061*** (-3.34)	-1.039*** (-3.28)
INTAN	-0.037 (-1.24)	-0.035 (-1.14)	-0.038 (-1.26)	-0.078** (-2.45)	-0.075** (-2.35)	-0.079** (-2.47)	-2.350* (-1.87)	-2.449* (-1.95)	-2.339* (-1.86)
DFI	-0.003 (-0.89)	-0.003 (-1.08)	-0.003 (-1.00)	0.003 (0.99)	0.002 (0.80)	0.003 (0.88)	0.047 (0.38)	0.067 (0.54)	0.060 (0.48)
RD	0.001** (2.09)	0.001*** (2.58)	0.001** (2.27)	0.001 (1.06)	0.001 (1.58)	0.001 (1.25)	-0.029* (-1.82)	-0.036** (-2.26)	-0.032** (-1.99)
Intercept	0.055** (2.00)	0.053* (1.93)	0.066** (2.39)	0.049* (1.72)	0.047* (1.66)	0.062** (2.15)	-1.918* (-1.69)	-1.717 (-1.52)	-2.263** (-1.98)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj.R^2(%)	14.84	14.76	14.96	11.61	11.52	11.76	11.91	11.82	11.99
Obs.	5699	5699	5699	5706	5706	5706	5810	5810	5810

Notes: This table provides details of the social network effect on the tax aggressive from Eq. (4). The regression model is as follows:

$$TAXAGG_{i,t} = \alpha_0 + \beta_1 NETWORK_{i,t-1} + \beta_2 SIZE_{i;t-1} + \beta_3 MB_{i,t-1} + \beta_4 LEVER_{i,t-1} + \beta_5 NOL_{i,t-1} + \beta_6 \Delta NOL_{i,t-1} + \beta_7 ROA_{i,t-1} + \beta_8 ESUB_{i,t-1} + \beta_9 PPE_{i,t-1} + \beta_{10} INTAN_{i,t-1} + \beta_{11} DFI_{i,t-1} + \beta_{12} RD_{i,t-1} + YearDummies + IndustryDummies + \varepsilon_{i,t}$$

Where the dependent variable is tax aggressive ($TAXAGG_{i,t}$) for company i in t year. The independent variables include corporate social network ($NETWORK_{i,t-1}$), firm size ($SIZE_{i;t-1}$), market to book value ($MB_{i,t-1}$), leverage ratio ($LEVER_{i,t-1}$), net operating loss ($NOL_{i,t-1}$) and change in net operating loss ($\Delta NOL_{i,t-1}$), return on asset ($ROA_{i,t-1}$), equity method ($ESUB_{i,t-1}$), total plant, property and equipment ($PPE_{i,t-1}$), intangible assets ($INTAN_{i,t-1}$), income from foreign operations ($DFI_{i,t-1}$), research and development activities ($RD_{i,t-1}$). *YearDummies* and *IndustryDummies* represent the industry and year dummy variables, respective. The tax aggressive ($TAXAGG_{i,t}$) includes the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. Corporate social network ($NETWORK_{i,t-1}$) represent for the CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). All variables are defined in Appendix A. The *t*-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Endogenous Concern

The Effect of Corporate Governance

External and internal monitoring may play an important role in firm tax aggressive activity. Minnick and Noga (2010), Lanis and Richardson (2011) argued that quality of corporate governance has significant impact on tax aggressive activity. Therefore, unobservable variables that are common to corporate governance can also generate a positive relation between corporate social network and tax aggressive (the omitted-variable bias). We follow Hartzell and Starks (2003), Cremers and Nair (2005), Del Guercio et al. (2008), E. H. Kim and Lu (2011) adopt board size (BS), independent director (ID) and board and supervisor ownership (BSO) variables which assume play in sharping governance to our following regression model as following:

$$TAXAGG_{i,t} = \alpha_0 + \beta_1 NETWORK_{i,t-1} + \beta_2 GOVERNANCE_{i;t-1} + \beta_3 SIZE_{i;t-1} + \beta_4 MB_{i,t-1} + \beta_5 LEVER_{i,t-1} + \beta_6 NOL_{i,t-1} + \beta_7 \Delta NOL_{i,t-1} + \beta_8 ROA_{i,t-1} + \beta_9 ESUB_{i,t-1} + \beta_{10} PPE_{i,t-1} + \beta_{11} INTAN_{i,t-1} + \beta_{12} DFI_{i,t-1} + \beta_{13} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \varepsilon_{i,t} \quad (5)$$

Where the governance variables ($GOVERNANCE_{i;t-1}$) includes the number of board members ($BS_{i;t-1}$), the percentage of independent directors ($ID_{i;t-1}$), the percentage of board and supervisor ownership ($BSO_{i;t-1}$). All variables are defined in Appendix A [The remaining control variables are the same as those in Eq. (4)]

Table 4 presents separate results using closeness (Panel A) and degree (Panel B) to measure the social network of CEOs and CFOs. Our results show the coefficients of BS has a positive significant impact on ETR2 at 0.002 at least the 5% level and negative significant impact on BTD, ranging from -0.123 to -0.143 at the 1% level. While BSO has a significant impact on ETR1 at 0.001 at the 5% level. In addition, the coefficients of ID range from 0.137 to 0.142, significance at the 1% levels on BTD for both panels. These results indicate that corporate governance variables have significant impact on tax aggressive activity. Table 4 show that CEOs have a significantly positive impact on two tax effective measures (ETR1 and ETR2) and negative impact on BTD at the 1% level for all our sample. For the CFO's social network, the coefficients of CFOs have significantly negative impact on all ETR1, ETR2 and positive impact to BTD at least the 5% level. These results of Table 4 are similar to our prior empirical results.

Table 4: The effect of corporate governance

Dependent variables	ETR1			ETR2			BTD		
Panel A: Social network measured by Closeness									
CFO Closeness	0.011*** (3.00)		0.009*** (2.76)	0.013*** (3.82)		0.012*** (3.56)	-0.526*** (-4.06)		-0.501*** (-3.85)
CFO Closeness		-0.016*** (-3.47)	-0.015*** (-3.26)		-0.018*** (-3.68)	-0.016*** (-3.41)		0.543*** (2.85)	0.485*** (2.55)
BS	0.001 (1.48)	0.001* (1.83)	0.001 (1.52)	0.002** (2.42)	0.002*** (2.87)	0.002** (2.48)	-0.123*** (-4.15)	-0.140*** (-4.74)	-0.125*** (-4.21)
ID	-0.001 (-0.55)	-0.001 (-0.58)	-0.001 (-0.57)	0.001 (0.32)	0.000 (0.28)	0.001 (0.29)	0.141** (2.51)	0.141** (2.51)	0.142** (2.53)
BSO	0.001** (-2.39)	0.001** (-2.36)	0.001** (-2.29)	0.001 (-1.33)	0.001 (-1.32)	0.001 (1.23)	0.014*** (3.51)	0.015*** (3.57)	0.014*** (3.43)
SIZE	0.001 (0.35)	0.001 (0.70)	0.001 (0.20)	0.004*** (2.62)	0.004*** (3.15)	0.004** (2.47)	0.379*** (6.78)	0.347*** (6.29)	0.385*** (6.88)
MB	-0.001** (-2.44)	-0.001** (-2.45)	-0.001** (-2.42)	-0.001*** (-2.72)	-0.001*** (-2.73)	-0.001*** (-2.69)	0.048*** (4.26)	0.048*** (4.24)	0.048*** (4.24)
LEVER	-0.001 (-0.66)	-0.001 (-0.74)	-0.001 (-0.68)	-0.002 (-1.43)	-0.002 (-1.54)	-0.002 (-1.46)	0.119** (2.35)	0.125** (2.46)	0.120** (2.37)
NOL	0.036*** (9.14)	0.037*** (9.41)	0.037*** (9.22)	0.029*** (6.95)	0.030*** (7.28)	0.029*** (7.03)	-0.738*** (-4.51)	-0.780*** (-4.77)	-0.749*** (-4.58)
ΔNOL	-0.876*** (-3.79)	-0.900*** (-3.89)	-0.900*** (-3.90)	-0.184 (-0.76)	-0.207 (-0.86)	-0.208 (-0.87)	9.853 (1.04)	10.534 (1.11)	10.634 (1.12)
ROA	0.003*** (16.27)	0.003*** (16.06)	0.003*** (16.20)	0.002*** (12.43)	0.002*** (12.16)	0.002*** (12.36)	0.064*** (9.76)	0.066*** (10.01)	0.065*** (9.82)
ESUB	0.019*** (8.29)	0.020*** (8.65)	0.019*** (8.32)	0.013*** (5.67)	0.014*** (6.08)	0.013*** (5.71)	-0.406*** (-4.35)	-0.446*** (-4.80)	-0.409*** (-4.38)
PPE	-0.031*** (-4.01)	-0.032*** (-4.13)	-0.032*** (-4.14)	-0.033*** (-4.11)	-0.034*** (-4.22)	-0.034*** (-4.25)	-1.103*** (-3.48)	-1.093*** (-3.44)	-1.073*** (-3.38)
INTAN	-0.036 (-1.20)	-0.037 (-1.21)	-0.038 (-1.26)	-0.080** (-2.49)	-0.079** (-2.47)	-0.081** (-2.54)	-2.443* (-1.94)	-2.578** (-2.05)	-2.416* (-1.92)
DFI	-0.003 (-0.93)	-0.003 (-1.13)	-0.003 (-1.05)	0.003 (0.99)	0.002 (0.78)	0.003 (0.87)	0.035 (-0.28)	0.055 (0.45)	0.047 (0.38)
RD	0.001* (1.95)	0.001** (2.39)	0.001** (2.15)	0.001 (0.83)	0.001 (1.34)	0.001 (1.03)	-0.023 (-1.45)	-0.030* (-1.89)	-0.026 (-1.59)
Panel B: Social network measured by Degree									
CEO Degree	0.010*** (3.27)		0.011*** (3.46)	0.011*** (3.32)		0.012*** (3.55)	-0.361*** (-2.78)		-0.390*** (-3.00)
CFO Degree		-0.010** (-2.46)	-0.011*** (-2.71)		-0.012*** (-2.58)	-0.013*** (-3.12)		0.360** (2.30)	0.401*** (2.56)
BS	0.001 (1.43)	0.001* (1.91)	0.001 (1.53)	0.002** (2.45)	0.002*** (2.97)	0.002*** (2.56)	-0.130*** (-4.39)	-0.143*** (-4.83)	-0.134*** (-4.50)
ID	-0.001 (-0.55)	-0.001 (-0.52)	-0.001 (-0.51)	0.001 (0.32)	0.001 (0.35)	0.001 (0.36)	0.140** (2.49)	0.137** (2.43)	0.137** (2.43)
BSO	0.001** (-2.48)	0.001** (-2.42)	0.001** (-2.41)	0.001 (-1.44)	0.001 (-1.37)	0.001 (-1.37)	0.015*** (3.63)	0.015*** (3.61)	0.015*** (3.56)
SIZE	0.001 (0.20)	0.001 (0.87)	0.001 (0.12)	0.004*** (2.59)	0.005*** (3.32)	0.004** (2.49)	0.371*** (6.58)	0.338*** (6.14)	0.374*** (6.64)
MB	-0.001** (-2.38)	-0.001** (-2.44)	-0.001** (-2.34)	-0.001*** (-2.67)	-0.001*** (-2.72)	-0.001*** (-2.62)	0.048*** (4.20)	0.048*** (4.24)	0.047*** (4.16)
LEVER	-0.001 (-0.67)	-0.001 (-0.71)	-0.001 (-0.66)	-0.002 (-1.48)	-0.002 (-1.51)	-0.002 (-1.47)	0.122** (2.41)	0.124** (2.45)	0.122** (2.41)
NOL	0.037*** (9.27)	0.037*** (9.38)	0.037*** (9.33)	0.030*** (7.15)	0.030*** (7.25)	0.030*** (7.20)	-0.767*** (-4.69)	-0.779*** (-4.76)	-0.776*** (-4.75)
ΔNOL	-0.889*** (-3.85)	-0.892*** (-3.86)	-0.910*** (-3.94)	-0.198 (-0.82)	-0.202 (-0.84)	-0.223 (-0.93)	10.236 (1.07)	10.167 (1.07)	10.864 (1.14)

ROA	0.003*** (16.31)	0.003*** (16.04)	0.003*** (16.25)	0.002*** (12.43)	0.002*** (12.14)	0.002*** (12.36)	0.065*** (9.79)	0.066*** (10.02)	0.065*** (9.85)
ESUB	0.019*** (8.35)	0.020*** (8.67)	0.019*** (8.38)	0.014*** (5.77)	0.014*** (6.10)	0.014*** (5.80)	-0.425*** (-4.56)	-0.450*** (-4.84)	-0.429*** (-4.60)
PPE	-0.031*** (-3.97)	-0.032*** (-4.11)	-0.032*** (-4.11)	-0.033*** (-4.07)	-0.034*** (-4.20)	-0.034*** (-4.22)	-1.115*** (-3.51)	-1.098*** (-3.46)	-1.081*** (-3.41)
INTAN	-0.038 (-1.25)	-0.035 (-1.16)	-0.038 (-1.27)	-0.080** (-2.50)	-0.077** (-2.42)	-0.081** (-2.52)	-2.532** (-2.01)	-2.609** (-2.07)	-2.515** (-2.00)
DFI	-0.003 (-0.94)	-0.003 (-1.11)	-0.003 (-1.05)	0.003 (0.97)	0.003 (0.80)	0.003 (0.86)	0.035 (0.29)	0.053 (0.43)	0.048 (-0.39)
RD	0.001* (1.92)	0.001** (2.37)	0.001** (2.09)	0.001 (0.86)	0.001 (1.35)	0.001 (1.06)	-0.025 (-1.53)	-0.031* (-1.91)	-0.027* (-1.69)

Notes: This table provides details of the effects of corporate social network on tax aggressiveness by using firm fixed-effect regressions to control for the influences of unknown firm characteristics that affect social network and tax avoidance. The regression model as follow:

$$TAXAGG_{i,t} = \alpha_0 + \beta_1 NETWORK_{i,t-1} + \beta_2 GOVERNANCE_{i;t-1} + \beta_3 SIZE_{i;t-1} + \beta_4 MB_{i,t-1} + \beta_5 LEVER_{i,t-1} + \beta_6 NOL_{i,t-1} + \beta_7 \Delta NOL_{i,t-1} + \beta_8 ROA_{i,t-1} + \beta_9 ESUB_{i,t-1} + \beta_{10} PPE_{i,t-1} + \beta_{11} INTAN_{i,t-1} + \beta_{12} DFI_{i,t-1} + \beta_{13} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \varepsilon_{i,t}$$

Where the dependent variable is tax aggressive ($TAX_{i,t}$) for company i in t year. The independent variables include corporate social network ($NETWORK_{i,t-1}$), corporate governance ($GOVERNANCE_{i;t-1}$), firm size ($SIZE_{i;t-1}$), market to book value ($MB_{i,t-1}$), leverage ratio ($LEVER_{i,t-1}$), net operating loss ($NOL_{i,t-1}$) and change in net operating loss ($\Delta NOL_{i,t-1}$), return on asset ($ROA_{i,t-1}$), equity method ($ESUB_{i,t-1}$), total plant, property and equipment ($PPE_{i,t-1}$), intangible assets ($INTAN_{i,t-1}$), income from foreign operations ($DFI_{i,t-1}$), research and development activities ($RD_{i,t-1}$). *YearDummies* and *IndustryDummies* represent the industry and year dummy variables, respective. The tax aggressive measures ($TAXAGG_{i,t}$) include the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. Corporate social network ($NETWORK_{i,t-1}$) represent for the CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). Corporate governance ($GOVERNANCE_{i;t-1}$) includes the number of board members (*BS*), the percentage of independent directors (*ID*), the percentage of board and supervisor ownership (*BSO*). All variables are defined in Appendix A. The *t*-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively

Instrumental Regression

For reducing the endogeneity problem caused by missing variables, we use the instrumental variables model, and the adopted instruments include corporate governance variables. First, we adopt two governance variables related to internal monitoring by the board. We adopt the number of board member (BS) in our regression. Prior studies indicate that larger board tend to be less effective monitors (Eisenberg et al., 1998; Jensen, 1993; Lipton & Lorsch, 1992; Yermack, 1996). Next, we include the percentage of independent directors on the board (ID). The monitoring role played by independent directors has been widely documented. For example, Weisbach (1988) finds that CEO turnover following poor performance is positively related to the fraction of outside directors. Second, to capture the strength of external monitoring, we follow Hartzell and Starks (2003) and estimate institutional ownership concentration by the percentage of institutional ownership (IO). Previous studies document the important role institutional investors play in shaping corporate governance (Cremers & Nair, 2005; Del Guercio et al., 2008; Edmans et al., 2011; Hartzell & Starks, 2003; E. H. Kim & Lu, 2011). The selection of instruments is due to the reason well-governed firms prefer to have highly-skilled executives and directors such as highly externally-tied executives and directors (Benton, 2017; Davis & Greve, 1997; Fracassi & Tate, 2012; Mcdonald et al., 2008). The regression model as following:

1st Stage:

$$NETWORK_{i,t-1} = \alpha_0 + \gamma_1 INSTRU_{i;t} + \beta_1 SIZE_{i;t-1} + \beta_2 MB_{i,t-1} + \beta_3 LEVER_{i,t-1} + \beta_4 NOL_{i,t-1} + \beta_5 \Delta NOL_{i,t-1} + \beta_6 ROA_{i,t-1} + \beta_7 ESUB_{i,t-1} + \beta_8 PPE_{i,t-1} + \beta_9 INTAN_{i,t-1} + \beta_{10} DFI_{i,t-1} + \beta_{11} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \eta_{i,t} \quad (6)$$

2st Stage: (7)

$$TAXAGG_{i,t} = \alpha_0 + \delta_1 \widehat{NETWORK}_{i,t-1} + \beta_1 SIZE_{i;t-1} + \beta_2 MB_{i,t-1} + \beta_3 LEVER_{i,t-1} + \beta_4 NOL_{i,t-1} + \beta_5 \Delta NOL_{i,t-1} + \beta_6 ROA_{i,t-1} + \beta_7 ESUB_{i,t-1} + \beta_8 PPE_{i,t-1} + \beta_9 INTAN_{i,t-1} + \beta_{10} DFI_{i,t-1} + \beta_{11} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \varepsilon_{i,t}$$

In the first-stage, the dependent variable is corporate social network ($NETWORK_{i,t-1}$). The instrumental variables ($INSTRU_{i;t}$) includes the percentage of institution ownership ($IO_{i;t-1}$), the number of board members ($BS_{i;t-1}$), the percentage of independent director ($ID_{i;t-1}$). In the second-stage, the dependent variable is the tax aggressive measure, which is defined as ETR1, ETR2 and BTD. The remaining control variables are the same as those in Eq.(4). All variables are defined in Appendix A.

We separate the results using closeness in Panel A and degree in Panel B to measure the social network of CEOs and CFOs. Table 5 shows the results of the instrumental variable regression in which we use two-stage least squares to estimate the coefficients. In the first stage, our IVs have the significant effect on corporate social network. The findings are largely consistent with our expectations. The second-stage regression results reported in Table V show the robustness of our findings to controlling for endogeneity and simultaneity biases. The coefficients of
$\widehat{CEO}$ are significantly positive impact on ETR1 and ETR2, ranging from 0.080 to 0.117, significance at the 1% levels and significantly negative effect, ranging from -5.460 to -6.708, significance at the 1% level on BTD. It strongly implies and consistent evidence that higher CEO's social network are associated with lower degrees of tax avoidance.

Our results also present $\widehat{CFO}$ has a significantly negative impact on ETR1, ranging from -0.233 to -0.655 at the 5% level for our all sample. In addition, the coefficients of $\widehat{CFO}$ is significantly negative on ETR2 (as shown in Panel B of Table 5) at the 5% level and significantly positive on BTD (as shown in Panel A of Table 5) at the 1% level. The results of Table 5 are similar to our empirical results and there is no evidence of reverse causation or simultaneity bias. However, we consider instrumental variables regressions as one of several different techniques for controlling endogenous levels and ignoring variable biases and presenting other tests below.

Table 5: Endogenous concerns: Instrumental regression results

Variable	1st stage		2nd stage		1st stage		2nd stage		1st stage		2nd stage	
	CEO	CFO	ETR1	ETR1	CEO	CFO	ETR2	ETR2	CEO	CFO	BTD	BTD
Panel A: Social network measured by Closeness												
$\widehat{CEO}$			0.109*** (3.29)				0.117*** (3.63)				-5.460*** (-4.74)	
$\widehat{CFO}$				-0.651** (-2.27)				-0.702** (-2.28)				27.483*** (3.64)
IO	-0.001*** (-4.56)	0.001** (2.45)			-0.001*** (-4.33)	0.001** (2.18)			-0.002 (0.01)	0.001 (0.01)		
BS	0.021*** (7.10)	-0.001 (-1.21)			0.021*** (7.53)	-0.002* (-1.85)			0.027*** (8.59)	-0.003 (0.01)		
ID	0.001 (0.13)	0.000 (0.19)			0.002 (0.46)	-0.001 (-0.42)			-0.002 (-0.42)	0.003 (0.01)		
SIZE	0.087*** (14.85)	-0.021*** (-5.47)	-0.007** (-2.34)	-0.009 (-1.55)	0.086*** (14.78)	-0.020*** (-5.18)	-0.004 (-1.38)	-0.006 (-0.98)	0.094*** (15.84)	-0.020*** (-5.17)	0.777*** (6.39)	0.704*** (4.15)
MB	-0.001 (-0.89)	0.001 (0.68)	-0.001** (-2.00)	0.000 (-0.54)	-0.001 (-0.92)	0.001 (0.76)	-0.001** (-2.16)	0.000 (-0.55)	0.001 (-0.17)	0.001 (0.01)	0.050*** (3.91)	0.034 (1.38)
LEVER	-0.008 (-1.57)	-0.002 (-0.50)	0.001 (0.11)	-0.002 (-0.72)	-0.009* (-1.73)	-0.002 (-0.50)	-0.001 (-0.62)	-0.003 (-1.13)	-0.009* (-1.85)	-0.002 (-0.61)	0.062 (1.07)	0.176 (1.63)
NOL	0.070*** (4.33)	0.020* (1.77)	0.029*** (5.78)	0.049*** (5.02)	0.073*** (4.50)	0.017 (1.53)	0.021*** (3.99)	0.041*** (4.06)	0.054*** (3.24)	0.021* (1.84)	-0.447** (-2.27)	-1.289*** (-3.46)
ΔNOL	0.202 (0.21)	-1.573** (-2.38)	-0.912*** (-3.67)	-1.904*** (-2.91)	0.197 (0.21)	-1.461** (-2.22)	-0.214 (-0.83)	-1.209* (-1.78)	0.278 (0.29)	-0.408 (-0.54)	10.540 (0.99)	11.901 (0.56)
ROA	-0.003*** (-4.30)	-0.001 (-1.38)	0.003*** (14.31)	0.002*** (6.04)	-0.003*** (-4.36)	-0.001 (-1.44)	0.002*** (11.51)	0.002*** (4.07)	-0.002 (0.01)	-0.001 (0.01)	0.055*** (6.94)	0.085*** (5.93)
ESUB	0.073*** (7.86)	0.002 (0.32)	0.011*** (3.26)	0.021*** (4.39)	0.067*** (7.32)	0.003 (0.50)	0.006* (1.84)	0.016*** (3.25)	0.074*** (7.91)	0.003 (0.53)	-0.012 (-0.09)	-0.521*** (-2.65)
PPE	0.016 (0.50)	-0.070*** (-3.18)	-0.033*** (-4.02)	-0.076*** (-3.01)	0.023 (0.74)	-0.065*** (-2.98)	-0.035*** (-4.09)	-0.078*** (-2.98)	0.051 (1.58)	-0.067*** (-3.05)	-0.781** (-2.18)	0.729 (0.87)
INTAN	0.191 (1.55)	-0.133 (-1.54)	-0.055* (-1.66)	-0.118* (-1.65)	0.223* (1.78)	-0.115 (-1.32)	-0.101*** (-2.88)	-0.155** (-2.05)	0.356*** (2.80)	-0.096 (-1.11)	-0.589 (-0.41)	-0.013 (0.01)
DFI	-0.025** (-2.05)	-0.024*** (-2.79)	0.001 (0.02)	-0.018** (-2.00)	-0.020 (-1.62)	-0.023*** (-2.68)	0.005 (1.58)	-0.013 (-1.37)	-0.011 (-0.89)	-0.029*** (-3.40)	-0.026 (-0.19)	0.831** (2.44)
RD	0.009*** (5.86)	0.005*** (4.25)	0.001 (-0.39)	0.004** (2.52)	0.009*** (5.95)	0.005*** (4.10)	-0.001 (-1.25)	0.004** (2.25)	0.009*** (5.39)	0.004 (0.01)	0.023 (1.06)	-0.139*** (-3.16)
Panel B: Social network measured by Degree												
$\widehat{CEO}$			0.080** (2.16)				0.104*** (2.91)				-6.708*** (-3.87)	
$\widehat{CFO}$				-0.233** (-2.45)				-0.068 (-0.69)				13.710*** (4.96)
IO	-0.001* (-1.65)	0.001*** (4.97)			-0.001* (-1.66)	0.001*** (3.71)			-0.001 (0.01)	0.001 (0.01)		
BS	0.022***	0.003			0.022***	0.007*			0.021***	0.001		

	(7.36)	(1.13)			(7.47)	(1.84)			(6.46)	(0.02)		
ID	0.005 (0.82)	0.003 (0.80)			0.005 (0.95)	0.002 (0.53)			-0.002 (-0.33)	0.008** (2.41)		
SIZE	0.094*** (15.19)	-0.013*** (-2.67)	-0.006 (-1.52)	0.001 (0.86)	0.094*** (15.66)	-0.014*** (-2.85)	-0.005 (-1.20)	0.006 *** (4.08)	0.102*** (17.13)	-0.006 (-1.18)	0.956*** (5.12)	0.225 *** (2.79)
MB	-0.002** (-2.10)	0.001 (1.11)	-0.001* (-1.77)	0.001 (-1.24)	-0.002** (-2.12)	0.001 (1.17)	-0.001* (-1.76)	-0.001** (-2.32)	-0.002* (-1.80)	0.001 (0.01)	0.037*** (2.65)	0.036** (2.10)
LEVER	-0.006 (-1.15)	0.001 (-0.04)	0.001 (-0.21)	-0.001 (-0.56)	-0.006 (-1.13)	0.001 (0.01)	-0.001 (-0.96)	-0.002 (-1.59)	-0.005 (-1.06)	-0.001 (-0.34)	0.078 (1.27)	0.141* (1.85)
NOL	0.021 (1.27)	0.023* (1.77)	0.035*** (8.36)	0.042*** (7.95)	0.019 (1.15)	0.021 (1.60)	0.028 *** (6.18)	0.031*** (6.88)	0.004 (0.25)	0.028** (2.02)	-0.710*** (-3.64)	-1.111*** (-4.37)
ΔNOL	1.513 (1.58)	-1.831** (-2.38)	-1.010*** (-4.10)	-1.307*** (-3.91)	1.599* (1.67)	-1.799** (-2.34)	-0.357 (-1.36)	-0.309 (-1.03)	1.207 (1.24)	-0.672 (-0.82)	11.751 (1.03)	11.454 (0.80)
ROA	-0.003*** (-5.17)	-0.001** (-2.42)	0.003*** (13.18)	0.002*** (9.99)	-0.003*** (-5.18)	-0.001** (-2.18)	0.002*** (10.84)	0.002*** (9.59)	-0.003 (0.01)	-0.001 (0.01)	0.050*** (5.34)	0.089*** (8.50)
ESUB	0.058*** (6.19)	0.008 (1.07)	0.015*** (4.49)	0.022*** (7.26)	0.058*** (6.27)	0.009 (1.22)	0.008** (2.42)	0.015*** (5.77)	0.056*** (5.94)	0.014* (1.85)	-0.042 (-0.28)	-0.628*** (-4.33)
PPE	-0.011 (-0.34)	-0.097*** (-3.79)	-0.031*** (-3.86)	-0.053*** (-4.04)	-0.003 (-0.11)	-0.091*** (-3.56)	-0.032*** (-3.82)	-0.038*** (-3.18)	0.038 (1.18)	-0.078*** (-2.93)	-0.803** (-2.10)	-0.084 (-0.16)
INTAN	0.296** (2.36)	-0.055 (-0.54)	-0.059* (-1.76)	-0.044 (-1.16)	0.270** (2.13)	-0.035 (-0.34)	-0.104*** (-2.94)	-0.076** (-2.36)	0.255** (2.00)	-0.032 (-0.30)	-0.765 (-0.49)	-2.268 (-1.20)
DFI	-0.022* (-1.77)	-0.031*** (-3.13)	-0.001 (-0.32)	-0.010** (-2.14)	-0.021* (-1.65)	-0.029*** (-2.90)	0.005 (1.51)	0.001 (0.19)	-0.018 (-1.42)	-0.035*** (-3.34)	-0.083 (-0.55)	0.507** (2.46)
RD	0.010 *** (6.26)	0.008*** (5.91)	0.001 (0.11)	0.003*** (3.11)	0.010*** (6.26)	0.008*** (5.75)	-0.001 (-1.00)	0.001 (1.28)	0.009*** (5.77)	0.008 (0.01)	0.036 (1.39)	-0.137*** (-4.31)

Notes: This table provides details of the effects of corporate social network on tax aggressiveness by using instrumental regression. The model is:

1st Stage:

$$NETWORK_{i,t-1} = \alpha_0 + \gamma_1 INSTRU_{i;t} + \beta_1 SIZE_{i;t-1} + \beta_2 MB_{i,t-1} + \beta_3 LEVER_{i,t-1} + \beta_4 NOL_{i,t-1} + \beta_5 \Delta NOL_{i,t-1} + \beta_6 ROA_{i,t-1} + \beta_7 ESUB_{i,t-1} + \beta_8 PPE_{i,t-1} + \beta_9 INTAN_{i,t-1} + \beta_{10} DFI_{i,t-1} + \beta_{11} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \eta_{i,t}$$

2st Stage:

$$TAXAGG_{i,t} = \alpha_0 + \delta_1 \widehat{NETWORK}_{i,t-1} + \beta_1 SIZE_{i;t-1} + \beta_2 MB_{i,t-1} + \beta_3 LEVER_{i,t-1} + \beta_4 NOL_{i,t-1} + \beta_5 \Delta NOL_{i,t-1} + \beta_6 ROA_{i,t-1} + \beta_7 ESUB_{i,t-1} + \beta_8 PPE_{i,t-1} + \beta_9 INTAN_{i,t-1} + \beta_{10} DFI_{i,t-1} + \beta_{11} RD_{i,t-1} + Year\ Dummies + Industry\ Dummies + \varepsilon_{i,t}$$

In the first-stage, the dependent variable is corporate social network ($NETWORK_{i,t-1}$). The instrumental variables ($INSTRU_{i;t}$) includes the percentage of institution ownership ($IO_{i;t-1}$), the number of board members ($BS_{i;t-1}$), the percentage of board and supervisor ownership ($BSO_{i;t-1}$). In the second-stage, the dependent variable is the tax aggressive measures ($TAXAGG_{i,t}$). The control variables are firm size ($SIZE_{i;t-1}$), market to book value ($MB_{i,t-1}$), leverage ratio ($LEVER_{i,t-1}$), net operating loss ($NOL_{i,t-1}$) and change in net operating loss ($\Delta NOL_{i,t-1}$), return on asset ($ROA_{i,t-1}$), equity method ($ESUB_{i,t-1}$), total plant, property and equipment ($PPE_{i,t-1}$), intangible assets ($INTAN_{i,t-1}$), income from foreign operations ($DFI_{i,t-1}$), research and development activities ($RD_{i,t-1}$). YearDummies and IndustryDummies represent the industry and year dummy variables, respective. The tax aggressive ($TAXAGG_{i,t}$) includes the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. Corporate social network ($NETWORK_{i,t-1}$) represent for the CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). All variables are defined in Appendix A. The t-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Robustness Tests

Hanlon and Heitzman (2010) recommend that researchers carefully consider the appropriateness of tax avoidance measures. We consider whether other tax aggressive measures can significant impact to the corporate social network. A large of prior researches by Desai and Dharmapala (2006), Dyreng et al. (2008), Dyreng et al. (2010), Minnick and Noga (2010) use effective tax rate (total income tax expense divided by pretax income) and cash effective tax rate (cash tax expense divided by pretax income). As our knowledge, effective tax rate captures the effect on financial accounting earnings, meanwhile cash effective tax rate captures the effect of cash outlay. Therefore, financial reporting income and taxable income generally differ, effective tax rate and cash effective tax rate will also differ. We adopt effective tax rate (ETR3) and cash effective tax rate (ETR4) to examine the robustness test, as shown in Panel A of Table 6. Next, we consider how alternative measures of corporate social network impact on tax aggressive, as shown in Panel B of Table 6. We use Betweenness and Eigenvector to measure for corporate social network of CEOs and CFOs. We then include it as our independent variable in Eq.(4). Finally, to obtain a clearer picture on the relation between tax aggressive and corporate social network, we examine whether using other regression can lead to other results. We use Tobit regression to test our main hypothesis with ETR1, ETR2, ETR3, ETR4 as dependent variables, as shown in Panel C of Table 6. The selection Tobit regression due to the reason that our tax aggressive measures are truncated in range from zero to one, therefore, using Tobit regression is appropriate with our dependent variables.

In Panel A of Table 6, the coefficients of CEO closeness are significantly positive, ranging from 0.011 to 0.012, significance at the 1% level. CEO degree also have a significantly positive impact on ETR3, ETR4, from 0.009 to 0.013 at the 1% level. The results suggest that higher CEOs corporate social network can lead to a decrease in tax aggressive. In addition, the coefficients of CFO closeness range from -0.011 to -0.014, significant at the 5% and the 1% levels. CFO degree has a negative significant impact on ETR3, ETR4, from -0.006 to -0.009, at the 10% level. These results posit that an increase in CFOs social network can lead to an increase in tax evasion. The results in Panel A of Table 6 are similar to our main results, it indicates the association between tax aggressive and corporate social network is economically significant as well and the sign as we expected.

Panel B of Table 6 shows that CEOs betweenness have a positive impact on ETR1, ETR2 and negative impact on BTD, significance at the 1% level. In addition, the coefficients of CEOs eigenvector range from 0.011 to 0.012, significance at the 10% level on ETR1. These results indicate that a higher CEOs corporate social network lead a decrease tax aggressive. Our results show that CFOs betweeness have a significant negative impact on ETR1, ETR2 at the 1% significance level, and a significant positive impact on BTD at the 10% significance level. While the impact of CFOs eigenvector on our tax aggressive measures is not clearly. The results in Panel B of Table 6 are as our expectations and support to our main results that better CEO's (CFO's) corporate social network can significantly decrease (increase) tax aggressive.

In Panel C of Table 6, CEO closeness has significantly positive effects on all our tax aggressive measures, ranging from 0.030 to 0.123, significance at all the 1% level. The coefficients of CEO degree are also significant positive, ranging from 0.014 to 0.085, significance almost at the 1% level. Our results show the coefficients of CFO closeness are significant negative, ranging from -0.034 to -0.094, significance at the 1% level; CFO degree are also significant negative, ranging from -0.012 to -0.068. These results are similar to our prior empirical results, implying that better CEO's (CFO's) corporate social network is associated with lower (higher) tax aggressive

Table 6: Robustness tests

Panel A: Alternative Measure of Tax Avoidance

Dependent variables	ETR3						ETR4					
CEO Closeness	0.011*** (2.81)		0.011*** (2.71)				0.012*** (3.39)		0.012*** (3.22)			
CFO closeness		-0.012** (-2.32)	-0.011** (-2.20)					-0.014*** (-2.71)	-0.013** (-2.48)			
CEO Degree				0.013*** (3.65)		0.013*** (3.76)				0.009** (2.47)		0.009*** (2.63)
CFO Degree					-0.006* (-1.65)	-0.007* (-1.71)					-0.008* (-1.92)	-0.009** (-2.12)
SIZE	0.001 (-0.16)	0.001 (0.22)	0.000 (-0.27)	-0.001 (-0.52)	0.001 (0.33)	-0.001 (-0.57)	0.006*** (4.10)	0.007*** (4.73)	0.006*** (3.98)	0.006*** (4.17)	0.007*** (4.88)	0.006*** (4.11)
MB	-0.001** (-2.08)	-0.001** (-2.08)	-0.001** (-2.06)	-0.001** (-2.00)	-0.001** (-2.08)	-0.001** (-1.97)	0.000 (-1.12)	0.001 (-1.11)	0.000 (-1.09)	0.001 (-1.07)	0.001 (-1.11)	0.001 (-1.03)
LEVER	0.001 (-0.18)	0.001 (-0.25)	0.000 (-0.19)	0.001 (-0.17)	0.001 (-0.24)	0.001 (-0.17)	-0.001 (-0.80)	-0.001 (-0.88)	-0.001 (-0.82)	-0.001 (-0.83)	-0.001 (-0.87)	-0.001 (-0.82)
NOL	0.038*** (8.69)	0.039*** (8.88)	0.038*** (8.74)	0.038*** (8.74)	0.038*** (8.86)	0.038*** (8.77)	0.018*** (4.04)	0.019*** (4.34)	0.018*** (4.10)	0.019*** (4.21)	0.019*** (4.32)	0.019*** (4.25)
ΔNOL	-0.551** (-2.20)	-0.569** (-2.27)	-0.567** (-2.26)	-0.578** (-2.31)	-0.563** (-2.24)	-0.592** (-2.36)	-0.192 (-0.75)	-0.209 (-0.81)	-0.211 (-0.82)	-0.206 (-0.80)	-0.205 (-0.80)	-0.225 (-0.88)
ROA	0.003*** (15.21)	0.003*** (15.06)	0.003*** (15.17)	0.003*** (15.35)	0.003*** (15.05)	0.003*** (15.32)	0.003*** (17.26)	0.003*** (17.03)	0.003*** (17.21)	0.003*** (17.22)	0.003*** (17.01)	0.003*** (17.17)
ESUB	0.013*** (5.10)	0.013*** (5.40)	0.013*** (5.13)	0.012*** (5.08)	0.013*** (5.41)	0.012*** (5.10)	0.013*** (5.16)	0.014*** (5.53)	0.013*** (5.19)	0.013*** (5.30)	0.014*** (5.53)	0.013*** (5.31)
PPE	-0.024*** (-2.84)	-0.025*** (-2.93)	-0.025*** (-2.93)	-0.024*** (-2.82)	-0.024*** (-2.90)	-0.024*** (-2.91)	-0.037*** (-4.27)	-0.037*** (-4.36)	-0.037*** (-4.37)	-0.036*** (-4.24)	-0.037*** (-4.34)	-0.037*** (-4.34)
INTAN	-0.036 (-1.08)	-0.037 (-1.10)	-0.038 (-1.13)	-0.039 (-1.16)	-0.035 (-1.06)	-0.039 (-1.17)	-0.093*** (-2.71)	-0.093*** (-2.70)	-0.095*** (-2.76)	-0.093*** (-2.70)	-0.091*** (-2.65)	-0.093*** (-2.72)
DFI	0.002 (0.54)	0.001 (0.36)	0.001 (0.45)	0.002 (0.52)	0.001 (0.39)	0.001 (0.45)	0.001 (0.40)	0.001 (0.23)	0.001 (0.31)	0.001 (0.38)	0.001 (0.26)	0.001 (0.30)
RD	0.001* (1.88)	0.001** (2.24)	0.001** (2.01)	0.001* (1.75)	0.001** (2.22)	0.001* (1.88)	0.001 (1.51)	0.001* (1.94)	0.001* (1.65)	0.001 (1.57)	0.001* (1.94)	0.001* (1.71)
Intercept	0.054* (1.85)	0.065** (2.15)	0.068** (2.27)	0.062** (2.08)	0.056* (1.89)	0.070** (2.33)	-0.049 (-1.59)	-0.040 (-1.27)	-0.033 (-1.05)	-0.049 (-1.59)	-0.049 (-1.59)	-0.039 (-1.26)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj.R^2(%)	11.13	11.09	11.19	11.82	11.65	11.87	12.90	12.83		12.81	12.78	12.87
Obs.	5682	5682	5682	5682	5682	5682	5701	5701	5701	5701	5701	5701

Panel B: Alternative Measure of Social network

Dependent variables	ETR1						ETR2						BTD					
CEO Betweenness	0.013*** (4.25)		0.016*** (4.92)				0.013*** (4.20)		0.017*** (4.91)				-0.549*** (-4.34)		-0.633*** (-4.72)			
CFO Betweenness		-0.004 (-1.32)	-0.009*** (-2.80)					-0.005 (-1.41)	-0.010*** (-2.91)					0.043 (0.34)	0.254* (1.90)			
CEO Eigenvector				0.011* (1.69)		0.012* (1.70)				0.003 (0.59)		0.004 (0.72)				0.212 (1.05)		0.213 (1.05)
CFO Eigenvector					0.001 (-0.03)	-0.001 (-0.21)					-0.005 (-1.31)	-0.005 (-1.38)					0.009 (0.07)	-0.004 (-0.03)
SIZE	0.001 (0.71)	0.002* (1.70)	0.001 (0.64)	0.002 (1.48)	0.002 (1.63)	0.002 (1.48)	0.005*** (3.51)	0.006*** (4.55)	0.005*** (3.45)	0.006*** (4.06)	0.006*** (4.15)	0.006*** (4.03)	0.328*** (5.98)	0.274*** (5.12)	0.331*** (6.04)	0.268*** (4.97)	0.275*** (5.15)	0.268*** (4.97)
MB	-0.001** (-2.53)	-0.001*** (-2.60)	-0.001** (-2.47)	-0.001*** (-2.59)	-0.001*** (-2.61)	-0.001*** (-2.60)	-0.001*** (-2.76)	-0.001*** (-2.83)	-0.001*** (-2.71)	-0.001*** (-2.84)	-0.001*** (-2.88)	-0.001*** (-2.87)	0.051*** (4.47)	0.052*** (4.56)	0.050*** (4.43)	0.052*** (4.58)	0.052*** (4.56)	0.052*** (4.58)
LEVER	-0.001 (-0.59)	-0.001 (-0.68)	-0.001 (-0.56)	-0.001 (-0.73)	-0.001 (-0.74)	-0.001 (-0.73)	-0.002 (-1.47)	-0.002 (-1.56)	-0.002 (-1.44)	-0.002 (-1.56)	-0.002 (-1.57)	-0.002 (-1.57)	0.118** (2.32)	0.122** (2.40)	0.117** (2.30)	0.122** (2.40)	0.122** (2.40)	0.122** (2.40)
NOL	0.037*** (9.41)	0.038*** (9.53)	0.037*** (9.45)	0.038*** (9.41)	0.038*** (9.42)	0.038*** (9.40)	0.030*** (7.26)	0.030*** (7.37)	0.030*** (7.30)	0.030*** (7.11)	0.030*** (7.10)	0.029*** (7.09)	-0.782*** (-4.77)	-0.792*** (-4.83)	-0.785*** (-4.79)	-0.795*** (-4.84)	-0.791*** (-4.82)	-0.795*** (-4.84)
ΔNOL	-0.893*** (-3.87)	-0.885*** (-3.83)	-0.906*** (-3.93)	-0.915*** (-3.91)	-0.913*** (-3.90)	-0.915*** (-3.90)	-0.210 (-0.88)	-0.200 (-0.83)	-0.224 (-0.94)	-0.203 (-0.84)	-0.196 (-0.81)	-0.198 (-0.82)	10.537 (1.11)	10.087 (1.06)	10.871 (1.14)	9.896 (1.04)	10.038 (1.05)	9.899 (1.04)
ROA	0.003*** (16.18)	0.003*** (15.84)	0.003*** (16.11)	0.002*** (15.67)	0.002*** (15.64)	0.002*** (15.66)	0.002*** (12.39)	0.002*** (12.05)	0.002*** (12.33)	0.002*** (12.39)	0.002*** (12.43)	0.002*** (12.44)	0.069*** (10.49)	0.071*** (10.82)	0.069*** (10.51)	0.071*** (10.83)	0.071*** (10.80)	0.071*** (10.82)
ESUB	0.019*** (8.19)	0.019*** (8.49)	0.019*** (8.19)	0.020*** (8.66)	0.020*** (8.68)	0.020*** (8.66)	0.013*** (5.75)	0.014*** (6.04)	0.013*** (5.76)	0.014*** (5.93)	0.014*** (5.95)	0.014*** (5.94)	-0.407*** (-4.38)	-0.429*** (-4.62)	-0.407*** (-4.38)	-0.429*** (-4.62)	-0.428*** (-4.61)	-0.429*** (-4.62)
PPE	-0.033*** (-4.25)	-0.032*** (-4.15)	-0.033*** (-4.34)	-0.031*** (-4.07)	-0.031*** (-4.05)	-0.031*** (-4.07)	-0.034*** (-4.27)	-0.033*** (-4.16)	-0.035*** (-4.36)	-0.033*** (-4.07)	-0.033*** (-4.06)	-0.033 (-4.07)	-1.046*** (-3.30)	-1.084*** (-3.42)	-1.031*** (-3.25)	-1.088*** (-3.43)	-1.086*** (-3.42)	-1.088*** (-3.43)
INTAN	-0.040 (-1.32)	-0.035 (-1.17)	-0.041 (-1.37)	-0.036 (-1.19)	-0.035 (-1.16)	-0.036 (-1.19)	-0.080** (-2.53)	-0.076** (-2.38)	-0.082*** (-2.57)	-0.076** (-2.37)	-0.076** (-2.38)	-0.076** (-2.39)	-2.330* (-1.85)	-2.452* (-1.95)	-2.309* (-1.84)	-2.475** (-1.97)	-2.451** (-1.95)	-2.475** (-1.97)
DFI	-0.003 (-0.87)	-0.003 (-0.98)	-0.003 (-0.88)	-0.003 (-1.11)	-0.003 (-1.11)	-0.003 (-1.11)	0.003 (0.99)	0.003 (0.90)	0.003 (0.99)	0.003 (0.85)	0.003 (0.84)	0.003 (0.85)	0.054 (0.44)	0.056 (0.45)	0.055 (0.45)	0.056 (0.45)	0.055 (0.45)	0.056 (0.45)
RD	0.001** (2.05)	0.001** (2.42)	0.001** (2.24)	0.001** (2.34)	0.001** (2.36)	0.001** (2.33)	0.000 (1.23)	0.001 (1.61)	0.001 (1.44)	0.001 (1.30)	0.001 (1.26)	0.001 (1.25)	-0.030* (-1.87)	-0.034** (-2.12)	-0.032** (-1.99)	-0.034** (-2.12)	-0.034** (-2.10)	-0.034** (-2.12)
Intercept	0.058 (1.41)	0.037 (0.91)	0.059 (1.43)	0.011* (1.69)	0.014 (0.32)	0.018 (0.42)	0.048 (1.12)	0.027 (0.63)	0.049 (1.15)	0.003 (1.12)	0.033 (1.13)	0.032 (1.11)	-2.097* (-1.85)	-1.460 (-1.30)	-2.275*** (-2.00)	-1.465 (-1.30)	-1.447 (-1.29)	-1.465 (-1.30)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj.R^2(%)	14.47	14.22	14.57	14.18	14.14	14.17	11.21	10.96	11.33	10.79	10.77	10.79	12.03	11.75	12.07	11.76	11.75	11.75
Obs.	5641	5641	5641	5583	5583	5583	5649	5649	5649	5648	5648	5648	5810	5810	5810	5810	5810	5810

Panel C: Tobit regression												
Dependent variables	ETR1			ETR2			ETR3			ETR4		
CEO Closeness	0.032*** (3.81)		0.030*** (3.57)	0.094*** (4.94)		0.086*** (4.65)	0.050*** (4.40)		0.048*** (4.25)	0.123*** (5.10)		0.117*** (4.86)
CFO Closeness		-0.040*** (-3.68)	-0.037*** (-3.43)		-0.088*** (-3.83)	-0.078*** (-3.45)		-0.038*** (-2.55)	-0.033** (-2.25)		-0.094*** (-3.08)	-0.084*** (-2.77)
SIZE	0.003 (0.79)	0.004 (1.40)	0.002 (0.61)	0.026*** (3.70)	0.031*** (4.59)	0.024*** (3.57)	-0.002 (-0.47)	0.001 (0.29)	-0.002 (-0.55)	0.050*** (5.50)	0.058*** (6.22)	0.048*** (5.32)
MB	-0.002*** (-11.85)	-0.002*** (-11.80)	-0.002*** (-11.76)	-0.002*** (-7.59)	-0.002*** (-7.48)	-0.002*** (-7.59)	-0.002*** (-9.17)	-0.002*** (-9.21)	-0.002*** (-9.26)	-0.002*** (-5.33)	-0.002*** (-4.88)	-0.002*** (-5.22)
LEVER	0.001 (-0.16)	-0.001 (-0.31)	-0.001 (-0.21)	-0.010 (-1.15)	-0.012 (-1.29)	-0.010 (-1.18)	0.001 (0.17)	0.001 (0.03)	0.001 (0.10)	-0.014 (-1.13)	-0.014 (-1.16)	-0.014 (-1.18)
NOL	0.122*** (10.46)	0.125*** (10.74)	0.122*** (10.55)	0.209*** (8.00)	0.215*** (8.23)	0.208*** (8.16)	0.153*** (9.52)	0.159*** (9.77)	0.155*** (9.58)	0.140*** (4.82)	0.152*** (5.02)	0.142*** (4.86)
ΔNOL	-0.973** (-2.01)	-0.990** (-2.05)	-1.008** (-2.10)	-0.226 (-0.24)	-0.264 (-0.29)	-0.263 (-0.28)	-0.661 (-1.06)	-0.636 (-1.01)	-0.646 (-1.03)	-0.240 (-0.20)	-0.258 (-0.19)	-0.260 (-0.21)
ROA	0.007*** (15.97)	0.007*** (15.79)	0.007*** (15.96)	0.010*** (12.59)	0.010*** (11.68)	0.010*** (12.88)	0.009*** (14.39)	0.009*** (4.15)	0.009*** (14.34)	0.014*** (21.15)	0.014*** (15.68)	0.014*** (20.94)
ESUB	0.036*** (6.99)	0.038*** (7.36)	0.036*** (7.04)	0.045*** (4.56)	0.049*** (4.97)	0.045*** (4.65)	0.027*** (4.47)	0.031*** (4.93)	0.028*** (4.56)	0.062*** (4.60)	0.067*** (4.98)	0.064*** (4.68)
PPE	-0.086*** (-4.30)	-0.090*** (-4.46)	-0.089*** (-4.41)	-0.175*** (-3.97)	-0.172*** (-3.97)	-0.171*** (-3.98)	-0.080*** (-3.01)	-0.083*** (-3.08)	-0.082*** (-3.07)	-0.223*** (-4.13)	-0.240*** (-4.14)	-0.232*** (-4.27)
INTAN	-0.115 (-1.45)	-0.112 (-1.41)	-0.125 (-1.57)	-0.284* (-1.65)	-0.328* (-1.87)	-0.326* (-1.88)	-0.161 (-1.47)	-0.132 (-1.20)	-0.143 (-1.31)	-0.442* (-1.95)	-0.403* (-1.74)	-0.440* (-1.93)
DFI	-0.006 (-0.83)	-0.007 (-1.11)	-0.006 (-0.96)	0.015 (1.19)	0.013 (0.98)	0.014 (1.13)	0.005 (0.59)	0.003 (0.32)	0.004 (0.44)	0.018 (1.07)	0.013 (0.80)	0.017 (1.01)
RD	0.002** (1.99)	0.002** (2.44)	0.002** (2.18)	0.001 (0.58)	0.002 (1.21)	0.002 (0.90)	0.002 (1.61)	0.003** (2.08)	0.002* (1.74)	0.002 (0.84)	0.003 (1.13)	0.003 (0.96)
Intercept	-0.230*** (-3.28)	-0.202*** (-2.83)	-0.184*** (-2.58)	-0.892*** (-5.75)	-0.841*** (-5.37)	-0.788*** (-5.13)	-0.329*** (-3.41)	-0.315*** (-3.19)	-0.287*** (-2.94)	-1.674*** (-8.21)	-13.643*** (-7.66)	-1.579*** (-7.58)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Obs.	5699	5699	5699	5706	5706	5706	5682	5682	5682	5701	5701	5701

Dependent variables	ETR1			ETR2			ETR3			ETR4		
CEO Degree	0.033*** (4.06)		0.034*** (4.27)	0.081*** (4.46)		0.084*** (4.64)	0.045*** (4.08)		0.046*** (4.20)	0.085*** (3.72)		0.014*** (3.06)

CFO Degree		-0.027*** (-2.75)	-0.029*** (-3.05)		-0.059*** (-2.90)	-0.068*** (-3.30)		-0.024* (-1.76)	-0.027** (-2.03)		-0.057** (-2.07)	-0.012** (-2.18)
SIZE	0.002 (0.61)	0.005 (1.57)	0.002 (0.56)	0.026*** (3.68)	0.033*** (4.78)	0.025*** (3.61)	-0.002 (-0.51)	0.002 (0.43)	-0.002 (-0.57)	0.052*** (5.59)	0.060*** (6.49)	0.009*** (5.04)
MB	-0.002*** (-11.58)	-0.002 (-11.70)	-0.002*** (-11.77)	-0.002*** (-7.52)	-0.002*** (-7.57)	-0.002*** (-7.45)	-0.002*** (-9.16)	-0.002*** (-9.32)	-0.002*** (-9.25)	-0.002*** (-5.43)	-0.002*** (-5.29)	-0.003*** (-3.25)
LEVER	-0.001 (-0.19)	-0.001 (-0.26)	-0.001 (-0.20)	-0.012 (-1.26)	-0.012 (-1.29)	-0.011 (-1.26)	0.001 (0.03)	0.001 (0.07)	0.001 (0.10)	-0.014 (-1.17)	-0.013 (-1.13)	0.001 (-0.11)
NOL	0.123*** (10.59)	0.125*** (10.72)	0.124*** (10.70)	0.214*** (8.15)	0.216*** (8.18)	0.215*** (8.22)	0.156*** (9.61)	0.158*** (9.73)	0.156*** (9.72)	0.145*** (4.97)	0.155*** (5.18)	0.032*** (5.62)
ΔNOL	-0.991** (-2.04)	-0.993** (-2.05)	-1.004** (-2.09)	-0.245 (-0.26)	-0.249 (-0.27)	-0.277 (-0.30)	-0.668 (-1.06)	-0.650 (-1.03)	-0.670 (-1.07)	-0.253 (-0.21)	-0.265 (-0.21)	-0.323 (-0.99)
ROA	0.007*** (16.01)	0.007*** (15.79)	0.007*** (16.03)	0.010*** (12.50)	0.010*** (11.46)	0.010*** (12.42)	0.009*** (14.33)	0.009*** (14.15)	0.009*** (14.45)	0.014*** (22.67)	0.014*** (19.01)	0.006*** (22.76)
ESUB	0.036*** (7.12)	0.038*** (7.36)	0.036*** (7.20)	0.046*** (4.76)	0.050*** (4.95)	0.047*** (4.84)	0.029*** (4.67)	0.031*** (4.92)	0.029*** (4.74)	0.064*** (4.84)	0.067*** (4.95)	0.016*** (5.18)
PPE	-0.085*** (-4.23)	-0.090*** (-4.44)	-0.088*** (-4.38)	-0.170*** (-3.88)	-0.172*** (-3.95)	-0.173*** (-3.97)	-0.082*** (-3.05)	-0.086*** (-3.18)	-0.081*** (-3.06)	-0.222*** (-3.98)	-0.227*** (-3.93)	-0.065*** (-5.95)
INTAN	-0.116 (-1.47)	-0.111 (-1.40)	-0.114 (-1.46)	-0.280 (-1.62)	-0.286* (-1.65)	-0.300* (-1.73)	-0.152 (-1.38)	-0.145 (-1.30)	-0.142 (-1.31)	-0.428* (-1.85)	-0.454* (-1.93)	-0.169*** (-3.76)
DFI	-0.006 (-0.85)	-0.007 (-1.08)	-0.007 (-1.00)	0.015 (1.18)	0.013 (0.99)	0.013 (1.03)	0.004 (0.47)	0.003 (0.32)	0.003 (0.35)	0.016 (0.97)	0.014 (0.83)	0.001 (-0.12)
RD	0.002** (1.96)	0.002** (2.43)	0.002** (2.12)	0.001 (0.63)	0.002 (1.22)	0.002 (0.88)	0.002 (1.61)	0.003** (2.08)	0.002* (1.74)	0.002 (0.75)	0.003 (1.17)	0.001 (1.13)
Intercept	-0.221*** (-3.14)	-0.228*** (-3.22)	-0.191*** (-2.71)	-0.883*** (-5.68)	-0.898*** (-5.71)	-0.816*** (-5.26)	-0.322*** (-3.32)	-0.341*** (-3.46)	-0.289** (-2.99)	-1.688*** (-8.10)	-1.734*** (-8.12)	-0.120*** (-3.09)
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Obs.	5699	5699	5699	5706	5706	5706	5682	5682	5682	5701	5701	5701

Notes: This table provides the robustness results for the corporate social network effect on the tax aggressive. Our sample covers the period from January 1, 2010 to December 31, 2014. All variables are defined in Appendix A. In Panel A, the alternative measures of tax aggressive is effective tax rate (ETR3) and cash effective tax rate (ETR4), as our dependent variables in Eq. (4). In Panel B, we examine alternative measure of corporate social network. CEO betweenness and CEO eigenvector measured for social network of CEOs, CFO betweenness and CFO eigenvector measured for social network of CFOs, as our independent variables in Eq. (4). In Panel C, we provide the results of the Tobit regression. The t-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively

The Effects of CEO Power on Tax Aggressive

Powerful CEOs can use their power through appointment and recruitment activities. Fracassi and Tate (2012) showed that firms with more powerful chief executive officers (CEOs) are more likely to add new directors with pre-existing network ties to the CEO to minimize search costs or exploit inside information about director quality. We re-estimate the regressions including the interaction between CEO power and corporate social network to examine whether external network connections with powerful CEOs impact tax avoidance activities. The regression model as following:

$$\begin{aligned}\mathrm{TAXAGG}_{i,t} = \alpha_0 &+ \beta_1\mathrm{NETWORK}_{i,t-1} + \beta_2\mathrm{CEOPOWER}_{i;t} + \beta_3\mathrm{NETWORK}_{i;t-1}\mathrm{xCEOPOWER}_{i;t}\\ &+ \beta_4\mathrm{SIZE}_{i;t-1} + \beta_5\mathrm{MB}_{i,t-1} + \beta_6\mathrm{LEVER}_{i,t-1} + \beta_7\mathrm{NOL}_{i,t-1} + \beta_8\Delta\mathrm{NOL}_{i,t-1}\\ &+ \beta_9\mathrm{ROA}_{i,t-1} + \beta_{10}\mathrm{ESUB}_{i,t-1} + \beta_{11}\mathrm{PPE}_{i,t-1} + \beta_{12}\mathrm{INTAN}_{i,t-1} + \beta_{13}\mathrm{DFI}_{i,t-1}\\ &+ \beta_{14}\mathrm{RD}_{i,t-1} + \mathrm{YearDummies} + \mathrm{IndustryDummies} + \varepsilon_{i,t}\end{aligned} \quad (8)$$

where the dependent variable $\mathrm{TAXAGG}_{i,t}$ presents for ETR1, ETR2, BTD in company i in a given year. The main independent variables are corporate social network include social network of CEOs and CFOs, the CEO power dummy variable ($\mathrm{CEOPOWER}_{i;t}$), and their interaction. All variables are defined in Appendix A [The remaining control variables are the same as those in Eq. (4)]

Table 7 presents separate results using closeness (Panel A) and degree (Panel B) to measure the social network of CEOs and CFOs. Our results show that the coefficients of CEOPOWER range from -0.010 to -0.015, significance at least the 5% on ETR1 and ETR2 for both panel, indicating that firms with higher CEO power lead to higher tax aggressive, these results are line with the prior empirical results of Laguir and Staglianò (2014). We examine whether tax aggressive requirement serves as a channel by including the interaction terms, NETWORK × CEOPOWER, in our regression model. Table 7 shows the coefficients between CEO's social network and CEO power, ranging from 0.017 to 0.024, significantly positive impacts on ETR1 and ETR2 at the 1% level as shown in both panel A and panel B, and negative effect on BTD3 at the 10% level as shown in Panel B. These results indicate that the interaction between CEO power and CEO's social network lead a lower tax aggressive. Our finding also presents the coefficient of interaction between CEO power and CFO's social network is insignificant different from zero, implying that CEO power does not affect the relation between CFO's social network and tax aggressive.

Table 7: The effects of CEO power on tax aggressive

Dependent variables	ETR1		ETR2		BTD	
Panel A: Social network measured by Closeness						
CEO Closeness	0.002 (0.45)		0.008* (1.84)		-0.570*** (-3.64)	
CFO Closeness		-0.013** (-2.37)		-0.015*** (-2.63)		0.483** (2.11)
CEOPOWER	-0.015*** (-2.98)	0.008 (0.86)	-0.014*** (-2.73)	0.003 (0.34)	0.049 (0.27)	-0.168 (-0.44)
CEO Closeness x CEO POWER	0.024*** (3.69)		0.017*** (2.56)		-0.080 (-0.31)	
CFO Closeness x CEO POWER		-0.010 (-1.03)		-0.009 (-0.82)		0.223 (0.55)
SIZE	0.001 (0.88)	0.002 (1.21)	0.005*** (3.24)	0.005*** (3.82)	0.331*** (5.99)	0.287*** (5.31)
MB	-0.001*** (-2.65)	-0.001*** (-2.60)	-0.001*** (-2.83)	-0.001*** (-2.81)	0.052*** (4.55)	0.052*** (4.53)

LEVER	-0.001 (-0.69)	-0.001 (-0.70)	-0.002 (-1.56)	-0.002 (-1.62)	0.116** (2.28)	0.123** (2.42)
NOL	0.037*** (9.22)	0.038*** (9.46)	0.029*** (6.98)	0.030*** (7.29)	-0.753*** (-4.59)	-0.799*** (-4.87)
ΔNOL	-0.888*** (-3.85)	-0.910*** (-3.94)	-0.189 (-0.79)	-0.211 (-0.88)	10.254 (1.08)	10.860 (1.14)
ROA	0.003*** (15.87)	0.002*** (15.82)	0.002*** (12.19)	0.002*** (12.01)	0.069*** (10.55)	0.071*** (10.85)
ESUB	0.019*** (8.24)	0.020*** (8.54)	0.013*** (5.72)	0.014*** (6.10)	-0.384*** (-4.11)	-0.431*** (-4.64)
PPE	-0.032*** (-4.12)	-0.032*** (-4.19)	-0.033*** (-4.10)	-0.033*** (-4.15)	-1.050*** (-3.31)	-1.049*** (-3.31)
INTAN	-0.034 (-1.12)	-0.036 (-1.20)	-0.076** (-2.39)	-0.077** (-2.40)	-2.252* (-1.79)	-2.412* (-1.92)
DFI	-0.003 (-1.01)	-0.003 (-1.13)	0.003 (0.94)	0.002 (0.78)	0.048 (0.39)	0.071 (0.58)
RD	0.001** (2.17)	0.001*** (2.61)	0.001 (1.07)	0.001 (1.59)	-0.027* (-1.72)	-0.036** (-2.25)
Panel B: Social network measured by Degree						
CEO Degree	0.004 (1.14)		0.005 (1.22)		-0.247 (-1.57)	
CFO Degree		-0.010** (-2.07)		-0.012** (-2.39)		0.370* (1.91)
CEOPOWER	-0.010** (-2.16)	-0.002 (-0.26)	-0.014*** (-2.87)	-0.006 (-0.80)	0.249 (1.37)	0.098 (0.34)
CEO Degree x CEO POWER	0.018*** (2.84)		0.019*** (2.91)		-0.490* (-1.92)	
CFO Degreex CfO POWER		0.001 (0.12)		0.002 (0.25)		-0.073 (-0.22)
SIZE	0.001 (0.76)	0.002 (1.45)	0.005*** (3.29)	0.006*** (4.07)	0.315*** (5.67)	0.276*** (5.11)
MB	-0.001*** (-2.60)	-0.001*** (-2.60)	-0.001*** (-2.81)	-0.001*** (-2.80)	0.051*** (4.51)	0.052*** (4.53)
LEVER	-0.001 (-0.67)	-0.001 (-0.69)	-0.002 (-1.59)	-0.002 (-1.61)	0.121** (2.37)	0.123** (2.41)
NOL	0.037*** (9.33)	0.038*** (9.45)	0.030*** (7.15)	0.030*** (7.30)	-0.784*** (-4.78)	-0.799*** (-4.87)
ΔNOL	-0.897*** (-3.88)	-0.905*** (-3.91)	-0.199 (-0.83)	-0.209 (-0.87)	10.590 (1.11)	10.505 (1.10)
ROA	0.003*** (15.98)	0.002*** (15.78)	0.002*** (12.20)	0.002*** (11.98)	0.070*** (10.64)	0.071*** (10.87)
ESUB	0.019*** (8.24)	0.020*** (8.54)	0.014*** (5.80)	0.014*** (6.11)	-0.407*** (-4.37)	-0.434*** (-4.67)
PPE	-0.031*** (-4.07)	-0.032*** (-4.14)	-0.032*** (-4.06)	-0.033*** (-4.11)	-1.063*** (-3.35)	-1.062*** (-3.35)
INTAN	-0.036 (-1.20)	-0.035 (-1.14)	-0.076** (-2.39)	-0.075** (-2.35)	-2.409* (-1.91)	-2.446* (-1.94)
DFI	-0.003 (-0.98)	-0.003 (-1.07)	0.003 (0.90)	0.003 (0.82)	0.053 (0.43)	0.066 (0.54)
RD	0.001** (2.16)	0.001*** (2.58)	0.001 (1.15)	0.001 (1.59)	-0.030* (-1.86)	-0.036** (-2.27)

Notes: This table provides details of the effects of CEO power on the tax aggressiveness. The regression model is:

$$\begin{aligned} TAXAGG_{i,t} = \alpha_0 &+ \beta_1 NETWORK_{i,t-1} + \beta_2 CEOPOWER_{i;t} + \beta_3 NETWORK_{i;t-1} x CEOPOWER_{i;t} + \beta_4 SIZE_{i;t-1} + \beta_5 MB_{i,t-1} + \beta_6 LEVER_{i,t-1} \\ &+ \beta_7 NOL_{i,t-1} + \beta_8 \Delta NOL_{i,t-1} + \beta_9 ROA_{i,t-1} + \beta_{10} ESUB_{i,t-1} + \beta_{11} PPE_{i,t-1} + \beta_{12} INTAN_{i,t-1} + \beta_{13} DFI_{i,t-1} + \beta_{14} RD_{i,t-1} \\ &+ YearDummies + IndustryDummies + \varepsilon_{i,t} \end{aligned}$$

Where the dependent variable is tax aggressive ($TAX_{i,t}$) present for the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. The main independent variable is corporate social network includes

CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). CEOPOWER is dummy variable that equals one is the CEO is chair, zero otherwise. $NETWORK_{i;t-1}xCEO\_POWER_{i;t}$ is the interaction between social network and CEO power. All variables are defined in Appendix A. The *t*-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively

The Effects of CEOs Background on Tax Aggressive

The financial and/or accountant background of CEOs can play an important role in determining the quality of financial reporting, it helps them keep a deeper understanding of financial and accounting issues, they therefore can make proper accounting decisions and improve the financial reporting process. Custódio and Metzger (2014) show the statistic that 41% of the CEOs have previous experience in the financial industry or in a financial role. This study showed the significant effect of financial expert CEOs on firms' financial policies, such as cash holdings, debts, and share purchases. CEOs who have a professional qualification in finance and/or accounting do not participate directly in overseeing the accounting process and the preparation of financial statements but they may set the tone from the top and impact to the CFOs decision (Feng et al., 2011). However, to the best of our knowledge, the impact of CEOs' financial or accountant experience on tax aggressive are unexplored. In this section, we examine whether CEOs with financial and/or accounting backgrounds interact with corporate social network are associate with tax aggressive. We re-estimate our regressions model as following:

$$\begin{aligned} TAXAGG_{i,t} = \alpha_0 &+ \beta_1 NETWORK_{i,t-1} + \beta_2 CEOfin\_acc_{i;t} + \beta_3 NETWORK_{i;t-1}xCEOfin\_acc_{i;t} \\ &+ \beta_4 SIZE_{i;t-1} + \beta_5 MB_{i,t-1} + \beta_6 LEVER_{i,t-1} + \beta_7 NOL_{i,t-1} + \beta_8 \Delta NOL_{i,t-1} \\ &+ \beta_9 ROA_{i,t-1} + \beta_{10} ESUB_{i,t-1} + \beta_{11} PPE_{i,t-1} + \beta_{12} INTAN_{i,t-1} + \beta_{13} DFI_{i,t-1} \\ &+ \beta_{14} RD_{i,t-1} + YearDummies + IndustryDummies + \varepsilon_{i,t} \end{aligned} \quad (9)$$

where the dependent variable $TAXAGG_{i,t}$ presents for ETR1, ETR2, BTD in company i in a given year. The main independent variables are corporate social network include social network of CEOs and CFOs, CEO has financial and/or accounting background is as a dummy variable ($CEOfin\_acc_{i;t}$), and their interaction. All variables are defined in Appendix A [The remaining control variables are the same as those in Eq. (4)]

In Panel A of Table 8, the coefficients interaction between $CEOfin_{acc}$ and CFO closeness are significantly negative at 10% level on ETR1, and positive impact on BTD with the coefficient 1.268 at the10% level. As shown in panel B, we also find the coefficient of CFO degree and $CEOfin_{acc}$ is -0.027, negative significant at the 10% level on ETR1. These results indicate that firms with financial and/or accountant CEOs connect with CFO's social network are more likely implement tax aggressive. Our empirical results are consistent with prior researches and clarify more detail the interaction between CEO and CFO on tax avoidance (Hsieh et al., 2018).

Table 8: The effects of CEO background on tax aggressive

Dependent variables	ETR1		ETR2		BTD	
Panel A: Social network measured by Closeness						
CEO Closeness	0.010*** (3.13)		0.013*** (3.78)		-0.586*** (-4.43)	
CFO Closeness		-0.014*** (-2.94)		-0.017*** (-3.34)		0.450** (2.27)
CEOfin_acc	-0.001 (-0.10)	0.021 (1.36)	-0.003 (-0.31)	0.014 (0.84)	-0.017 (-0.05)	-1.166* (-1.81)
CEO Closeness x CEOfin_acc	0.001 (-0.02)		0.014 (1.05)		-0.207 (-0.4)	
CFO Closeness x CEOfin_acc		-0.028* (-1.65)		-0.011 (-0.64)		1.268* (1.80)
SIZE	0.001	0.002	0.005***	0.005***	0.331***	0.290***

	(0.82)	(1.21)	(3.32)	(4.04)	(6.04)	(5.41)
MB	-0.001*** (-2.60)	-0.001*** (-2.60)	-0.001*** (-2.78)	-0.001*** (-2.80)	0.052*** (4.54)	0.051*** (4.52)
LEVER	-0.001 (-0.61)	-0.001 (-0.67)	-0.002 (-1.52)	-0.002 (-1.62)	0.116** (2.29)	0.121** (2.39)
NOL	0.037*** (9.20)	0.037*** (9.45)	0.029*** (7.02)	0.030*** (7.33)	-0.755*** (-4.60)	-0.795*** (-4.85)
ΔNOL	-0.889*** (-3.85)	-0.899*** (-3.89)	-0.191 (-0.80)	-0.210 (-0.87)	10.322 (1.08)	10.429 (1.09)
ROA	0.003*** (16.05)	0.003*** (15.84)	0.002*** (12.29)	0.002*** (11.98)	0.069*** (10.55)	0.071*** (10.83)
ESUB	0.019*** (8.12)	0.019*** (8.51)	0.013*** (5.66)	0.014*** (6.11)	-0.384*** (-4.12)	-0.430*** (-4.63)
PPE	-0.031*** (-4.06)	-0.032*** (-4.19)	-0.032*** (-4.02)	-0.033*** (-4.15)	-1.055*** (-3.33)	-1.046*** (-3.30)
INTAN	-0.036 (-1.19)	-0.037 (-1.24)	-0.080** (-2.51)	-0.078** (-2.45)	-2.211* (-1.76)	-2.349* (-1.87)
DFI	-0.003 (-0.89)	-0.003 (-1.10)	0.003 (1.03)	0.002 (0.79)	0.045 (0.37)	0.068 (0.55)
RD	0.001** (2.11)	0.001*** (2.68)	0.001 (1.05)	0.001 (1.62)	-0.028* (-1.72)	-0.038** (-2.36)
Panel B: Social network measured by Degree						
CEO Degree	0.012*** (3.64)		0.012*** (3.65)		-0.384*** (-2.89)	
CFO Degree		-0.008* (-1.85)		-0.011** (-2.54)		0.345** (2.12)
CEOfin_acc	0.009 (0.87)	0.019 (1.46)	0.010 (0.91)	0.007 (0.54)	0.208 (0.53)	-0.070 (-0.14)
CEO Degree x CEOfin_acc	-0.017 (-1.35)		-0.008 (-0.64)		-0.552 (-1.07)	
CFO Degreex CEOfin_acc		-0.027* (-1.83)		-0.004 (-0.24)		-0.051 (-0.09)
SIZE	0.001 (0.69)	0.002 (1.45)	0.005*** (3.34)	0.006*** (4.29)	0.318*** (5.77)	0.274*** (5.11)
MB	-0.001*** (-2.55)	-0.001*** (-2.60)	-0.001*** (-2.73)	-0.001*** (-2.79)	0.051*** (4.48)	0.052*** (4.52)
LEVER	-0.001 (-0.61)	-0.001 (-0.64)	-0.002 (-1.54)	-0.002 (-1.61)	0.120*** (2.70)	0.123** (2.42)
NOL	0.037*** (9.34)	0.037*** (9.39)	0.030*** (7.21)	0.030*** (7.32)	-0.789*** (-4.81)	-0.801*** (-4.88)
ΔNOL	-0.905*** (-3.92)	-0.890*** (-3.85)	-0.211 (-0.88)	-0.208 (-0.86)	10.667 (1.12)	10.589 (1.11)
ROA	0.003*** (16.09)	0.002*** (15.81)	0.002*** (12.27)	0.002*** (11.96)	0.070*** (10.61)	0.071*** (10.88)
ESUB	0.019*** (8.20)	0.019*** (8.52)	0.014*** (5.76)	0.014*** (6.13)	-0.405*** (-4.35)	-0.435*** (-4.68)
PPE	-0.031*** (-4.03)	-0.032*** (-4.14)	-0.032*** (-4.01)	-0.033*** (-4.12)	-1.069*** (-3.37)	-1.060*** (-3.34)
INTAN	-0.035 (-1.15)	-0.035 (-1.15)	-0.078** (-2.43)	-0.076** (-2.38)	-2.282* (-1.81)	-2.440* (-1.94)
DFI	-0.003 (-0.91)	-0.003 (-1.04)	0.003 (0.99)	0.003 (0.81)	0.045 (0.36)	0.066 (0.54)
RD	0.001** (2.08)	0.001*** (2.68)	0.001 (1.07)	0.001 (1.61)	-0.029* (-1.82)	-0.036** (-2.26)

Notes: This table provides details of the effects of CEOs have financial and/or accounting background on the tax aggressiveness. The regression model is:

$$TAXAGG_{i,t} = \alpha_0 + \beta_1 NETWORK_{i,t-1} + \beta_2 CEOfin\_acc_{i;t} + \beta_3 NETWORK_{i;t-1} x CEOfin\_acc_{i;t} + \beta_4 SIZE_{i;t-1} + \beta_5 MB_{i,t-1} + \beta_6 LEVER_{i,t-1} + \beta_7 NOL_{i,t-1} + \beta_8 \Delta NOL_{i,t-1} + \beta_9 ROA_{i,t-1} + \beta_{10} ESUB_{i,t-1} + \beta_{11} PPE_{i,t-1} + \beta_{12} INTAN_{i,t-1} + \beta_{13} DFI_{i,t-1} + \beta_{14} RD_{i,t-1} + YearDummies + IndustryDummies + \varepsilon_{i,t}$$

Where the dependent variable is tax aggressive ($TAX_{i,t}$) present for the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. The main independent variable is corporate social network includes CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). *CEOfin_acc* is dummy variable that equals one is the CEO have accounting or financial background, zero otherwise. $NETWORK_{i;t-1}xCEO\text{fin\_acc}_{i;t}$ is the interaction between social network and CEOs has financial and/or accounting background. All variables are defined in Appendix A. The *t*-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

The Effects of CFOs Background on Tax Aggressive

Prior research has shown that CFOs are typically in charge of financial planning, budgeting, internal control, and financial reporting quality (Gore et al., 2007; Kaufman, 2003). In addition, CFOs have significant impact on accounting decisions such as choosing accounting methods and making accounting adjustments (Geiger & North, 2006; Gore et al., 2007; Mian, 2001). Previous researches also assumed that accounting background and related experience can influence financial reporting and CFO performance. A recent survey reports that, among the CFOs of Fortune 100 firms, 32% had worked as controllers and 9% had served in corporate audit (Russell Reynolds Associates, 2012) [More details refer at: Where do CFOs come from? Profiles and career patterns of Fortune 100 chief financial officers. Available at: [(http://www.russellreynolds.com/content/where-do-chief-financial-officers-cfos-come-from)]. Hoitash et al. (2016) show that firms with accountant CFOs are more risk averse, accountant CFOs are negatively (positive) impact with firm value in high-growth (low-growth) industries. We examine whether the interaction between CFOs' accounting backgrounds (accountant CFOs) and corporate social network are associate with tax aggressive. We re-estimate our regressions model as following:

$$\begin{aligned}\mathrm{TAXAGG}_{i,t} = {} & \alpha_0 + \beta_1\mathrm{NETWORK}_{i,t-1} + \beta_2\mathrm{CFOfin\_acc}_{i;t} + \beta_3\mathrm{NETWORK}_{i;t-1}\mathrm{xCFOfin\_acc}_{i;t} \\ & + \beta_4\mathrm{SIZE}_{i;t-1} + \beta_5\mathrm{MB}_{i,t-1} + \beta_6\mathrm{LEVER}_{i,t-1} + \beta_7\mathrm{NOL}_{i,t-1} + \beta_8\Delta\mathrm{NOL}_{i,t-1} + \beta_9\mathrm{ROA}_{i,t-1} \\ & + \beta_{10}\mathrm{ESUB}_{i,t-1} + \beta_{11}\mathrm{PPE}_{i,t-1} + \beta_{12}\mathrm{INTAN}_{i,t-1} + \beta_{13}\mathrm{DFI}_{i,t-1} + \beta_{14}\mathrm{RD}_{i,t-1} + \mathrm{YearDummies} \\ & + \mathrm{IndustryDummies} + \varepsilon_{i,t}\end{aligned} \quad (10)$$

where the dependent variable $TAXAGG_{i,t}$ presents for ETR1, ETR2, BTD in company i in a given year. The main independent variables are corporate social network include social network of CEOs and CFOs. CFOs has both financial and accounting background set as a dummy variable ($CFOfin\_acc_{i;t}$), and their interaction. All variables are defined in Appendix A [The remaining control variables are the same as those in Eq. (4)]

Table 9 shows that $CFOfin\_acc$ are significantly impact on our tax aggressive measures. The coefficients interaction between $CFOfin\_acc$ and CFO's corporate social network are significantly negative, ranging from -0.023 to -0.029, significant at least the 5% level on ETR1 and ETR2 for both panel A and panel B. We didn't find any evidence about the interaction between $CFOfin\_acc$ and CEO's corporate social network. As a result, CFOs which higher social network and accountant background are lower the effective tax rate, leading more like engage corporate tax aggressive.

Table 9: The effects of CFO background on tax aggressive

Dependent variables	ETR1		ETR2		BTD	
Panel A: Social network measured by Closeness						
CEO Closeness	0.004 (0.61)		0.008 (1.27)		-0.749*** (-3.23)	
CFO Closeness		0.002 (0.15)		0.005 (0.48)		0.669 (1.60)
CFOfin_acc	-0.009 (-1.62)	0.018* (1.69)	-0.003 (-0.30)	0.030*** (2.65)	-0.062 (-0.32)	0.182 (0.40)
CEO Closeness x CFOfin_acc	0.010		0.009		0.208	

	(1.47)		(1.27)		(0.78)	
CFO Closeness x CFOfin_acc		-0.023** (-2.05)		-0.029** (-2.45)		-0.142 (-0.30)
SIZE	0.001 (0.86)	0.002 (1.48)	0.005*** (3.28)	0.006*** (4.14)	0.329*** (6.00)	0.284*** (5.28)
MB	-0.001*** (-2.61)	-0.001*** (-2.63)	-0.001*** (-2.81)	-0.001*** (-2.84)	0.052*** (4.54)	0.052*** (4.53)
LEVER	-0.001 (-0.64)	-0.001 (-0.76)	-0.002 (-1.52)	-0.002* (-1.68)	0.115** (2.26)	0.123** (2.42)
NOL	0.037*** (9.22)	0.038*** (9.55)	0.029*** (7.00)	0.030*** (7.38)	-0.752*** (-4.58)	-0.800*** (-4.88)
ΔNOL	-0.882*** (-3.82)	-0.898*** (-3.89)	-0.182 (-0.76)	-0.190 (-0.79)	10.396 (1.09)	11.031 (1.15)
ROA	0.003*** (16.10)	0.002*** (15.77)	0.002*** (12.34)	0.002*** (11.93)	0.069*** (10.57)	0.071*** (10.84)
ESUB	0.019*** (8.10)	0.020*** (8.58)	0.013*** (5.63)	0.014*** (6.17)	-0.384*** (-4.12)	-0.429*** (-4.62)
PPE	-0.031*** (-4.09)	-0.032*** (-4.15)	-0.033*** (-4.09)	-0.033*** (-4.13)	-1.059*** (-3.34)	-1.052*** (-3.32)
INTAN	-0.036 (-1.18)	-0.037 (-1.21)	-0.078** (-2.46)	-0.079** (-2.47)	-2.265* (-1.80)	-2.435* (-1.93)
DFI	-0.002 (-0.82)	-0.003 (-1.12)	0.003 (0.98)	0.002 (0.68)	0.046 (0.38)	0.068 (0.55)
RD	0.001** (2.19)	0.001*** (2.66)	0.001 (1.08)	0.001* (1.65)	-0.027* (-1.67)	-0.036** (-2.23)
Panel B: Social network measured by Degree						
CEO Degree	0.004 (0.74)		0.004 (0.74)		-0.514** (-2.23)	
CFO Degree		0.007 (0.91)		0.003 (0.42)		0.278 (0.93)
CFOfin_acc	-0.008 (-1.57)	0.015** (2.05)	-0.002 (-0.41)	0.021** (2.39)	-0.031 (-0.16)	-0.037 (-0.12)
CEO Degree x CFOfin_acc	0.010 (1.45)		0.011 (1.57)		0.136 (0.51)	
CFO Degreex CFOfin_acc		-0.023*** (-2.55)		-0.020** (-2.11)		0.091 (0.26)
SIZE	0.001 (0.70)	0.002* (1.65)	0.004*** (3.22)	0.006*** (4.28)	0.316*** (5.73)	0.273*** (5.08)
MB	-0.001** (-2.54)	-0.001*** (-2.63)	-0.001*** (-2.75)	-0.001*** (-2.83)	0.051*** (4.48)	0.052*** (4.53)
LEVER	-0.001 (-0.66)	-0.001 (-0.75)	-0.002 (-1.56)	-0.002* (-1.65)	0.119** (2.34)	0.123** (2.42)
NOL	0.037*** (9.35)	0.038*** (9.50)	0.030*** (7.20)	0.030*** (7.33)	-0.787*** (-4.80)	-0.800*** (-4.88)
ΔNOL	-0.901*** (-3.90)	-0.891*** (-3.85)	-0.202 (-0.84)	-0.192 (-0.80)	10.811 (1.13)	10.457 (1.09)
ROA	0.003*** (16.14)	0.002*** (15.73)	0.002*** (12.35)	0.002*** (11.92)	0.070*** (10.61)	0.071*** (10.88)
ESUB	0.019*** (8.19)	0.020*** (8.58)	0.013*** (5.76)	0.014*** (6.15)	-0.404*** (-4.34)	-0.434*** (-4.67)
PPE	-0.031*** (-4.07)	-0.032*** (-4.13)	-0.033*** (-4.07)	-0.033*** (-4.12)	-1.073*** (-3.38)	-1.063*** (-3.35)
INTAN	-0.037 (-1.24)	-0.035 (-1.16)	-0.079** (-2.49)	-0.077** (-2.40)	-2.370* (-1.88)	-2.455* (-1.95)
DFI	-0.003 (-0.88)	-0.003** (-2.09)	0.003 (0.92)	0.002 (0.71)	0.045 (0.36)	0.066 (0.53)
RD	0.001** (2.17)	0.001*** (2.60)	0.001 (1.14)	0.001 (1.59)	-0.029* (-1.79)	-0.036** (-2.26)

Notes: This table provides details of the effects of CFO has both financial and accounting background on the tax aggressiveness. The regression model is:

$$TAXAGG_{i,t} = \alpha_0 + \beta_1 NETWORK_{i,t-1} + \beta_2 CFOfin\_acc_{i;t} + \beta_3 NETWORK_{i;t-1} x CFOfin\_acc_{i;t} + \beta_4 SIZE_{i;t-1} + \beta_5 MB_{i,t-1} + \beta_6 LEVER_{i,t-1} + \beta_7 NOL_{i,t-1} + \beta_8 \Delta NOL_{i,t-1} + \beta_9 ROA_{i,t-1} + \beta_{10} ESUB_{i,t-1} + \beta_{11} PPE_{i,t-1} + \beta_{12} INTAN_{i,t-1} + \beta_{13} DFI_{i,t-1} + \beta_{14} RD_{i,t-1} + YearDummies + IndustryDummies + \varepsilon_{i,t}$$

Where the dependent variable is tax aggressive ($TAX_{i,t}$) present for the long term effective tax rate (ETR1); long term cash effective tax rate (ETR2) and permanent book-tax differences (BTD) for company i in year t. The main independent variable is corporate social network includes CEO's social network (CEO closeness and CEO degree) and CFO's social network (CFO closeness and CFO degree). *CFOfin_acc* is dummy variable that equals one is the CFO have both accounting and financial background, zero otherwise. $NETWORK_{i;t-1} xCFO\mathrm{fin\_acc}_{i;t}$ is the interaction between social network and CFO has both financial and accounting background. All variables are defined in Appendix A. The *t*-values examine whether the regression coefficient is significantly different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

CONCLUSION

This study examines whether and how corporate social network are related to corporate tax aggressive. We constructed a comprehensive social network of Taiwan listed companies that arise from education background, prior experience, and other social activities such as training course, business forum, and meetings in the period span 5 year from 2010 to 2014. Our empirical results show that social network, is a statistically and economically meaningful determinant of the corporate tax aggressive. We also show the consistent evidence by using instrumental regression and robustness test to support for our main findings. Those funding imply that firms share tax strategies by social networks ties.

In summary, our study extents the literature by linking the social network to firm, a better understanding of the effect of network ties on tax aggressive. Our research contributes to prior literatures by following ways: Firstly, this paper is the first to document that individual's social capital associated with the position within the social network affects outcomes of tax avoidance; Secondly, we investigate both the impact of CEO and CFO network ties to corporate tax avoidance, CEOs/CFOs occupying higher places in the overall network hierarchy are less (more) perform tax avoidance strategies; Finally, our finding show the evidence of the interaction between powerful, academic background and network ties impact to tax evasion.

APPENDIX A: DEFINITIONS OF VARIABLES

Variable	Description
Tax aggressiveness	
Long term effective tax rate (ETR1)	Total income tax expense over last three year scaled by total pre-tax income over the same period of time based on Dyreng et al. (2008); Minnick and Noga (2010).
Long term cash effective tax rate (ETR2)	Total refund income tax to chief financial officer over last three year scaled by total pre-tax income over the same period of time based on Dyreng et al. (2008); Minnick and Noga (2010).
Effective tax rate (ETR3)	Total income tax expense scaled by total pre-tax income over the same based on Richardson and Lanis, (2007); Huseynov and Klamm (2012); Lennox, et al. (2013).
Cash effective tax rate (ETR4)	Total refund income tax to chief financial officer scaled by total pre-tax income based on Dyreng et al.,2008;Khurana and Moser2009; Chen et al. (2010).
Permanent book-tax differences (BTD)	Pre-Tax Income-(Income tax expense/ Law tax rate)]- [(Deferred income tax expense/law tax rate)] based on Frank, Lynch, and Rego (2009); Lennox, et al. (2013).
Social Network	
CEO closeness	CEO's social network which measured by closeness. CEO closeness set as an indicator that equal one if the value is higher than mean of it by each year, zero otherwise.
CFO closeness	CFO's social network which measured by closeness. CFO closeness set as an indicator that equal one if the value is higher than mean of it by each year, zero otherwise.
CEO degree	CEO's social network which measured by degree. CEO degree set as an indicator variable that equal one if the value is higher than mean of it by each year, zero otherwise.
CFO degree	CFO's social network which measured by degree. CFO degree set as an indicator variable that equal one if the value is higher than mean of it by each year, zero otherwise.
CEO betweenness	CEO's social network which measured by betweenness. CEO betweenness set as an indicator that equal one if the value is higher than mean of it by each year, zero otherwise.
CFO betweenness	CFO's social network which measured by betweenness. CFO betweenness set as an indicator variable that equal one if the value is higher than mean of it by each year, zero otherwise.
CEO eigenvector	CEO's social network which measured by eigenvector. CEO eigenvector set as an indicator variable that equal one if the value is higher than mean of it by each year, zero otherwise.
CFO eigenvector	CFO's social network which measured by eigenvector. CFO eigenvector set as an indicator variable that equal one if the value is higher than mean of it by each year, zero otherwise.
Firm characteristics	
SIZE	Logarithm of market value of equity for a firm at the beginning of a year.
MB	Market-to-book ratio for a firm in a given year, measured as market value of equity divided by book value of equity.
LEVER	Leverage for a firm in a given year, measured by long-term debt scaled by lagged total assets.

NOL	Indicator variable that equals 1 if tax loss carry forward is positive at the beginning of year t, and 0 otherwise.
ΔNOL	Change in loss carry forward for firm in a given year, scaled by lagged total assets.
ROA	Return on assets for a firm in a given year, scaled by total assets.
ESUB	Equity income in earning for a firm in a given year, scaled by lagged total assets.
PPE	Property, plant, and equipment for a firm in a year, scaled by lagged total assets.
INTAN	Intangible assets for a firm in a given year, scaled by lagged total assets.
DFI	Foreign pre-tax income scaled by lagged total assets; Missing values of foreign pre-tax income are set to zero.
RD	Natural logarithm of research and development activities expenditure. Missing values of research and development expense are set to zero.
Governance characteristics	
ID	The percentage of directors who are independent.
BS	The number of board members.
IO	The percentage of institution ownership.
BSO	The percentage board and supervisor ownership.
CEOPOWER	Equal to one if the CEO is chair, zero otherwise.
CEOfin_acc	Equal to one if CEO have a financial and/or accountant background, zero otherwise.
CFOfin_acc	Equal to one if CFO have both accountant and financial background, zero otherwise.

REFERENCES

Adler, P. S. (2001). Market, hierarchy, and trust: The knowledge economy and the future of capitalism. *Organization science, 12*(2), 215-234.

Adler, P. S., & Kwon, S.-W. (2002). Social capital: Prospects for a new concept. *Academy of management review, 27*(1), 17-40.

Ahuja, G., & Katila, R. (2004). Where do resources come from? The role of idiosyncratic situations. *Strategic Management Journal, 25*(8-9), 887-907.

Amit, R., & Schoemaker, P. J. H. (1993). Strategic assets and organizational rent. *Strategic Management Journal, 14*(1), 33-46.

Athanassiou, N., & Nigh, D. (2002). The impact of the top management team's international business experience on the firm's internationalization: Social networks at work. *Management International Review, 42*(2), 157-182.

Austin, C. R., & Wilson, R. J. (2015). *Are Reputational Costs a Determinant of Tax Avoidance?* Paper presented at the 2013 American Taxation Association Midyear Meeting: Tax Avoidance in an International Setting.

Baker, W. E., & Faulkner, R. R. (2004). Social networks and loss of capital. *Social Networks, 26*(2), 91-111.

Bebchuk, L. A., & Fried, J. M. (2006). Pay without performance: Overview of the issues. *Academy of Management Perspectives, 20*(1), 5-24.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy, 70*(5, Part 2), 9-49.

Belliveau, M. A., O'Reilly III, C. A., & Wade, J. B. (1996). Social capital at the top: Effects of social similarity and status on CEO compensation. *Academy of Management Journal, 39*(6), 1568-1593.

Benton, R. A. (2017). The Decline of Social Entrenchment: Social Network Cohesion and Board Responsiveness to Shareholder Activism. *Organization science, 28*(2), 262-282.

Bizjak, J., Lemmon, M., & Whitby, R. (2009). Option backdating and board interlocks. *The Review of Financial Studies, 22*(11), 4821-4847.

Brown, J. L. (2011). The Spread of Aggressive Corporate Tax Reporting: A Detailed Examination of the Corporate-Owned Life Insurance Shelter. *The Accounting Review, 86*(1), 23-57.

Brown, J. L., & Drake, K. D. (2014). Network Ties Among Low-Tax Firms. *The Accounting Review, 89*(2), 483-510.

Burt, R. S. (1997). A note on social capital and network content. *Social Networks, 19*(4), 355-373.

Butler, A. W., & Gurun, U. G. (2012). Educational networks, mutual fund voting patterns, and CEO compensation. *The Review of Financial Studies, 25*(8), 2533-2562.

Cai, Y., & Sevilir, M. (2012). Board connections and M&A transactions. *Journal of Financial Economics, 103*(2), 327-349.

Cao, Q., Maruping, L. M., & Takeuchi, R. (2006). Disentangling the Effects of CEO Turnover and Succession on Organizational Capabilities: A Social Network Perspective. *Organization science, 17*(5), 563-576.

Carrie R. Leana, I., & Buren, H. J. V. (1999). Organizational Social Capital and Employment Practices. *Academy of management review, 24*(3), 538-555.

Chen, S., Chen, X., Cheng, Q., & Shevlin, T. (2010). Are family firms more tax aggressive than non-family firms? *Journal of Financial Economics, 95*(1), 41-61.

Chi, S. S., Pincus, M., & Teoh, S. H. (2014). Mispricing of Book-Tax Differences and the Trading Behavior of Short Sellers and Insiders. *The Accounting Review, 89*(2), 511-543.

Chidambaran, N., Kedia, S., & Prabhala, N. (2011). CEO director connections and corporate fraud.

Coff, R. W. (2002). Human Capital, Shared Expertise, and the Likelihood of Impasse in Corporate Acquisitions. *Journal of Management, 28*(1), 107-128.

Cohen, L., Frazzini, A., & Malloy, C. (2008). The Small World of Investing: Board Connections and Mutual Fund Returns. *Journal of Political Economy, 116*(5), 951-979.
Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly, 17*(1), 1-25.
Collins, C. J., & Clark, K. D. (2003). Strategic Human Resource Practices, Top Management Team Social Networks, and Firm Performance: The Role of Human Resource Practices in Creating Organizational Competitive Advantage. *Academy of Management Journal, 46*(6), 740-751.
Cremers, K. J. M., & Nair, V. B. (2005). Governance Mechanisms and Equity Prices. *The Journal of Finance, 60*(6), 2859-2894.
Custódio, C., & Metzger, D. (2014). Financial expert CEOs: CEO′s work experience and firm′s financial policies. *Journal of Financial Economics, 114*(1), 125-154.
Davis, G. F., & Greve, H. R. (1997). Corporate Elite Networks and Governance Changes in the 1980s. *American Journal of Sociology, 103*(1), 1-37.
Del Guercio, D., Seery, L., & Woidtke, T. (2008). Do boards pay attention when institutional investor activists "just vote no"? *Journal of Financial Economics, 90*(1), 84-103.
Desai, M. A., & Dharmapala, D. (2006). Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics, 79*(1), 145-179.
Dichev, I. D., Graham, J. R., Harvey, C. R., & Rajgopal, S. (2013). Earnings quality: Evidence from the field. *Journal of Accounting and Economics, 56*(2, Supplement 1), 1-33.
Duchin, R., & Sosyura, D. (2013). Divisional Managers and Internal Capital Markets. *The Journal of Finance, 68*(2), 387-429.
Dyreng, S. D., Hanlon, M., & Maydew, E. L. (2008). Long-Run Corporate Tax Avoidance. *The Accounting Review, 83*(1), 61-82.
Dyreng, S. D., Hanlon, M., & Maydew, E. L. (2010). The Effects of Executives on Corporate Tax Avoidance. *The Accounting Review, 85*(4), 1163-1189.
Edmans, A., Goldstein, I., & Zhu, J. (2011). *Contracting with Synergies*.
Eisenberg, T., Sundgren, S., & Wells, M. T. (1998). Larger board size and decreasing firm value in small firms. *Journal of Financial Economics, 48*(1), 35-54.
El-Khatib, R., Fogel, K., & Jandik, T. (2015). CEO network centrality and merger performance. *Journal of Financial Economics, 116*(2), 349-382.
Engelberg, J., Gao, P., & Parsons, C. A. (2012). Friends with money. *Journal of Financial Economics, 103*(1), 169-188.
Feng, M., Ge, W., Luo, S., & Shevlin, T. (2011). Why do CFOs become involved in material accounting manipulations? *Journal of Accounting and Economics, 51*(1), 21-36.
Fich, E. M., & Shivdasani, A. (2006). Are busy boards effective monitors? *The Journal of Finance, 61*(2), 689-724.
Fracassi, C. (2017). Corporate Finance Policies and Social Networks. *Management Science, 63*(8), 2420-2438.
Fracassi, C., & Tate, G. (2012). External Networking and Internal Firm Governance. *The Journal of Finance, 67*(1), 153-194.
Frank, M. M., Lynch, L. J., & Rego, S. O. (2009). Tax Reporting Aggressiveness and Its Relation to Aggressive Financial Reporting. *The Accounting Review, 84*(2), 467-496.
Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215-239.
GE, W., Matsumoto, D., & Zhang, J. L. (2011). Do CFOs Have Style? An Empirical Investigation of the Effect of Individual CFOs on Accounting Practices*. *Contemporary Accounting Research, 28*(4), 1141-1179.
Geiger, M. A., & North, D. S. (2006). Does Hiring a New CFO Change Things? An Investigation of Changes in Discretionary Accruals. *The Accounting Review, 81*(4), 781-809.
Geletkanycz, M. A., & Hambrick, D. C. (1997). The External Ties of Top Executives: Implications for Strategic Choice and Performance. *Administrative Science Quarterly, 42*(4), 654-681.

Gore, A., Matsunaga, S., & Yeung, P. (2007). The relation between financial monitoring and incentives for chief financial officers. *Unpublished Working Paper, George Washington University*.

Graham, J. R., Hanlon, M., Shevlin, T., & Shroff, N. (2014). Incentives for Tax Planning and Avoidance: Evidence from the Field. *The Accounting Review, 89*(3), 991-1023.

Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics, 60*(2), 187-243.

Grant, R. M. (1996). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization science, 7*(4), 375-387.

Hambrick, D. C., & Mason, P. A. (1982). The Organization as a Reflection of Its Top Managers. *Academy of Management Proceedings, 1982*(1), 12-16.

Hanlon, M. (2005). The Persistence and Pricing of Earnings, Accruals, and Cash Flows When Firms Have Large Book-Tax Differences. *The Accounting Review, 80*(1), 137-166.

Hanlon, M., & Heitzman, S. (2010). A review of tax research. *Journal of Accounting and Economics, 50*(2), 127-178.

Hanlon, M., & Slemrod, J. (2009). What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics, 93*(1-2), 126-141.

Hannan, M. T., & Freeman, J. (1977). The Population Ecology of Organizations. *American Journal of Sociology, 82*(5), 929-964.

Harris, D., & Helfat, C. (1997). Specificity of CEO human capital and compensation. *Strategic Management Journal, 18*(11), 895-920.

Hartzell, J. C., & Starks, L. T. (2003). Institutional investors and executive compensation. *The Journal of Finance, 58*(6), 2351-2374.

Hayes, R. M., & Schaefer, S. (2000). Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance. *The RAND Journal of Economics, 31*(2), 273-293.

Hoitash, R., Hoitash, U., & Kurt, A. C. (2016). Do accountants make better chief financial officers? *Journal of Accounting and Economics, 61*(2), 414-432.

Horton, J., Millo, Y., & Serafeim, G. (2012). Resources or power? Implications of social networks on compensation and firm performance. *Journal of Business Finance & Accounting, 39*(3-4), 399-426.

Hsieh, T.-S., Wang, Z., & Demirkan, S. (2018). Overconfidence and tax avoidance: The role of CEO and CFO interaction. *Journal of Accounting and Public Policy, 37*(3), 241-253.

Hwang, B.-H., & Kim, S. (2009). It pays to have friends. *Journal of Financial Economics, 93*(1), 138-158.

Inger, K. K. (2014). Relative Valuation of Alternative Methods of Tax Avoidance. *The Journal of the American Taxation Association, 36*(1), 27-55.

Jack Dorminey, A. Scott Fleming, Mary-Jo Kranacher, & Richard A. Riley, J. (2012). The Evolution of Fraud Theory. *Issues in Accounting Education, 27*(2), 555-579.

Jensen, M. C. (1993). The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems. *The Journal of Finance, 48*(3), 831-880.

Jiang, J., Petroni, K. R., & Yanyan Wang, I. (2010). CFOs and CEOs: Who have the most influence on earnings management? *Journal of Financial Economics, 96*(3), 513-526.

Kaufman, K. (2003). Value-added financial management: The new CEO job description. *Executive Insight, May*.

Kilduff, M., & Tsai, W. (2003). *Social networks and organizations*: Sage.

Kim, E. H., & Lu, Y. (2011). CEO ownership, external governance, and risk-taking. *Journal of Financial Economics, 102*(2), 272-292.

Kim, J.-B., Li, Y., & Zhang, L. (2011). Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics, 100*(3), 639-662.

Kogut, B., & Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization science, 3*(3), 383-397.

Krishnan, G. V., Raman, K. K., Yang, K., & Yu, W. (2011). CFO/CEO-Board Social Ties, Sarbanes-Oxley, and Earnings Management. *Accounting Horizons, 25*(3), 537-557.

Laguir, I., & Staglianò, R. (2014). *How do powerful CEOs manage corporate tax aggressiveness?* (Vol. 34).

Lanis, R., & Richardson, G. (2011). The effect of board of director composition on corporate tax aggressiveness. *Journal of Accounting and Public Policy, 30*(1), 50-70.

Larcker, D. F., So, E. C., & Wang, C. C. Y. (2013). Boardroom centrality and firm performance. *Journal of Accounting and Economics, 55*(2), 225-250.

Lennox, C., LISOWSKY, P., & PITTMAN, J. (2013). Tax Aggressiveness and Accounting Fraud. *Journal of Accounting Research, 51*(4), 739-778.

Lin, N. (2001). Guanxi: A conceptual analysis. *Contributions in Sociology, 133*, 153-166.

Lipton, M., & Lorsch, J. W. (1992). A Modest Proposal for Improved Corporate Governance. *The Business Lawyer, 48*(1), 59-77.

Lisowsky, P. (2009). Inferring U.S. Tax Liability from Financial Statement Information. *The Journal of the American Taxation Association, 31*(1), 29-63.

Malmendier, U., & Tate, G. (2009). Superstar ceos. *The Quarterly journal of economics, 124*(4), 1593-1638.

Mcdonald, M. L., Khanna, P., & Westphal, J. D. (2008). Getting Them to Think Outside the Circle: Corporate Governance, Ceos' External Advice Networks, and Firm Performance. *Academy of Management Journal, 51*(3), 453-475.

McGuire, S. T., Omer, T. C., & Wilde, J. H. (2014). Investment Opportunity Sets, Operating Uncertainty, and Capital Market Pressure: Determinants of Investments in Tax Shelter Activities? *The Journal of the American Taxation Association, 36*(1), 1-26.

Mian, S. (2001). On the choice and replacement of chief financial officers. *Journal of Financial Economics, 60*(1), 143-175.

Mills, L. F., Nutter, S. E., & Schwab, C. M. (2013). The Effect of Political Sensitivity and Bargaining Power on Taxes: Evidence from Federal Contractors. *The Accounting Review, 88*(3), 977-1005.

Minnick, K., & Noga, T. (2010). Do corporate governance characteristics influence tax management? *Journal of Corporate Finance, 16*(5), 703-718.

Phillips, R., Freeman, R. E., & Wicks, A. C. (2003). What Stakeholder Theory is Not. *Business Ethics Quarterly, 13*(4), 479-502.

Rego, S. O., & Wilson, R. (2009). Executive compensation, tax reporting aggressiveness, and future firm performance. In *Working Paper, University of Iowa*.

Rego, S. O., & Wilson, R. (2012). Equity Risk Incentives and Corporate Tax Aggressiveness. *Journal of Accounting Research, 50*(3), 775-810.

Renneboog, L., & Zhao, Y. (2011). Us knows us in the UK: On director networks and CEO compensation. *Journal of Corporate Finance, 17*(4), 1132-1157.

Shackelford, D. A., & Shevlin, T. (2001). Empirical tax research in accounting. *Journal of Accounting and Economics, 31*(1), 321-387.

Subrahmanyam, A. (2008). Social networks and corporate governance. *European Financial Management, 14*(4), 633-662.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8): Cambridge university press.

Weisbach, M. S. (1988). Outside directors and CEO turnover. *Journal of Financial Economics, 20*, 431-460.Wilson, R. J. (2009). An Examination of Corporate Tax Shelter Participants. *The Accounting Review, 84*(3), 969-999.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics, 40*(2), 185-211.

DECISION SCIENCES INSTITUTE
Role of Irrational Opposition in Wearable Health Tracker Adoption in Corporate Wellness Programs

Nima Molavi
Department of Information Operations and Technology Management,
College of Business and Innovation, University of Toledo, Toledo, Ohio, USA
Nima.molavi@utoledo.edu

James Y. Strebler
Department of Information Operations and Technology Management,
College of Business and Innovation, University of Toledo, Toledo, Ohio, USA
James.strebler@utoledo.edu

Suvankar (Sam) Ghosh
Department of Information Operations and Technology Management,
College of Business and Innovation, University of Toledo, Toledo, Ohio, USA
Suvankar.ghosh@utoledo.edu

ABSTRACT

This paper extends the UTAUT 2 by adding the concept of irrational opposition to adoption of wearable health tracker in a corporate wellness setting. Our proposed model incorporates all the UTAUT 2 constructs and adds the new idea of irrational opposition. UTAUT 2 focused on those ideas that can be categorized as rationally processed by the user. However, not all human behavior is rational. We seek to identify the impact of irrationality on the decision as well as potential drivers of that irrationality. In this paper, a preliminary instrument is developed and validated to measure irrational opposition.

KEYWORDS: Healthcare management, UTAUT 2, Consumer technology adoption, Irrational opposition, Health trackers

INTRODUCTION

Corporations seek to improve worker health due to the cost of medical care for workers. Because of this, corporate wellness programs are common in many larger corporations. One potential area for technology in this case is the use of wearable health trackers. These trackers can provide information about heart rate of the subject throughout the day. By using these trackers, the subjects may be able to gather data that they can use to improve their health. In addition, the subject may share this information with their primary care physician.

Voluntary adoption of these health trackers may help with effectiveness of the corporation's wellness program. However, participation in both wellness programs, as well as the use of a health tracker are usually voluntary. This paper seeks to identify those factors, which contribute to a subject using a health tracker. Conversely, this may provide insight into obstacles to adoption.

Because of the similarities to choices on consumer-level technology adoption decisions, the bulk of our theoretical model is based on Unified Theory of Acceptance and Use of Technology 2 (Venkatesh et al., 2012). The constructs in UTAUT 2 are largely rational or social. However, there is also an irrational basis for resistance to some actions. Because of this, based on Prototype/Willingness Model and Automotive Model we added the construct of “Irrational Opposition” in order to capture some of the less rational potential reasons for opposition to adoption.

This paper seeks to:
1. Apply a UTAUT 2 based model to a real-world application
2. Add to the theoretical model by attempting to measure irrationality

LITERATURE REVIEW

In this section most important theories and models related to the adaption of health trackers as well as irrational opposition behavior are reviewed. Three streams of literature have reviewed: (1) Technology Acceptance Models, (2) Theories related to Irrational Opposition, and (3) Adoption of Wearable Health Trackers. The following subsections summarize the most relevant literature in above aforementioned streams.

Technology Acceptance Models

Technology acceptance studies is a major field of interest amongst IS researchers. Studies in this stream analyzing the mechanisms and antecedents of IT artifacts adaptions. Since health trackers are one of the most novel IT artifact, their adaption can be investigated through the lens of technology acceptance models.

Many different models are proposed by various researchers to explain the antecedents and mechanism of new technologies adaption. Davis (1989), in one of the most fundamental papers of this field, proposed Technology Acceptance Model (TAM). In TAM model, two famous antecedents are used to predict the IT adoption: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). PU and PEOU predict the Behavioral Intention which directly influences Actual Behavior. This model was derived from The Theory of Planned Behavior developed by Ajzen (1991).

After the introduction of TAM, researchers aimed to increase the predictability of the model or fit the model into new contexts. Given these objectives, new constructs were added to TAM and new relationships were suggested. Venkatesh and Davis (2000) developed TAM 2 by considering Theory of Reasoned Action (Ajzen and Fishbein 1980; Fishbein and Ajzen 1975) as a complement for TAM. They added subjective norm, voluntariness, and image as social influence constructs, accompanied by job relevance, output quality, and results demonstrability as cognitive instrumental process constructs.

In another substantial work, Venkatesh et al. (2003) studied most prominent information technology acceptance models and combined their constructs. Their effort led to the development of a unified theory, called Unified Theory of Acceptance and Use of Technology (UTAUT). UTAUT model consists of Performance Expectancy, Effort Expectancy, Social Influence and Facilitating Conditions as the antecedents of behavioral intention. Moreover, they have considered four moderators, namely, gender, age, experience and voluntariness, all of which were already studied and supported in previous studies (Venkatesh and Morris, 2000).

Venkatesh et al. (2012) extended UTAUT model and developed UTAUT 2 addressing the technology adoption in the consumer context. Hedonic motivation, Price value and Habit were added to this model in order to fit the UTAUT into consumer context. Additionally, Voluntariness of use is also eliminated as it was not relevant any more. UTAUT 2 can be a great foundation for this research as health trackers are belonging to consumer context. In this regard, this model is selected for studying the adaption of health trackers.

Irrational Opposition

Most of the aforementioned researches are highly focused on behavioral intention. It is widely proved that accurately measuring behavioral intention is necessary to predict the behavioral intention. However, a question arises here: Is knowing the behavioral intention also sufficient to predict the actual behavior?

Warshaw and Davis (1985) argued that the behavioral intention and actual behavior should be disentangled. They defined and explained that Behavioral expectations can also be used to complement the behavioral intention in predicting the actual behavior. According to these findings, this paper tries to study related theories and identify a new factor which could complement the behavioral intention to predict the actual behavior.

Webb and Sheeran (2006) also examined that whether changing behavioral intentions engenders behavior change. Conducting a meta-analysis, they determined that the relationship between behavioral intention and behavior is not always consistent.

Number of theories are proposed to explain the inconsistency of behavioral intention and actual behavior. Automotive Model argues that sometimes the concept of intentional behavior is not equal to goal-directed behavior (Bargh, J. A. 1990). The core idea of this model is that behavior sometimes are unintentional (Gollwitzer and Bargh 2005). So, it is possible that the intention alone, could not predict the behavior thoroughly.

Another related theory is Prototype/Willingness Model (Gerrard et al. 2008) which suggests that sometimes humans are acting a behavior in a non-logical way. In these conditions decision makers perform a behavior which is not rational. For instance, people know that smoking is harmful but they do it. In this case, they have intention to not smoke, but they are willing to do that.

This paper tries to measure all of these mechanisms using a new construct called Irrational Opposition. This construct refers to all of the decisions made by humans that are not based on solid logic or in a rational way.

Adoption of Wearable Health Trackers

Wearable health trackers are referring to the electronic devices being wore by individuals and are able to track health measures, such as heart beat, blood pressure, or oxygen saturation. Smart watches and Fitbit are great examples of wearable health trackers which are commonly being used nowadays.

In one of the earliest researches focusing on the adoption of wearable health trackers (WHTs), Nasir and Yurder (2015) investigated the perceptions of users and physicians about WHTs. TAM model has been used and extended by adding perceived risk and compatibility constructs, in order to fit to the context. However, this study did not provide any empirical support for the proposed hypotheses and WHT acceptance model.

In another similar work, Lunney, Cunningham, and Eastin (2016) tried to blend TAM with TPB model in order to develop a framework for the context of wearable fitness technologies adoption. They used structural equation modeling (SEM) technique in order to test their hypotheses. Based on the results, they claimed that wearable fitness technologies use is significantly related to perceived health outcomes of these devices. As one of the suggestions for future researches, they have indicated that there is a need to focus on motivation, as well as habit formation through goal-setting.

Zhang et al. (2017) aimed to explore factors that influence the WHTs adoption intention. They considered technical attributes, health attributes, and consumer attributes. They have also used TAM model and integrated it with health belief model, snob effect and conformity and reference group theory. The results suggested that adoption of WHTs is being influenced by all of the three considered attributes simultaneously, namely, technical, health, and consumer. Authors have also suggested to pay more attention to the special groups of people who may have different characteristics in their intention to use WHTs.

Lee and Lee (2018) examined the factors influencing individual's intention to adopt a WHT. Considered factors in this research are interpersonal influence, personal innovativeness, and self-efficacy, attitudes toward a wearable fitness tracker, health interests, and perceived expensiveness of the device. Subjects of this research were students which are highly different with the subjects of our paper who are employees of a firm. Based on a logistic regression analysis, it is observed that consumer attitudes, personal innovativeness, and health interests had significant and positive impact on the intention of WHT adoption. Moreover, they have compared two groups of subjects who differed in the knowledge about the WHTs and observed that intention to adopt was stronger among who were aware of WHTs than among who were not aware.

Finally, in one of the most recent studies, Shin et al. (2019) conducted a systematic literature review on the topic wearable activity trackers and categorized the researches into six major themes which are as follows:

1. Acceptance, adoption, and abandonment
2. Behavior Change
3. Patient treatment and medical settings
4. Technology focus
5. Self-monitoring data centered
6. Privacy

Based on the analysis of 78 articles in the acceptance, adoption, and abandonment theme. Authors suggest that this theme of research is bridging "Technology" and "Users" which is identical to other technology acceptance contexts. They have argued that WHTs are highly complex devices which are needed to be understood comprehensively from various perspectives. In this regards, authors called for interdisciplinary awareness about the current landscape of WHTs use.

Theoretical Contributions

This paper aims to add irrational opposition construct to UTAUT 2, in order to increase the predictability power of model in the Health Tracker context. Theoretical contributions of this paper are as follows:

1. Introducing Irrational Opposition construct and its antecedents to technology acceptance literature.
2. Adding Irrational Opposition construct to UTAUT 2 model and analyze its possible relationships with UTAUT 2 constructs.
3. Considering two different paths (Rational Path and Irrational Path) in UTAUT 2 framework.

The rest of this paper is as follows: in section 3 research model is delineated and the related hypotheses are introduced. In section 4 the methodology is proposed and the instrument development procedure is explained. In section 5 the future works are described and the action plan is defined. Lastly, section 6 conclude this paper.

RESEARCH MODEL AND HYPOTHESES DEVELOPMENT

The model is based primarily on the established UTAUT2 model (Venkatesh, 2012). We have also added "image" and "compatibility" based on the Moore and Benbasat, 1991.

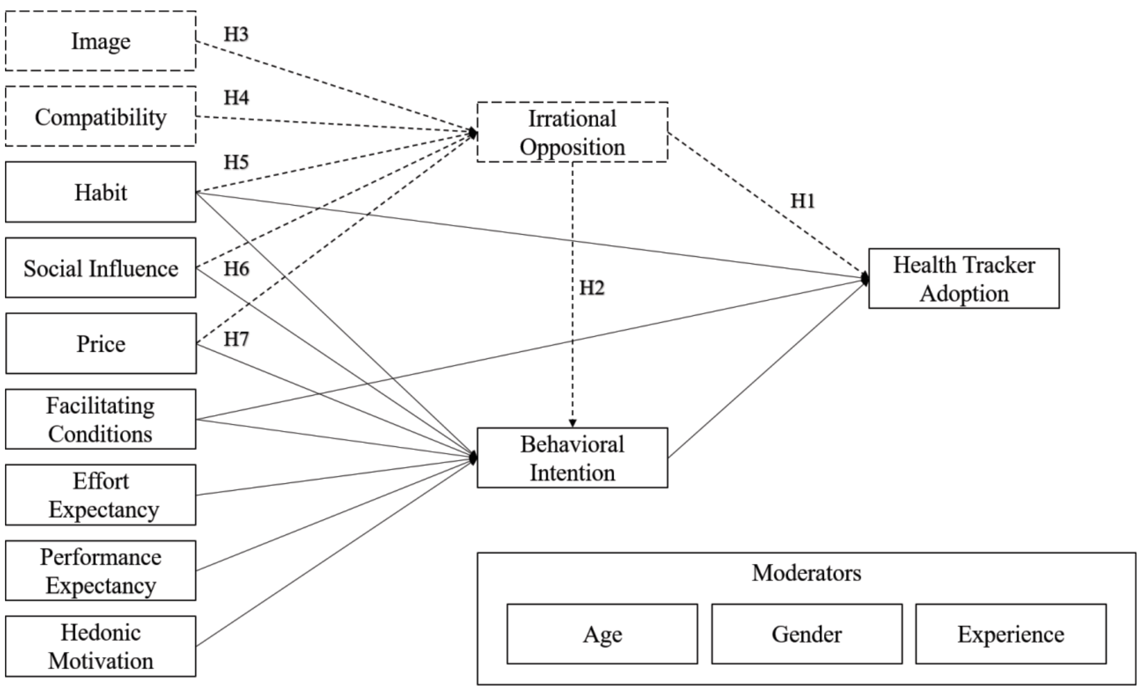


Figure 1- Research Model

The additional construct of "irrational opposition" is based on the concept Behavioral Willingness from Gerrard et al. (2008). The purpose of adding irrationality to the UTAUT 2 model is to attempt to increase the explanatory power of the model. It is notable that, to keep research model simple, the impact of moderator factors on the relationship are not drawn.

The first hypothesis of this paper is related to the main idea of this paper: the impact of irrational opposition on use behavior. Based on Prototype/Willingness Model and Automotive Model, which are reviewed in the previous section, the behavior of decision makers is not always rational. The reason for their behavior could have roots in their irrationality. Moreover, as the irrational opposition is always some type of resistance to a specific behavior, therefore in our study, it has destructive effect on the adoption of wearable health trackers. Consequently, the first hypothesis of this paper is developed as follows:

> Hypothesis 1: *Irrational opposition is negatively associated with the adoption of wearable health trackers.*

On the other hand, irrational opposition can decrease the intention to adopt wearable health tracker. This impact is not the same with the impact of irrational opposition on use behavior as it performs in another layer of behavior. In some cases, the irrational opposition may fundamentally decrease the intention of doing something which is rationally reasonable. This hypothesis is also consistent with the Prototype/Willingness Model.

> Hypothesis 2: *Irrational opposition is negatively associated with the intention to adopt wearable health trackers.*

The next five developed hypotheses are related to the antecedents of irrational opposition. The first antecedent is Image which came from both the paper of Moore and Benbasat (1991). In our paper, the image is defined as the perceptions of subjects about a person who uses health tracker. Some people may think that if someone uses health trackers it means that he or she is not healthy enough or even unattractive. However, some other people may think that if a person uses health tracker he is smart and attractive. This image of a person who uses health tracker will influence the irrational opposition. In this condition, their inappropriate thoughts about the user of health tracker may form an opposition to use wearable health trackers.

> Hypothesis 3: *Image of people about a person who uses health tracker risk is negatively related to irrational opposition.*

Next hypothesis is related to the compatibility of health trackers with people's job and lifestyle. For instance, if a person greatly uses his or her hands during work, the health trackers which should be worn around the wrist will not be compatible with the job. Hence, if using wearable health trackers are not compatible with people's lifestyle or job, they will immediately reject using that. Individuals do not want to use those health trackers because they prefer their lifestyle or routine over tracking their health condition. This opposition is not reasonable because in many cases people can slightly change their lifestyle, and change how they do their job or even how they use health trackers and satisfy requirements of both. In this regards, if wearable health tracker does not seem compatible with the individuals' lifestyle or routine tasks, they may exhibit irrational opposition to use health trackers.

> Hypothesis 4: *Compatibility of wearable health tracker with individual's lifestyle or job is negatively associated with irrational opposition*.

In addition to the influence of added constructs on irrational opposition, the original UTAUT2 constructs may also have impacts on the irrational opposition. We hypothesized that Habit, Social Influence, Price may also influence the irrational opposition. In many studies using Prototype/Willingness Model, it is proved that social influence will form the willingness of a behavior which could be in contrast with their actual intention. Habit can also be influential on

irrational opposition. For instance, consider a person who has not worn any watch or bracelet at all. This person does not want to use wearable health trackers because his habit is in sharp contrast with the adoption of these devices. Price can also lead to the formation of irrational opposition. Individuals may initially get disappointed by the price of a product without understanding and considering its real potential and benefits. If the knowledge of a person about wearable health trackers and their potential benefits is limited, then their high price may raise an irrational opposition to adopt them.
Based on above discussion, the final hypotheses of this paper are developed as follows:

Hypotheses 5: *Social influence is negatively related to the irrational opposition.*
Hypotheses 6: *Habit is negatively related to the irrational opposition.*
Hypotheses 7: *Price is negatively related to the irrational opposition.*

METHODOLOGY

The questions used to measure the UTAUT 2 constructs were adopted directly from the questions used in the UTAUT 2 model. However, the construct of Irrational Opposition was new to this model. Because of this, new measurement questions were developed according to the following methodology.

15 new questions were developed and printed on 3x5 cards. Raters were given the cards and asked to categorize them according to which construct they belonged to. The raters consisted of one current research faculty member and two Ph.D. program students. On the first pass, the correlation of questions to the desired construct was poor with only 28.89% of questions correctly being categorized.

For the second pass, new, established constructs were added and the phrasing of the new questions was changed to reduce ambiguity and bring the phrasing more in line with affirmative statements. On the second pass, 88.89% of the questions were categorized with the appropriate construct. This percentage of correctly categorized questions was acceptable.

Next, of the questions relating to the new construct, they were ordered from best to worst, in order to reduce the number of final questions to four. In order to facilitate this task, the authors provided the raters with 4 general conceptual facets to be measured for the construct. Each rater then selected the question that they felt was the best question in order to represent the conceptual area.

The four conceptual facets were: avoidance of beneficial activities, avoidance of information, resentment of company and social perception or image.

After the raters ordered the questions, a winner for each conceptual facet was selected, resulting in a total of 4 questions to measure the constructs.

The UTAUT 2 questions and the 4 new questions are to be administered via a survey using Survey Monkey. The subjects will be selected by queries to number of companies.

FUTURE WORK

As it is explained in the previous section, this paper will use survey as the methodology to test the hypotheses. In this regard, the next step of this research will be data collection. To collect the data, the population is identified as all of the employees in the United States. The number of firms should be selected from the companies who are willing to initiate health programs in near future or have already implemented these programs.

Afterwards, first, a pilot test can be done with a small random sample of the employees within one firm to test the validity and wording of questionnaire. The wording of questionnaire will be revised after the pilot test and then the survey will be distributed through survey monkey to the entire sample for the final phase.

In the next step, the collected data will be cleaned and used to test the developed hypotheses. Based on the results of hypotheses testing, the possible explanation for the results will be investigated and explained in order to provide insights about the topic.

RESULTS-CONTINGENT MANAGERIAL IMPLICATIONS

In this section, the managerial implications of this paper are discussed. One of the objectives of this research is to provide insights for healthcare managers and help them identify practices which can enhance the adoption of health trackers. If the results support the hypotheses, a number of suggestions are proposed that can be used to reduce the negative impact of irrational opposition to health tracker adoption. These results-contingent suggestions are discussed below:

1. Companies should emphasize the value of the product and clarify its benefits in comparison with the price. It should be shown that the price of a health tracker is reasonable considering the benefits from them. Companies can also provide some financial aid (such as loans) for employees to help them to buy the health trackers. It is also notable that since the price is also one of the antecedents for behavioral intention, by implementing the suggested practices, behavioral intention can also be enhanced directly.

2. The second possible way to reduce the irrational opposition could be by promoting the image of health tracker users. This can be achieved through advertisements, flyers, banners, and campaigns. The main idea is to show that the individuals who adopt the health trackers are smart, popular, and cool. This also needs the commitment of senior and junior managers. They should use health trackers during their work to show to the employees that even their managers and supervisors are using these products. By implementing these suggestions, the image of health trackers can be enhanced and consequently, the irrational opposition may be reduced.

3. In order to reduce the irrational opposition, companies can study the existing work processes and tasks within the company and, if it is possible, modify them to be compatible with the adoption of health trackers. By doing that, employees initially see the compatibility of their work processes with health tracker adoption and the probability of their opposition due to the insubstantial compatibility issue will be reduced.

4. Lastly, companies can also help in changing the employees’ habits. They can allocate incentives for first months of health tracker adoption in order to transform health tracker usage into a habit. Employees initially may adopt the health tracker to receive the participation incentives, but after a while, using health trackers may be turn into a habit for them. For instance, if employees adopt smart watches, they can get used to other features of smart watches and therefore continue using them permanently. In other words, by considering incentives for the participants, companies can try to push the health trackers into employees’ daily routine and habits. Incentives could also mitigate the effect of health

trackers' high price. Consequently, incentives may reduce the irrational opposition through two different paths: (1) Turning health tracker into a habit of employees, and (2) Reduce the negative effect of health trackers' high price.

In contrast, if the results do not support the role of irrational opposition to health tracker adoption, it can be concluded that the subjects are highly rational when they decide on the technologies concerning their health and wellness. If this occurs, further research can investigate the role of irrational opposition in other types of health-related information technologies, in order to see whether the observed rational behavior is limited to health trackers or if it is generalizable to other health-related information technologies.

CONCLUSION

Wearable health trackers are one of the most influential novel ITs that are capable of improving humans' health substantially. In the recent years, companies have tried to utilize health trackers and combine them with health reward programs, in order to increase enhance the health condition of their employees. However, it is not clear to what extent employees are willing to participate in the health programs and adopt health trackers. In this regard, this paper expanded the unified theory of acceptance and use of technology (UTAUT 2) in order to fit the model to health tracker contexts.

This study proposes a theoretical model for the adoption of wearable health trackers as part of a corporate wellness program. A new construct called Irrational Opposition is introduced which can prevent the wider adoption of wearable health trackers. Moreover, Image and Compatibility constructs were added to the model, which are identified in previous studies. It is hypothesized that image, compatibility, social influence, habit, and price may be antecedents of irrational opposition.

To be able to test the hypotheses, measurement instrument is needed for each construct. As the irrational opposition is a new construct, a preliminary instrument is developed and validated using the Q-sort methodology. For all other constructs, items were collected from previous studies and reworded to fit the context. The next step of this study will be a pilot test to refine the instrument. Afterwards, a large-scale survey will be conducted to measure each construct using developed and validated instruments to test the hypotheses.

To conclude, reviewing literature on technology acceptance and wearable health trackers and extending UTAUT 2, this paper brings forward the concept of irrational opposition and proposes a valid measurement instrument for it considering the context of wearable health trackers adoption in corporate health programs.

APPENDIX

Instrument

<u>Compatibility (CO)</u>

CO1. Using a health tracker is compatible with all aspects of my health.
CO2. Using a health tracker is completely compatible with my current situation
CO3. I think that using a health tracker fits well with the way I like to stay in shape.
CO4. Using a health tracker fits into my exercise style

Performance Expectancy (PE)

PE1. I find a health tracker useful in my daily life.
PE2. Using health tracker increases my chances of achieving things that are important to me.
PE3. Using health tracker helps me accomplish fitness goals more quickly.
PE4. Using health tracker increases my ability to maintain good health.

Effort Expectancy (EE)

EE1. Learning how to use a health tracker is easy for me.
EE2. My interaction with a health tracker is clear and understandable.
EE3. I find health trackers easy to use.
EE4. It is easy for me to become skillful at using a health tracker.

Social Influence (SI)

SI1. People who are important to me think that I should use a health tracker.
SI2. People who influence my behavior think that I should use a health tracker.
SI3. People whose opinions that I value prefer that I use a health tracker.

Facilitating Conditions (FC)

FC1. I have the resources necessary to use a health tracker.
FC2. I have the knowledge necessary to use a health tracker.
FC3. A health tracker is compatible with other technologies I use.
FC4. I can get help from others when I have difficulties using a health tracker.

Hedonic Motivation (HM)

HM1. Using a health tracker is fun.
HM2. Using a health tracker is enjoyable.
HM3. Using a health tracker is very entertaining.

Price Value (PV)

PV1. Health trackers are reasonably priced.
PV2. Health tracker are a good value for the money.
PV3. At the current price, health trackers provide a good value.

Habit (HT)

HT1. The use of health trackers has become a habit for me.
HT2. I am addicted to using a health tracker.
HT3. I must use a health tracker.
HT4. Using a health tracker has become natural to me.

Behavioral Intention (BI)

BI1. I intend to continue using a health tracker in the future.

BI2. I will always try to use a health tracker in my daily life.
BI3. I plan to continue to use a health tracker frequently.

Irrational Opposition (IO)

IO1. Although both healthy and unhealthy workers are participating, if I participate in a company health-related program, my co-workers may think I am not healthy.
IO2. I am not willing to use health tracker, even if it could increase my chances of survival in an emergency.
IO3. I do not want more information on my physical health or well-being.
IO7. Any initiative that the company produces is probably bad for employees.

Use

Please choose your usage frequency for each of the following:
a) Health tracker
b) Blood Pressure monitor
c) Fitness Software
d) Diet Journal
e) Fitness Journal
f) Exercise Equipment
Note: Frequency ranged from “never” to “many times per day.”

Image

“Imagine a one of your male/female colleagues is not participating in the health tracker program. How would you describe him/her using the following characteristics?”

1. Cool
2. Smart
3. Independent
4. Popular
5. Self-Conscious
6. Unattractive
7. Immature
8. Dull
9. Confused
10. Nerd

Age

Age was measured in years.

Gender

Gender was coded using a 0 or 1 dummy variable where 0 represented women.

Experience

Experience was measured in months.

Use

Use was measured as a formative composite index of both variety and frequency of mobile Internet use.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour.

Ahuja, M. K., & Thatcher, J. B. (2005). Moving beyond intentions and toward the theory of trying: effects of work environment and gender on post-adoption information technology use. MIS quarterly, 29(3), 427-459.

Bargh, J. A. (1990). Goal and intent: Goal-directed thought and behavior are often unintentional. Psychological Inquiry, 1(3), 248-251.

Casselman, J., Onopa, N., & Khansa, L. (2017). Wearable healthcare: Lessons from the past and a peek into the future. Telematics and Informatics, 34(7), 1011-1023.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, and behavior: An introduction to theory and research. Reading, Mass.: Addison Wessley.

Gerrard, M., Gibbons, F. X., Houlihan, A. E., Stock, M. L., & Pomery, E. A. (2008). A dual-process approach to health risk decision making: The prototype willingness model. Developmental review, 28(1), 29-61.

Gollwitzer, P. M., & Bargh, J. A. (2005). Automaticity in goal pursuit (pp. 624-646).

Hu, P. J., Chau, P. Y., Sheng, O. R. L., & Tam, K. Y. (1999). Examining the technology acceptance model using physician acceptance of telemedicine technology. Journal of management information systems, 16(2), 91-112.

Lee, S. Y., & Lee, K. (2018). Factors that influence an individual's intention to adopt a wearable healthcare device: The case of a wearable fitness tracker. Technological Forecasting and Social Change, 129, 154-163.

Lunney, A., Cunningham, N. R., & Eastin, M. S. (2016). Wearable fitness technology: A structural investigation into acceptance and perceived fitness outcomes. Computers in Human Behavior, 65, 114-120.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. Information systems research, 2(3), 192-222.

Nasir, S., & Yurder, Y. (2015). Consumers' and physicians' perceptions about high tech wearable health products. Procedia-Social and Behavioral Sciences, 195, 1261-1267.

Shin, G., Jarrahi, M. H., Fei, Y., Karami, A., Gafinowitz, N., Byun, A., & Lu, X. (2019). Wearable Activity Trackers, Accuracy, Adoption, Acceptance and Health Impact: A Systematic Literature Review. Journal of biomedical informatics, 103153.

Taylor, S., & Todd, P. (1995). Assessing IT usage: The role of prior experience. MIS quarterly, 561-570.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS quarterly*, 115-139.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS quarterly, 36(1), 157-178.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychological bulletin, 132(2), 249.

Warshaw, P. R., & Davis, F. D. (1985). Disentangling behavioral intention and behavioral expectation. Journal of experimental social psychology, 21(3), 213-228.

Zhang, M., Luo, M., Nie, R., & Zhang, Y. (2017). Technical attributes, health attribute, consumer attributes and their roles in adoption intention of healthcare wearable technology. International journal of medical informatics, 108, 97-109.

DECISION SCIENCES INSTITUTE

Identify Short-Term Trend in Support Services' Demand for Technology

Askar H. Choudhury
Illinois State University
Email: choudhury@ilstu.edu

Nathan S. Hartman
Illinois State University
Email: nshartm@ilstu.edu

Ted A. Coussens
Illinois State University
Email: tacouss@ilstu.edu

ABSTRACT

The purpose of this study is to explore the underlying support services demand as way to improve the operational processes of educational technology. This allows for an alignment between the short-term service schedule and the expected users demand for short-term support. Enterprise system data at a large U.S. university was used to model the demand for service support in educational technology. Significant negative coefficients supported our claims that service demand, training requirements, and in-class repairs slowly decline during the semester, whereas lamp replacement demand grows as the semester progress. The idea of operational analytics to measure short-term trend in service demand was used as empirical signal technology managers could use in technology management.

KEYWORDS: *Educational Technology; Technology Management; Support Services Demand; Support Service Management; Cross-directional Trend*

INTRODUCTION

Technology and the system itself are maintained through support service operations, including incident and maintenance management techniques in a manner common to information technology (IT) service management conducted within most organizations (McNaughton, Ray, & Lewis, 2010). Technology is essential in accomplishing the primary organizational operations (Croteau & Bergeron, 2001; Mithas, Ramasubbu, & Sambamurthy, 2011; Tallon, Kraemer, & Gurbaxani, 2000) and has become a tangible resource utilized by organizations to enhance their capabilities. Instructional technology, however, can be more prone to frequent incident and maintenance requests as it is often a technology originally designed to be employed in business office settings and was adapted for use in classrooms. Instructional technologies are similar in form, but operationally function differently than its original design. For example, unlike office use situations instructional technology in each classroom is prone to be used by variety of different instructors within each day, on different days of the week, from semester to semester. Therefore, technology support services demand surges from low to high or high to low at different times within the semester. Thus, the support service structure of information technology in traditional applications is potentially different from similar technology used in unique environment, such as educational technology. As Antonenko, et al. (2017) pointed out that much of the educational technology development has been driven by the considerations of what can be developed using the existing technologies, rather than what should be developed for the users need. Cunningham & Kwakkel (2017) explored forecasting of technological change in the context of high-tech organizations. Stormi, et al., (2018) presented a forecasting model that predicts demand for industrial services which is directly related to the installed base, such as repair, maintenance, and remanufacturing. Giannì & Franceschini, (2003) proposed a model in e-commerce that analyses the ways in which managers can obtain information regarding customers need over time to perform more personalized service. Ishak & Newton (2018) in their study indicated that user resistance to new technology feasibly can be predicted. These studies do not investigate the short-term trend direction that could be of useful to offset resource requirement and help to improve the resource planning efficiently. As the service landscape changes with increasing technological demands, and as the field adopts a stronger transformative approach towards the entire service eco-system, undertaking service demand planning and their support management becomes critical (Subramony, et al., 2017). Unfortunately, little attention has been devoted towards the investigation of support services demand and quantifying the magnitude of the effect of support services demand in educational technology management.

In this regard, broader industry technology management literature may be a guide for understanding the impact on the support services demand. Unfortunately, the broader industry technology management differs from educational technology in its environment and types of service required (Banerjee & Igbaria, 1993). For example, the short-term trend in support services demand is different in its creation in educational technology system that occurs within the semester. This induces the concept of differences in technology use is dependent on service categories resulted from semester variations in educational environment as oppose to quarterly variations in broader industry technology environment. Thus, we conceptualize that there is a short-term (upward or downward) trend in demand for support service within a semester in educational technology environment is specific to a service category. This within the semester increasing or decreasing demand is referred to as short-term trend as oppose to long-term trend that exists over the years.

LITERATURE REVIEW

Determining the frequency to which systems once adopted are maintained and the frequency to which technology support staffs are called into a classroom for routine and/or emergency support associated with successes attained in classrooms can broaden the understanding of technology support management. This area might seem intuitive as equipment manufacturers provide predictions for the frequency of upkeep, but the landscape of universities creates complicated technology integrations because the technology is often different from one department to another and different from one time to another with the variability being when the technology was installed and the prevalent technology at that time to suit the demands of the departments. In addition, hardware and/or technology systems frequently require unexpected on-site modification, once adopted to attain successful operation and create a challenge to the management.

The literature on the integration of technology and its management in the university classrooms has focused on technology adoption (Kopcha & Sullivan, 2006), technology use (Shelton, 2014), and success of students' technology-enhanced learning (Thinnukool, 2018). To some lesser degree researchers have focused on the process used by universities to select technology for the long-term and adoption of technology, a common research area, is one of the barriers associated with the efficient integration and management of instructional technologies (Reid, 2014). Ironically, the amount and frequency of technological support service provided by staff required to attain satisfactory functionality across a university's classroom technologies are key predictors to the faculty's use and adoption. In many cases, the institutional policy along with individual academic departments might drive the adoption of new instructional technology (McPherson & Whitworth, 2008; Shelton, 2014). However, the decisions for technology adoption or its use are usually made by two different stakeholders (Lanzolla & Suarez, 2012). In their study, they have noted that the benefit of technology adoption will not materialize unless it is being used. Salmon and Angood (2013) observed midlevel administrators pursue cost-efficient solutions to satisfy a majority of technology users' need.

However, the question has grown from, if the resources invested will enhance education to, if installed, when and how often will the university be called upon to service the technology and what are the key determinants (Choudhury, Hartman, and Coussens, 2018). The optimal mixture of factors and the frequency of support dedicated to these technologies have not been well considered in previous research studies. In fact, to our knowledge no empirical research study has been done in understanding or estimating the support services demand in educational technology. The most identifiable trend associated with instructional technology has been the increasing amount and different types of technology used by students and faculty. While the users of technology are often identified as a major barrier to the adoption and use of technology (Butler & Sellbom, 2002; Reid, 2014), sometime due to anxiety associated with the technology use (Igbaria & Parasuraman, 1989), we propose the frequency to which a university's staff must dedicate support services to these technologies may also be a barrier.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE
Active-ating Agile Project Management with Team-Based Learning

Jeffrey P. Landry
Department of Information Systems and Technology
School of Computing
University of South Alabama
Email: jlandry@southalabama.edu

David M. Bourrie
Department of Information Systems and Technology
School of Computing
University of South Alabama
Email: dbourrie@southalabama.edu

David Ratisher
Workforce Development Division
Pasco-Hernando State College
Email: dratisher@gmail.com

ABSTRACT

Prior work demonstrated that a project management course for information systems and technology majors has a specific curricular role for teaching agile project management concepts and practices. This paper expands on the prior work, describing a course module for delivering agile content. As agile methods were and still are a marked departure from more traditional methods, we propose that agile therefore requires an equally non-traditional teaching method. Team-Based Learning, an active learning pedagogy, was used to engage students in the spirit of agile software development. Said pedagogy is described and its effectiveness supported with in-course assessment results.

KEYWORDS: Agile, Project Management, Team-Based Learning

INTRODUCTION

The question of where to teach agile in the information systems curriculum was partially addressed in our prior paper (Landry & McDaniel 2016), where we contended that a course in information systems (IS) project management was a good choice. Our approach followed three principles in defining agile content for a project management course: (1) agile alongside, (2) agile across, and (3) agile project, not product. Not wishing to tout agile as a silver bullet, and recognizing the variety of different approaches to IS project management, we did not scrap the traditional methods, but taught agile alongside them. We found that agile topics might span all ten PMBOK® areas, and so we decided to attempt to identify topics across all ten PMBOK areas of project scope, cost, schedule, quality, communication, risk, resource, integration, procurement, and stakeholder management (Project Management Institute, 2017). Finally, keeping in mind that the target course was IS project management, we decided to avoid the non-project management agile topics, such as pair programming, and focus instead on only the project management-related agile topics.

Having provided an answer to the *where* question, we then addressed the related *what* question, by identifying the specific topics across the PMBOK areas, as well as defining a preliminary set of seven learning outcomes for agile project management (PM). The topics we chose were daily standups, Kanban boards, and burndown charts. We integrated these into a cohesive set of exercises. The learning outcomes for agile coverage were as follows:

1. Evaluate the suitability of agile methods for use in a given project and organization context
2. Analyze project proposals using multiple techniques
3. Compare and contrast traditional versus agile project management
4. Demonstrate an awareness of agile project management basic terminology and concepts across multiple PMBOK areas.
5. Apply agile PM principles and techniques for managing in multiple PMBOK areas
6. Discuss the soft skills and abilities of project managers
7. Complete team-based work applying the principles and tools of project management

Building on the approach of the prior paper, we now further examine the teaching and learning of agile methods in the project management curriculum. In this paper, we consider the question, *Is Team-Based Learning an effective pedagogy for educating students in agile project management?* This is an important question for educators in computing disciplines where project management is taught. As agile project management adoption continues to increase, it is important that graduates be prepared to flourish in this new environment. Students should have a strong background in agile methods, knowing that they are different from traditional methods, and that this difference isn't trivial or arbitrary. As a professional, they must be able to thrive in both traditional and agile environments, understand the differences, and if necessary, make decisions regarding the methodology to use. The selection and use of a project management methodology will affect future project successes and failures. The discovery of effective teaching methods, and their improvement, is essential to the preparation of future agile project management professionals.

BACKGROUND

The background for our approach stems from three areas: agile in project management, active learning in agile project management, and team-based learning.

Agile in Project Management

Agile, according to the Project Management Institute, is best defined by the Agile Manifesto, which it considers to be the "original and official definition of agile values and principles" (Alliance 2017, p. 150). In the manifesto (Beck et al., 2001), the 17 signing authors introduced their ground-breaking software development approach with the declaration:

> We are uncovering better ways of developing software by doing it and helping others do it. Through this work, we have come to value:
>
> Individuals and interactions over processes and tools
> Working software over comprehensive documentation
> Customer collaboration over contract negotiation
> Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more.

The authors go on to define 12 guiding principles behind the Agile Manifesto:

- Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.
- Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.
- Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale.
- Business people and developers must work together daily throughout the project.
- Build projects around motivated individuals. Give them the environment and support they need and trust them to get the job done.
- The most efficient and effective method of conveying information to and within a development team is face-to-face conversation.
- Working software is the primary measure of progress.
- Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely.
- Continuous attention to technical excellence and good design enhances agility.
- Simplicity--the art of maximizing the amount of work not done--is essential.
- The best architectures, requirements, and designs emerge from self-organizing teams.
- At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.

In the nearly 18 years since the publication of the Manifesto, the approach has been implemented in various methods, including eXtreme Programming (XP) and Scrum (Lindstrom and Jeffries, 2004). For many years it was difficult to find reliable numbers on agile adoption, or empirical evidence of agile's effectiveness. Within the last five years, however, an empirical study of 1,002 projects found that agile methods had a positive impact on two dimensions of project success: efficiency and overall stakeholder satisfaction against organizational goals (Serrador and Pinto, 2015).

Agile has been identified as belonging to six published standards relevant to information systems and project management education. While the roots of agile go back as far as the turn of the millennium, the IS 2002 and MSIS 2006 (Gorgone et al., 2006) curriculum models provide no mention of agile methods. By the end of the first decade of the 2000's, though, agile was making headway. The IS 2010 model curriculum (Topi et al., 2010) suggests that "agile methods" be taught in the Systems Analysis and Design course. The Project Management Body of Knowledge (PMBOK®), published by the widely recognized Project Management Institute (PMI), was updated to include coverage of agile, offering an Agile Certified Practitioner (PMI-ACP®) certification (Project Management Institute, 2017), and the Agile Practice Guide (Alliance, 2017). The Schwalbe (2019) and Marchewka (2015) textbooks on project management provide coverage of agile methods. Finally, a preliminary report on an CIS curriculum model (Longenecker et al., 2016) and MSIS 2016 (Topi et al., 2017) list project management as a key outcome, and there are mentions in several places that students should gain a novice understanding of agile methods, including as a topic in the prescribed IS project management course.

Active Learning in Agile Project Management

Three prior studies have found empirical support for the effectiveness of active learning methods in teaching agile project management concepts. McAvoy and Sammon (2005) had students engaged in a workshop exercise where students debated the potential effectiveness of agile methods (Extreme Programming) in two project case scenarios. The teams then carried out a weighted scoring exercise, applying critical adoption factors. A total of 92% (22 of 240 students reported that the active learning approach was useful in increasing their understanding of agile methods. Schmitz (2018) developed a role-play simulation of the Scrum methodology called Scrummy to introduce agile project management concepts and practices to three cohorts of health informatics majors. He reasoned that an active, team-oriented, face-to-face learning method that solved problems with tacit knowledge was a good fit with the actual values and practices of agile project management. His results indicated that student self-efficacy for concepts introduced in the role-play exercise were higher than for concepts taught in lecture only and that self-efficacy post-test scores were statistically higher than pre-test. Another study set out to compare active learning methods to lecturing in an IT project management context. Sibona, Pourreza, and Hill (2018) used an origami (Japanese paper-folding) exercise as the basis for an active learning exercise on Scrum. Their results found that most students preferred active learning to lecture, found the active learning exercise to be more engaging, and that when both methods were used, that lecture come first, followed by active learning. The authors recommended the lecture→active approach, due to the variety of learning styles encountered and a reasonable amount of support for the effectiveness of lecture in the sample.

Team-Based Learning

The active learning pedagogy chosen for the project management courses sampled in this study is Team-Based Learning (TBL). TBL is a specific technique for using teams in courses that is sometimes confused with the general use of teams in courses. The specific TBL initiative adopted by the co-authors was devised by Larry Michaelsen and others and is described in their publications (Michaelsen, Knight, & Fink, 2004; Michaelsen & Sweet, 2008; Parmalee & Michaelsen, 2010). TBL has been used across a wide range of disciplines, including the natural sciences, social and behavioral sciences, legal studies, the humanities, and business (Knight, 2004). TBL can be used in classes of various sizes, making large classes feel smaller (Michaelsen, 2004). TBL has also been used successfully in international situations (Cragin, 2004) and applied in the electronic classroom (Freeman, 2004).

According to Michaelsen, for TBL to be successful, key practices must be faithfully incorporated, including the following:

- backward design
- team formation
- pre-class individual learning
- readiness assurance process
- application activities
- peer evaluation

Each of the TBL practices listed above were incorporated into the project management course and in delivering the agile project management lessons. A brief description of each TBL method follows.

Backward Design

With backward design, the teacher begins by asking: “what do we want our students to be able to do with the knowledge by the end of the course and afterwards?” Then, the teacher defines good course objectives, and activities that enable application of knowledge towards the objectives.

Team Formation

TBL relies on the use of permanent, balanced teams of 5-9 students each.

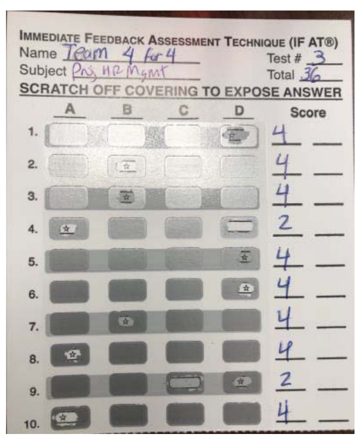
Pre-Class Individual Learning

In order for the students to have high-quality experiences during the team activities, each student must come prepared. This independent learning phase usually consists of an assigned reading.

Readiness Assurance Process

Following pre-class learning students experience the readiness assurance process, or RAP. The RAP consists of readiness assurance tests (RATs), which include individual and team tests, called iRATs and tRATs, along with grade appeals and clarifications. The tRAT It is the same test as the iRAT but given to the collaborating teams. The teams complete items on a scratch-off card, like a bingo game card, which gives them immediate feedback and the ability to answer in multiple attempts (for fewer and fewer points). The appeals and clarification phases provide a review of what was missed. Figure 1 provides an example of the Information Feedback Assessment Technique (IFAT®) that is used for the tRAT score sheet.

Figure 1. IF-AT®



Application Activities

Once knowledge has been assessed/assured through the RATs, students are then prepared for applying that knowledge. Application activities may be completed by teams and/or individuals. Team activities stimulate learning through collaboration and critical thinking within the team and competition and discussion across teams to bring about deeper individual learning in the applied knowledge areas.

Peer Evaluation

Anonymous peer evaluation is used to help make team members accountable to each other. The peer evaluations are part of each individual's grade.

Team-based learning is a suitable pedagogy for a project management class with coverage of agile methods alongside traditional methods. Teamwork is an important aspect of project management, and TBL relies heavily on collaborative learning. TBL's use of permanent teams lends itself to activities that use teams, such as daily standups. TBL's emphasis on collaboration and critical thinking is helpful for getting students to objectively analyze project management methods, such as comparing traditional methods to agile methods. Schmitz (2018) similarly pointed out the values that underlie agile methodologies—specifically, the preference for collaboration through face-to-face interaction that is described in the Agile Manifesto. He used an active learning approach—role-play simulation—for teaching and learning agile concepts and practices.

APPROACH

The topical content chosen for the IS project management courses was based on the idea that the agile development process includes not only software development activities, but project management as well. As we wanted to actively engage students in an applied agile experience, we looked for agile practices that could be carried out, or performed, in nearly authentic instances. We choose three activities: daily standup meetings, Kanban boards, and burndown charts. Daily standup meetings, also called daily scrums, are brief, once-per-day collaborative get-togethers where a development team addresses three sets of issues: progress from the prior day, intentions in the current day, and obstacles encountered or anticipated. A Kanban board is a visualization tool that depicts the status of task progress through a left-to-right flow of a task card across status "swimlanes". A burndown chart is a bar chart depicting daily work remaining for a timeboxed period. Taken together, these three agile practices focus on project schedule management, providing an alternative approach to traditional project scheduling. Furthermore, these activities span across multiple PMBOK areas, including project resource, scope, and communication management. We designed the lesson in such a way as to integrate the three activities into a single set of related activities.

The approach we took in teaching agile concepts and practices in the IS project management curriculum was to create a course lesson module using the TBL pedagogy. The following section describes our coverage and delivery of agile project management using the components of TBL.

Backward Design

We started with the question of "what do we want our students to be able to do with their knowledge of agile methods by the end of the course and afterwards?" This question was largely answered by our prior paper (Landry & McDaniel, 2016), which concluded that agile coverage across PMBOK areas, alongside traditional PM, and project not product would be emphasized, and by the prior section of this paper, which operationalized this coverage through an integrative problem-solving assignment on daily standups, Kanban boards, and burndown charts. The agile-related learning outcomes we developed for the three agile PM sub-topics were as follows.

By the end of the course, the student will be able to:

- Walk through a daily standup meeting, identifying, assigning, and reporting completed tasks while addressing obstacles.
- Update a Kanban board, given the results of that day's standup meeting.
- Create a burndown chart, calculating average velocity and estimated completion date.

Key takeaways include the need to assign (or volunteer for) work, check on its progress, and document its completion systematically; how to use a simple tool like a whiteboard to create a visual of work in its various states of completion; and how a team's recent productivity is useful for predicting future progress.

Team Formation

A total of 59 students were enrolled in three course sections of project management—two sections of an undergraduate course and one section of a graduate course—in the spring of 2019. At the beginning of the semester a total of seven (7) undergraduate and five (5) graduate teams, comprised of between four to six members each, were formed. Teams were formed to be demographically and academically balanced. Every attempt was made to not isolate a female member. There was a mix of U.S.-born and international students in the graduate course. Most importantly, strong academic performers were dispersed.

Pre-Class Individual Learning

All three sections were provided with readings and some lecture for preparation. Textbook readings (Schwalbe 2019 and Marchewka 2015) were assigned to the undergraduate and graduate students, respectively. All three sections were assigned reading materials published on the course management site. Students in the graduate course were assigned two articles (Lindstrom and Jeffries 2004; Serrador & Pinto 2015) to read in addition to the other content. The lecture and readings across all three sections introduced the students to the concept of how agile is different from traditional software development; reviewed key aspects of the Agile Manifesto; and covered topics such as iterative planning, timeboxing, incremental releases, individual roles, daily standups, Kanban boards, story points, burndown charts, velocity, and sprint estimation.

Readiness Assurance Process

Prior to the readiness assurance process, the students in each section were given written guidance on RAT coverage, as follows.

"<u>RAT-2b Agile Planning</u>

Our RAT #2 will take place in our class meeting this week. It will follow the format of prior RATs, with the iRAT followed by the tRAT, and then appeals and clarifications. The 10 items on agile planning will cover the following topics: agile manifesto, product backlog, iterative, incremental, sprints, daily stand-ups, kanban, velocity, burndown, timeboxing, and story points."

Of the ten agile items on the RAT, we are most interested in those mapped directly to upcoming application activities. These were five items closely linked to the application of agile concepts for daily standups, Kanban, and burndown:

1. In agile development projects, ___ is the primary measure of progress. (Ans: d)
   a. daily stand-up c. Gantt
   b. iteration d. working software

2. A unitless metric for the relative size of a requirement is ___. (Ans: b)
   a. escaped defects c. daily standups
   b. story points d. timeboxing

3. A project whose sprint velocity is 9 story points per day at the end of Day 4, and with 54 story points remaining will finish at the end of what day? (Ans: d)
   a. 13 c. 5
   b. 6 d. 10

4. At the end of Day 1 of a sprint there are 75 story points worth of work remaining, and at the end of Day 6 there are 50 story points of work remaining. What is the project's current velocity, in units of story points per day? (Ans: b)
   a. 25 c. 4
   b. 5 d. 50

5. In the middle of a sprint, a Kanban board with three swimlanes looks as follows: 7 cards in the left-most lane, and 5 cards in the right-most lane. What's the status of this project? (Ans: c)
   a. Everything is going as it should
   b. The team has run out of work to do, and so the sprint should conclude
   c. The team is not working on anything, even though there are tasks to be completed
   d. Nothing has gotten done yet

Application Activities

The students in all three sections received three application activities in total, and the graduate class received a fourth. The common application activities were: (1) a team application activity—agile project planning role play, (2) midterm exam problems, and (3) final exam problems. All application activities covered daily standups, Kanban boards, burndown charts, and computation of velocity and estimated day of completion. The graduate class also completed a homework assignment on creating an agile PM spreadsheet.

<u>Team Application Activity: Agile Project Planning Role Play</u>

In the class period following the RAP, students were given a team application activity on Agile Project Planning. The activity was a role play. One of the teams in the class section volunteered to role play, while the other teams observed. Students read their part from a script given out to the class. The script was written to incorporate agile manifesto values and principles. The roles were Narrator, Scrum Master, Assistant (to the scrum master), and three other Team Members. Together, the team walked through several days of daily standup meetings, following the script, and standing in front of the class, using the whiteboard. The scrum master would ask for a volunteer, and a team member would raise their hand and volunteer, reading their line. The assistant would write up a Kanban card complete with time estimate and team member's name. Then, the card would be placed in the (hopefully) correct swimlane among the three, labeled "to-do", "doing" and "done" from left-to-right. Some team members reported work completed, and the assistant then had to move a Kanban card from one swimlane to the next. A team member might report an obstacle that would lead to a discussion and possible actions taken. All of this was in the script. Mistakes were possible. By the end of the last day, the team in front of class, along with the teams observing, drew a burndown chart with work remaining bars for each day of the sprint. The assistant was asked to complete, with the team's help, the velocity and estimated day of completion. Each of the observing teams were then given a list of questions to answer in their teams.

Midterm Problems

The second application activity was in the form of problems given on the midterm exam. There were 20 points (20%) of problems given to each class section. Like the team application activity, these problems asked students to complete a Kanban board, draw a Kanban card, draw a burndown chart, and calculate velocity; and estimated day of completion, showing their work.

Final Exam

Similar to the midterm exam, the students in all three sections were given problems to complete. The final exam problems were worth between 13% (graduate class) and 15% (undergraduate classes) of the course grade.

Peer Evaluation

The students were evaluated on their semester-long team performance, using a peer evaluation instrument administered by Catme.org (Loignon et al., 2017; Ohland et al., 2012; Loughry et al., 2007). There were four items, each evaluated on a five-point scale, plus peer-to-peer comments. Students were required to differentiate scores across two dimensions: every item had to have at least one student scored as less than 5, and every student had to be less than 5 on at least one item. As previously mentioned, the peer evaluation grade weight was 8-10% of the final grade.

RESULTS

In this section, we will discuss the results of the iRAT, tRAT, team application activity, midterm exam, final exam, and peer assessment by using both quantitative and qualitative analysis to compare sections and aggregate mastery levels. The results are based on 57 students because one student from each of the undergraduate sections was absent the day of the agile readiness assessments.

iRAT

A one-way between subjects analysis of variance (ANOVA) was conducted to compare the effect of iRAT scores for the three class sections. Table 1 provides the means and standard deviations for evaluations conducted in the three classes. There was a significant effect on the iRAT scores on the class sections at the $p<.05$ level [$F(2, 54) = 5.11$, $p = .009$]. Post hoc comparisons using the Tukey HSD test indicate that the mean score for the graduate class and the undergraduate class with 18 students were significantly different. However, the undergraduate class with 11 students did not significantly differ from the other classes.

In Pre-Class Individual Learning, there is not a definitive way to figure out what each student did. However, the wide variance in iRAT scores combined with the qualitative feedback from students in post-hoc surveys indicate that not all students did their pre-RAT preparation.

tRAT

A one-way between subjects ANOVA was conducted to compare the effect of tRAT scores for the three class sections. There was a significant effect on the tRAT scores on the class sections at the $p<.05$ level [$F(2, 54) = 15.42$, $p = .000$]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the graduate class and both undergraduate classes were significantly different. However, the two undergraduate class did not significantly differ from one another.

Table 1: Readiness assessment, midterm and final exam results

	Classes					
	Undergraduate (N=11)		Undergraduate (N=18)		Graduate (N=28)	
Evaluation	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
iRAT	52.73%	29.78%	45.28%	31.03%	71.61%	26.04%
tRAT	82.73%	4.67%	83.89%	16.85%	98.21%	3.90%
Midterm	88.64%	24.92%	78.05%	24.92%	89.55%	13.44%
Final	78.79%	16.55%	77.22%	22.35%	89.74%	18.60%

As is typically the case with iRATs, individual performance varied widely, and although mean iRAT scores for each class section seem to differ, these differences were not statistically significant. Also as expected, tRAT scores were higher than the individual iRAT means. The iRAT-to-tRAT means by section went from 52.73.% to 82.73%, 45.28% to 83.89%, and 71.61% to 98.21%. Collaborating teams performed at a higher level than the individual student.

Team Application Activity

Although the team application activity was not graded, there is qualitative evidence that students were engaged in the activity and thought critically. There was evidence of (a) correct performance, (b) identification and correction of mistakes, (c) critical analysis of the case, and (d) insightful observations.

The volunteer teams in all three sections were able to walk-through the script and complete their assigned tasks with little or no intervention by the instructor. Nearly all of their actions were correct. Visible mistakes were identified by the observing teams and corrected for everyone to see. Here is a summary of performance on each of the activity questions:

1. Is the Day 5 Kanban board correct?

Issue surfaced: A volunteer team in one section made a mistake by assigning a Kanban card to the "to-do" swimlane, rather than the "doing." This enabled a clarification: if a task is identified and immediately assigned to an individual, it goes right to the "doing" swimlane.

2. Is the Assistant's estimate of the work remaining correct?

Issue surfaced: There was a discrepancy in velocity calculation. Some teams calculated by subtracting the sum of work completed from the work remaining at Day 1, then dividing by the number of days elapsed. However, this was a mistake, as some work was added during the sprint, and this additional workload was not accounted for. The correct calculation was to subtract work remaining on the last day from the work remaining on the first day, then dividing by the elapsed days.

3. When will the work be completed? Will it be finished by the end of the sprint?

Result: All teams realized that the work planned would not be completed by the end (Day 14) of the sprint, based on the now-corrected calculation of estimated day of completion, compared to the actual last day of the sprint given to them.

4. What, if anything, do you think went wrong? What could the Scrum Master have done about it?

Issue surfaced: In the undergraduate classes the teams realized that in the script, the scrum master did not ask the "what obstacles?" question, and that there was one task that stayed in the "doing" swimlane for a long time. They reasoned that surfacing obstacles might have cleared the way for the task to be completed.

Insightful questions: The graduate class also asked insightful, critical questions that led to discussion. One team asked if an obstacle was found to delay a sprint, and if that obstacle required an infrastructure fix, would that additional work qualify for a new Kanban card and result in additional work completed. The prevailing recommendation was to not count the infrastructure fix as work completed. A team's productivity, or velocity, should be based on completing its own work—such as software development for a software team. Another team asked what should happen to tasks that were not completed correctly. The recommendation was to move the task back to a prior swimlane ("doing") rather than writing a new "correction" card, although arguments were presented for both options.

Midterm Exam

The midterm exam questions that were used to assess the student's retention of the agile topics include the following:

- Assume that in an agile project sprint, tasks T1 and T2 have already been completed; Miriam is working on task T3; Fred has started task T4; and tasks T5, and T6 have not yet been assigned for anyone to work on. Draw a Kanban board correctly depicting the status of all these activities. (5 pts.)

- Draw a detailed Kanban card (not board) for a task T9-design database, whose size is estimated to be 20 story points, and with Bill and Sue assigned to work on it. (5 pts)

- At the end of Day 3 of a sprint there are 90 story points worth of work remaining, and at the end of Day 6 there are 30 story points of work remaining. What is the project's current velocity, in units of story points per day? (Show your work) (5 pts.)
- A project whose sprint velocity is 15 story points per day at the end of Day 4, and with 90 story points remaining, will finish at the end of what day? (Show your work) (5 pts.)

A one-way between subjects ANOVA was conducted to compare the effect of midterm exam scores for the three class sections and no significant effect was found at the $p<.05$ level [$F(2, 54) = 2.52$, $p = .090$]. The undergraduate class with 18 students had scores range from 15% to 100% with three students having scores between 15% and 45%. The other undergraduate sections scores ranged between 70% and 100% for the same questions.

Overall, 85.96% of students, or 49 out of 57, performed at a satisfactory level or better on the midterm's agile assessment activity (where satisfactory is defined as the minimal passing level, of 70% for undergraduates and 80% for graduate students).

Final Exam

The final exam questions (not shown) were variants of the midterm questions. A one-way between subjects ANOVA was conducted to compare the effect of final exam scores for the three class sections and no significant effect was found at the $p<.05$ level [$F(2, 54) = 2.63$, $p = .082$]. Results show that for two of the class sections, final exam performance matched that of the midterm. The undergraduate class with 11 students, scored significantly lower on the final exam section [$t(10) = 15.788$, $p = .000$) as compared to the midterm exam. The final exam scores for both undergraduate sections were statistically the same ($t(27) = 0.20$, $p = .843$).

Overall, 75% of students, or 42 out of 56, performed at a satisfactory (minimal passing) level or better on the culminating assessment activity—the agile application activity items on the final exam. The 85.96% to 75% regression of the mean performance is not unexpected. The agile material was older material covered in the first half of the course. Students were studying for and taking multiple, comprehensive final exams. Additionally, some students admitted to allocating study time strategically towards the courses where their effort would make the most difference in their final grade.

Peer Evaluations

The mean peer evaluation score was 92.5 with scores ranging from 81 to 100 on a scale whereby 100, 90, 80, 70, and 60 correspond to 5, 4, 3, 2, and 1 on a five-point scale.

DISCUSSION

For about ten years now, agile has made its way into IS curricula through its appearance in specific model curricula and textbooks. We contend that agile coverage in a project management course should include agile topics specifically related to project management and that cut across multiple PMBOK areas. The three areas we cover—daily standups, Kanban boards, and burndown charts—can be delivered in sequence and tie together well. They bring together PMBOK areas of scope, schedule, communication, and resource management.

The focus of this paper was to test whether an active learning pedagogy—Team-Based Learning—would be effective in delivering agile concepts and practices. It was thought that a pedagogy that used teams, collaboration, and face-to-face problem-solving would be fitting for learning about agile practices and values.

Results from three independent sections of two different project management courses provided evidence that students, both at the undergraduate and graduate levels, progressed as expected through a team-based learning pedagogy to eventually master agile learning outcomes. Individual performance on the iRAT varied, ranging from means of 52.73% to 71.61% across the three sections. Collaborative teams scored at means of 82.73% to 98.21% on team RATs. Individual performance on agile content on the midterm exceeded iRAT performance with means ranging from 78.06% to 89.55%, and did not drop off significantly on the final exam, as means ranged from 77.22% to 89.74%. While 85.96% of students mastered agile PM at the midterm, 75% did so by the final exam. The noticeable drop-off from the midterm to the final exam was evaluated further. It turned out that it was almost entirely due to undergraduate students. A total of 22 of 27 graduate students, or 81.4%, scored at a minimal (80%) passing level on the final exam agile items, while 19 of 29 undergraduate students, or 65.5%, scored at a minimal (70%) passing level. The proportion of passing students we expect for a learning outcome is usually 80%. This total was achieved by the midterm exam, but not for the final. For the final, the graduates achieved this, but the undergraduates did not.

Qualitative evidence observed during team application activities indicated that student teams were engaged, performing at a mostly correct level, were learning from mistakes, and were insightful. We attribute the design of the exercises, which enabled hands-on authentic problem solving and collaboration, enabled the discovery of insights and the detection and correction of mistakes.

Peer evaluation scores indicated that perceived student preparation and performance in the teams varied, and that, overall, a high degree of satisfaction with teammate performance was reported. The satisfaction with teammate performance is indirect evidence that students perceived the pedagogy to be effective.

Together, the pattern of student performance through the team-based learning activities, the midterm performance, the final exam performance (partially), and the qualitative evidence, along with reported team satisfaction provides evidence that our active learning approach for delivering agile methods coverage was effective.

IMPLICATIONS

The implication for this research is that effective agile project management can be taught and student learning assessed using the team-based learning pedagogy. In future studies, researchers might consider introducing variables such as student learning styles, student self-efficacy, and combinations of lecture and active learning methods to more fully explain variance in student performance. For undergraduate student retention, the problem of the final exam drop-off in performance is a phenomenon to be explored.

FUTURE DIRECTIONS

Future directions include development of higher order evaluative questions or case scenarios for the final exam that challenge students to compare traditional project management methods against agile methods. The current study demonstrated that students, especially at the graduate level, gained enough insight so as to be prepared for the higher order thinking. What remains is how to design a useful exercise, such as a case study or role play, that enables students to compare traditional and agile methods.

REFERENCES

Alliance, A. (2017). Agile Practice Guide, Project Management Institute, 2017: Agile Practice Guide (Vol. 1). Bukupedia.

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... & Kern, J. (2001). Manifesto for agile software development, http://agilemanifesto.org.

Cragin, J. (2004). Team-based learning in international situations. In L. Michaelsen, A. Knight, and L. Fink (Eds.), *Team-Based Learning: A Transformative Use of Small Groups in College Teaching*, Sterling, VA: Stylus Publishing.

Freeman, M. A. (2004). Team-based learning in a course combining in-class and online interaction. In L. Michaelsen, A. Knight, and L. Fink (Eds.), *Team-Based Learning: A Transformative Use of Small Groups in College Teaching*, Sterling, VA: Stylus Publishing.

Gorgone, J. T., Gray, P., Stohr, E. A., Valacich, J. S., & Wigand, R. T. (2006). MSIS 2006: model curriculum and guidelines for graduate degree programs in information systems, *ACM SIGCSE Bulletin*, 38(2), pp. 121-196.

Knight, A. B. (2004). Team-based learning: A strategy for transforming the quality of teaching and learning. In L. Michaelsen, A. Knight, and L. Fink (Eds.), *Team-Based Learning: A Transformative Use of Small Groups in College Teaching*, Sterling, VA: Stylus Publishing.

Landry, J., & McDaniel, R. (2016). Agile preparation within a traditional project management course. ***Information Systems Education Journal***, 14(6), 27.
Lindstrom, L., & Jeffries, R. (2004). Extreme programming and agile software development methodologies. *Information Systems Management*, 23(3), 41-52.

Loignon, A. C., Woehr, D. J., Thomas, J. S., Loughry, M. L., Ohland, M. W., & Ferguson, D. (2017). Facilitating peer evaluation in team contexts: The impact of frame-of-reference rater training. *Academy of Management Learning & Education*, 16(4), 562-578.

Longenecker, H. E. Jr., Babb, J., Waguespack, L. J., Tastle, B., & Feinstein, D. (2016). An update on the CIS curriculum project, *2016 Proceedings of the EDSIG Conference*. ISSN: 2473-3857, Las Vegas, Nevada USA v2 n4072.

Loughry, M. L., Ohland, M. W., & Moore, D. D. (2007). Development of a theory-based assessment of team member effectiveness. *Educational and Psychological Measurement*, 67, 505-524.

Marchewka, J. T. (2015). *Information Technology Project Management*, Fifth Edition, Wiley, ISBN-13: 978-1-119-03160-4.

McAvoy, J., & Sammon, D. (2005). Agile methodology adoption decisions: An innovative approach to teaching and learning. *Journal of Information Systems Education*, 16(4), 409.

Michaelsen, L. K. (2004). Team-based learning in large classes. In L. Michaelsen, A. Knight, and L. Fink (Eds.), *Team-Based Learning: A Transformative Use of Small Groups in College Teaching*, Sterling, VA: Stylus Publishing.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2004). *Team-Based Learning: A Transformative Use of Small Groups in College Teaching*, Stylus Publishing, LLC, Sterling, VA.

Michaelsen, L.K. & Sweet, M. (2008). Backward design: Excerpt from chapter two, In Michaelsen, L. K., Parmalee, D., Levine, R. & McMahon, K. (Ed.), *Team-Based Learning for Health Professions Education: A Guide to Using Small Groups for Improving Learning*, Sterling, VA: Stylus Publishing.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., Layton, R. A., Pomeranz, H. R., & Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self and peer evaluation. *Academy of Management Learning & Education*, 11 (4), 609-630.

Parmalee, D. X. & Michaelsen, L. K. (2010). Twelve tips for doing Team-Based Learning (TBL). *Medical Teacher*, 32, 118-122.

Project Management Institute (2017). *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*. 6th ed. Newtown Square, Pennsylvania: Project Management Institute, Inc.

Project Management Institute (2011). *PMI Agile Certified Practitioner Examination Content Outline*, Retrieved March 3, 2015 from http://www.pmi.org/Certification/~/media/Files/PDF/Agile/PMI_Agile_Certification_Content_Outline.ashx.

Schwalbe, K. (2019). *Information Technology Project Management*, Ninth Edition, Cengage Learning, Inc., Boston, MA.

Schmitz, K. (2018). A three cohort study of role-play instruction for agile project management. Journal of Information Systems Education, 29(2), 93-103.

Serrador, P., & Pinto, J. K. (2015). Does agile work?—A quantitative analysis of agile project success. *International Journal of Project Management*, 33(5), 1040-1051.

Sibona, C., Pourreza, S., & Hill, S. (2018). Origami: An active learning exercise for Scrum project management. Journal of Information Systems Education, 29(2).

Topi, H., Karsten, H., Brown, S. A., Carvalho, J. A., Donnellan, B., Shen, J., ... & Thouin, M. (2017). MSIS 2016 global competency model for graduate degree programs in information systems. *Communications of the Association for Information Systems*, 40(1), MSIS-i.

Topi, H., Valacich, J., Wright, R., Kaiser, K., Nunamaker, Jr., J., Sipior, J., & de Vreede, G. (2010). IS 2010: Curriculum guidelines for undergraduate degree programs in information systems. *Communications of the Association for Information Systems*, 26(18).

DECISION SCIENCES INSTITUTE
Designing Classroom Project Management for the Virtual Student Team

Diane Kutz
University of South Florida
Email: dckutz@mail.usf.edu

Matthew Mullarkey
University of South Florida
Email: mmullarkey@usf.edu

ABSTRACT

Virtual teams are a reality in today's work environment. Students can learn to navigate the virtual team environment in properly designed university courses. Many courses in project management are offered online and naturally employ technology to support these virtual team environments. As such, they provide an opportunity for the design of virtual team workspaces. This paper explores a series of iterative interventions used to design a collaborative virtual student team environment. Using a guided, emergent design approach we also propose a means to evaluate the effectiveness of this environment with regard to team cohesion, team commitment, and team collaborative climate.

KEYWORDS: Project Management, Technology Enabled Collaboration, Virtual Team Project Management, Online Project Management Classroom, Virtual Student Team

INTRODUCTION

The motivation for this research arose from (1) business demand that employees demonstrate an ability to operate in virtual team environments, and, (2) university requirement for needed improvements in educating students in online project management courses. The researchers' experience suggested that whether these students were undergraduates with limited work experience or graduate students with substantial work experience, the challenges were relatively similar – how do students effectively collaborate in an online environment to accomplish team-based assignments in a course? And, how might their fulfillment of virtual team assignments more fully prepare them for the needs of employers in the digital economy of distributed asynchronous team environments?

As project management faculty, the researchers observed that it appeared that the majority of students approached online virtual team assignments with a parsing activity using a form of 'divide and conquer' to complete the teamwork assignments in the course. In this approach, students segmented and separated the component parts of the project work required and set off to independently complete their requisite piece of work. At a final point in time, approaching the due date of the assignments, students then compiled the pieces and submitted the finished project. We found that this approach was less than ideal for both the quality of the project completed and for teaching students essential behaviors for collaboration in a virtual project team environment. For example, this parsing approach effectiveness was highly conditional on the virtual team's ability to actually determine the proper component parts of a project activity at the outset of the project. And, perhaps most importantly, often resulted in team conflict, missed deadlines, low satisfaction with group work, and poor team performance overall.

Instructors that routinely experienced difficulties with virtual team performance expressed a desire to develop and establish an effective collaborative virtual team environment. Researchers using a guided, emergent elaborated action design research approach in situ worked with these instructors to iteratively design, build, evaluate and implement an innovative approach to a technology-centric solution to more effective student virtual team performance.

This problem is practice-inspired because it comes from the need for universities and employers to train adults to successfully implement project work in virtual team environments. If the university is to train students to be effective in the virtual environment that is today's project management environment, then the university must teach the students how to effectively function in a virtual team environment.

This problem is research-ingrained because research has shown that positive team member relationships have a positive impact on team effectiveness and the design of effective online classroom environments to replicate work environments is essential to undergraduate and graduate student success in the workplace. Fernandez and Jawadi (2015) stated that "high quality relationships have been identified as a key factor in enhanced team performance." Virtual teams suffer from the lack of physical proximity that leads to organically developed relationships, and this is a hindrance on the development of good team member relationships (Handy, 1995), (Townsend, DeMarie, & Hendrickson, 2000). Therefore, a substitute for this physical proximity and organically developed relationships needs to be established in order to facilitate strong team relationships. And, the university project management course must offer a means for students to learn this essential capability.

LITERATURE REVIEW

Literature Review of Virtual Teams

The body of research on project management, effectiveness, and relationship building in traditional co-located teams is expansive (Pauleen, 2003). However, increasingly project managers are finding themselves charged with managing a project team where either all or some of their team members are working at a location other than the project manager's locale. "Virtual teams are here, and they are here to stay" (Bell & Kozlowski, 2002). In a 2014 article by Keith Ferrazzi, published in Harvard Business Review, Ferrazzi stated, "When my firm, Ferrazzi Greenlight, recently surveyed 1,700 knowledge workers, 79% reported working always or frequently in dispersed teams" (Ferrazzi, 2014).

There are many definitions of virtual teams with most definitions focusing on the nature of the dispersed locations and use and usefulness of technology (Martins, Gilson, & Maynard, 2004). For this study, virtual teams are defined as "teams whose members use technology to varying degrees in working across locational, temporal, and relational boundaries to accomplish an interdependent task" (Martins et al., 2004). Leadership in virtual teams requires a different mix of skills than those required of leaders in co-located teams (Pauleen, 2003). Virtual team leaders are challenged by having to essentially interpret personal and contextual signals from team members through the use of electronic communications with no visual cues or signals ("Global Work, Bridging Distance, Culture and Time," 1995).

Literature Review of Trust and Communication in Project Management

Trust is a vital ingredient when people work together (Ferrazzi, 2012). "Building and maintaining trust in the traditional, physical workplace is difficult enough, but the process is even tougher in a virtual environment, where people often have to work with people they haven't met in person" (Ferrazzi, 2012). According to Ferrazzi, upon initial formation team members tend to trust early, and they operate under what he calls "swift trust; however this swift trust will wain over time, but it can be replaced by "interpersonal trust" if the team successfully bonds and builds this more lasting trust (Ferrazzi, 2012).

According to Ferrell and Kline (2018), "In order for virtual teams to reach peak performance, trust and effective communication among team members must be developed." They further describe high trust teams where:

- Team members expect other members to follow through on commitments and interdependencies.
- There is a willingness to be vulnerable when uncertainty arises.
- Team members are confident that others take action for the good of the team (Ferrell & Kline, 2018).

However, by the very nature of their dispersion and reliance on technology, virtual teams encounter a multitude of barriers when it comes to building trust. They go on to state that barriers "are contextual and predominantly arise from the inherent challenges of geographical dispersion" and they discuss the importance of focusing on team building activities that focus on the sharing of personal values and experiences as well as the importance of communicating frequently (Ferrell & Kline, 2018).

Ferrell and Kline (2018) highlight that virtual teams need to use both communication-focused technology and coordination-focused technology effectively. They classify technology used in an attempt to replicate face-to-face communication, such as email, instant messaging, telephone, and video conference, as communication-focused technology. Whereas they classify technology focused on improving the coordination of tasks, such as project management software, desktop virtualization, file sharing tools, scheduling tools, and co-creation tools, as coordination-focused technology. They assert that "communication-focused technology is the backbone of trust-building communication in virtual teams, coordination-focused technology refers to a complementary set of tools that can be leveraged more peripherally to enhance trust and coordination among team members" (Ferrell & Kline, 2018).

Literature Review of Team Development and Team Building

Kozlowski and Ilgen (2006) make a distinction between team development and team building. They state that team building refers to formal efforts to make modifications to existing processes, whereas team development is a more informal process initiated by the team members to establish social structure and work processes that are effective for the team. Research typically views team development as an initial forming process of a team and, because teams are not always static in membership, as a continuous process over the life of the team (Bell & Kozlowski, 2002).

Hackman (1992) focused on the importance of factors that are considered “glue” that bonds team members to each other. When Steve Jobs was asked about teamwork how he envisions teams working together in a 1995 interview, he shared a parable of rocks in a rock tumbler (Elmer-Dewitt, 2011). Jobs said that “… through that group of incredibly talented people bumping up against each other, having arguments, having fights sometimes, making some noise, and working together they polish each other and they polish the ideas, and what comes out are these really beautiful stones” (Elmer-Dewitt, 2011).

The Jobs’ parable offers not only the parallel between rocks in a rock tumbler and team members on a team but also the lesson of the liquid and grit. Jobs’ neighbor in the analogy did not just put rocks in the tumbler, rather he “…put them in the can with a little bit of liquid and a little bit of grit powder…” (Elmer-Dewitt, 2011). The liquid and grit are the professional relationships built between team members working closely together that result in interpersonal trust and lead to effective team performance. Bell and Kozlowski (2002) state that “team orientation represents the affective bonds that connect members to the team and its mission”. More recent literature purports that team relationship building is comprised of four distinct approaches that may be combined (Aga, Noorderhaven, & Vallejo, 2016). They find that these approaches are goal-setting, developing interpersonal relations, clarifying roles, and employing problem-solving techniques (Aga et al., 2016). Ferrazzi (2012) makes the point that:

> “When assembling a virtual team, managers often assume that people will mainly be interested in what their fellow team members can do, as opposed to who they are as individuals. Wrong! When looking at a resume, CV, or bio of someone, people will often latch onto personal details, such as hobbies and other outside interests, including charities the person supports. Why? Because they want to get a better sense of that individual and to see if they might have anything in common” (Ferrazzi, 2012).

Bruce Tuckman (1965) introduced the stages of team development model, identifying that teams go through four distinct phases: forming, storming, norming, and performing. Each of these phases includes some amount of relationship building, with the greatest amount happening in the forming phase. The researchers purport that if a team spends adequate time in the forming phase, establishing relationships, that their journey through storming and norming will be brief and they will move on to performing much faster than if they neglect to take the time and effort to adequately navigate the forming phase.

Teams progress through these stages naturally in co-located teams as team members interact spontaneously throughout the day; however, leaders must take a more active role in regards to relationship development in virtual teams where spontaneous communication is rare, and team communication exists over computer-mediated technologies (Liao, 2017). “The importance of relationship building in a virtual environment and methods to build relationships are significant factors when practitioners engage in virtual work” (Pauleen, 2003). There have been numerous studies on relationship building in traditional co-located teams (Pauleen, 2003) and organizations spend a lot of money on team building training and exercises, yet most of this training has been focused on co-located teams. Nonetheless, the prevalence and reliance on virtual teams are growing, and it is vital for organizations, and by extension instructors, to gain a comprehensive understanding of the importance of team building and its relationship to project success.

Discussion and Research Questions

Project teams must operate in virtual environments where team members are separated geographically, temporally, and technologically. In these environments, effective teams must possess a different set of capabilities needed to build trust, establish relationships, communicate, and coordinate the project work. Existing literature further suggests that these capabilities must be exercised throughout the course of the project for the most productive team effort. Consequently, university instructors must consider approaches that build these capabilities proactively and avoid the typical student approach to shortcut virtual team project management.

The secondary literary review research, however, offers little in the way of project management course design to inform instructors seeking to educate students in virtual team effectiveness in project management courses. Our work with students in existing project management courses offered an opportunity to methodically introduce and evaluate a series of course improvements designed to increase virtual team effectiveness through improved student team performance.

Our research asks the following research questions:

RQ1: How can successful virtual team project management environments be adapted to an educational project management course environment?

RQ2: How can we co-create and design with students an effective virtual team project management environment?

RQ3: What might the evaluation approach for the elements of this new virtual team project management environment take?

METHODS

The researchers used an action research grounded approach to co-create and co-evaluate with students the design of an adaptive virtual team project management environment. We chose a guided, emerge design-centered approach consistent with the design science research (DSR) paradigm (Hevner, March, Park, & Ram, 2004). DSR acknowledges that information systems and technology-enhanced socio-technical systems often result from an iterative approach to the build and evaluation of multiple alternatives. When conducted in situ with participants of the socio-technical system, an action design science research (ADR) approach is ideal for principles of researchers' and participants' reciprocal shaping and mutually influential roles (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011). Using an elaborated ADR (eADR) method (Mullarkey & Hevner, 2019) we worked with students to iterate through stages of Diagnosis, Design, and Implementation establishing the critical principles of an effective virtual project management team environment in a university course. Throughout the process, we evaluated each artifact of this socio-technical system leading to a number of modifications that became part of the planning and execution of successive iterations of the system.

eADR Interventions and Diagnosis
As a soci-technical system creation and evaluation was critical to both the online course and the virtual project management team environment, it proved critical to design interventions

with student and researcher (including faculty) participants that were informed by theory and prior research in a Diagnosis stage. The first intervention was to institute a requirement that all student teams host an initial video conference meeting, via an online platform of their choice, focused only on team introductions and professional relationship building prior to any discussion of team assignments. The teams were also required to post minutes of this meeting to their team homepage in the Canvas Learning Management System for instructor review.

The goal of these initial meetings focused on establishing a collaborative atmosphere with open communications among team members. This intervention was successful but it led to a need to establish a more complete model for an effective collaborative virtual teamwork, hereby entitled Collaborative Virtual Team Project Management Classroom Environment. From the initial diagnosis in this eADR stage, we explicated the requirements of this environment and determined implementation and evaluation plans for the said environment, focusing on continuous improvement each semester in the courses taught by one of the researchers. “A design science approach places emphasis on achieving clarity in the goals and underlying theoretical constructs for a new artifact and carefully evaluating how well the new artifact meets those goals” (McLaren, Head, Yuan, & Chan, 2011).

As an output of the Diagnosis stage, the researchers established a model of the collaborative virtual team project management classroom environment (Figure 1). Each of the interventions listed in this model are design components expected to improve team cohesion, commitment, and collaborative climate leading to improved project management effectiveness. The interventions selected correspond to the communication-focused technology and coordination-based technology presented by Ferrell and Kline (2018).

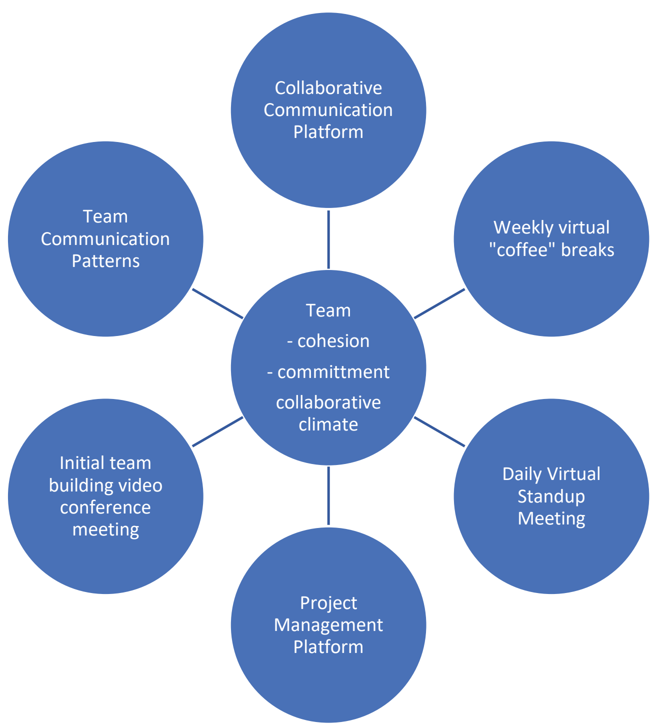


Fig. 1. Collaborative Virtual Team Project Management Classroom Environment

eADR Design Interventions

Graduate Students enrolled in MG 670-I03 (Leading High-Performance Teams) during the spring 2019 semester were tasked with serving as Subject Matter Experts (SMEs) to recommend an adaptive virtual team project management environment. They were required to conduct an evaluation in teams during each virtual team intervention in order to identify key performance indicators (KPI's) and Objectives and Key Results (OKR's) for each of the proposed interventions. Their research assignment appears in Appendix A, and the student team composition is listed in Appendix B. The student submissions were reviewed, and their intervention recommendations appear below, in order of intervention. Because we could involve multiple project teams, we could integrate evaluations from each and generate a richer triangulation of the evaluations across subject teams.

Team Intervention #1: Initial Team Building Video Conference Meeting – Team #4

Team #4 discussed the importance of an initial team building video conference meeting to establish initial rapport and to set expectations and team goals. They proposed that Zoom would be the best video conferencing platform upon which to hold this meeting as it is easy to use, affordable, and allows for a large number of participants.

Team #4 stressed that it was important that the goal of the initial meeting was to bring about informal iteration among team members – to see each other, self-introduction, pursue any common interests and hobbies. They also recommended that the meeting include games, a personality test, and an informal discussion on a topic such as a favorite place to vacation, a favorite movie, or another topic of interest. Interestingly, this is consistent with the work Team #5. These items served as "ice-breakers" to get everyone talking and sharing items that are not project related but which helped build team rapport.

Team Intervention #2: Daily Virtual Standup Meeting – Team #1

Team #1 wrote about their daily "standup meeting". The goal of the typical standup meeting is to quickly cover project progress since the prior meeting, discuss intended progress over the next period and to address any 'blockages' to progress or 'stuck' tasks that needed to be addressed.

As this was a virtual team, the standup meeting was held via a video conferencing platform. The team considered Google Hangouts and Skype, and Canvas Big Blue Button (Integrated conferencing in the Canvas Learning Management System (LMS)) as options, but they indicated that the project team should select the platform that best suits their needs as they all have relatively similar features. Most specifically, they pointed out that the team should choose one that team members are already familiar with in order to reduce the learning curve.

Team #1's evaluation indicated that they felt that a daily virtual standup meeting would be difficult and not necessary for students in a project management course so they proposed a weekly video standup meeting. The goal of the standup meeting would remain the same, but the frequency would be weekly.

Team Intervention #3: Weekly Virtual Video "Coffee Break" – Team #5

The Virtual Video Coffee Break is supposed to be a time that is dedicated to a virtual team to recreate the spontaneous conversations that would traditionally take place in an office environment. The CEO of GitLab, a remote-only company with 200 employees, compares this type of intervention to "the time someone working in an office might spend chatting while walking to meetings, grabbing a coffee in the break room, or having lunch together in person" (People Matters Editorial Team, 2017).

Team #5 entitled this intervention the "WeeklyTeam Recess." Team #5 purported that, instead of a specific weekly meeting focused on spontaneous conversations, the team should implement this intervention at the beginning of every meeting. They recommended that teams spend at least 10 minutes on this activity and that it takes place on the virtual conference platform whereupon the team holds their other team meetings.

Team Intervention #4: Team Collaborative Communication Environment – Team #3

The Collaborative Communication Environment (CCE) consists of the tools selected by the team in order to interact and accomplish tasks in a virtual setting. The makeup of a CCE is dependent on a multitude of variables, such as team size and project complexity with the following as components:

- Basic Messaging / Text Communication
- E-mail
- Phone – person to person and conference call
- Video Chat
- Small Video Conference (no add-ons, just basic video)
- Large Video Conference (with add-ons, such as screen and file share)
- Project Management Platform
- Document Control and Change/Revision Tracking
- Centralized Team Operations Center
- Virtual Water Cooler

Team #3 proposed a CCE Component Selection Matrix after several iterative evaluations of the components involved (Table 1).

Component	**Recommendation**
Basic Message / Text Communication	GroupMe
E-mail	Canvas Course Platform
Phone – Person to person	Cell Phones
Phone – Conference Call	Freeconferencecall.com
Video Conference	Zoom
Project Management Platform	Wrike
Document Control & Change Revision Tracking	Google Docs
Virtual Water Cooler	Yammer or Sococo

Table 1. CCE Component Selection Matrix

Team #3 stressed that while this matrix serves as a basic recommendation, consideration should be made to ensure that all team members will have the necessary access and training to utilize each tool.

The above components are pretty traditional but the Virtual Water Cooler component is a relatively unique component worthy of additional discussion. According to Paris (2016), virtual water coolers serve three purposes:

- It gives a team a chance to bond.
- It can inspire new ideas.
- It reinforces ties and a sense of belonging (Paris, 2016).

Team #3 indicated that "Water coolers allows the team to step outside of the typical work environment of e-mails and phone calls and collaborate with each other to achieve common ground or bounce ideas off one another". They proposed the use of Yammer or Sococo. The transition to some virtual water coolers may be easier for those familiar with social media platforms as Paris (2016) indicated that Yammer is very similar to Facebook. Sococo actually "… replicates an actual office with conference rooms, offices, and hangout spots" (Paris, 2016).

Team Intervention #5: Project Management Software Tool – Team #6

Team #6 researched many different project management software platforms before coming to their recommendation of using Wrike. They selected this product because of its flexibility, clear visibility of key information regarding not only the big picture but also the little details, their integrated forms, as well as their impeccable integration with email and Google Apps.

Team Intervention #6: Team Communication Patterns – Team #2

Team #2 first defined communication patterns as to how information is disseminated between team members. They referenced an MIT-sponsored study of 21 different organizations over seven years, that identified three key elements of communications patterns: energy, engagement, and exploration (Pentland, 2012). This same study noted that social time is attributed to over 50% of the positive responses in team communication patterns (Pentland, 2012).

Team #2 concluded the following in regards to team communication patterns:

- Ensure face-to-face communication is implemented in a minimum of 30 minutes per week utilizing online video conference tools discussing assignments and projects.
- Incorporate in one video session out of the month, socialization and do not discuss school or assignments. Use this session for bonding and establishing trust and openness within the team.
- Create a log and document after each virtual session and include the percentage of the collaboration that contributed to the assignment or project. Track results until the project is completed.
- Include clarification exercises into video sessions for example: after someone asks a question, clarify and confirm what was asked to establish clarity and understanding. This can be included in the documentation log for tracking.
- Ensure recap meeting notes are established after each face-to-face meeting and confirm action items and next steps for each team member.

RESULTS

The various project team iterations through the eADR Design stage lead to the following model of a collaborative virtual team project management environment.

Design Principle 1: Initial Team Building Video Conference Meeting

This activity is a virtual project team requirement that the newly formed virtual team, relatively quickly upon team formation, conduct an initial team building video conference with the specified purpose of establishing initial relational rapport. This video conference will be hosted on the Zoom platform with a requirement that all participants share their video camera. It will be focused specifically on informal communication among team members and facilitated with the use of ice breaker activities, such as the two truths and a lie activity or asking questions such as where would you vacation if money was not a consideration. It is imperative that this meeting does not focus on project details, but rather focuses on establishing initial team rapport.

Design Principle 2: Weekly Virtual Standup Meeting

This activity is a requirement that the team spends a short period of time every week dedicated to discussing the accomplishments the day before, the plan for the day ahead, and any issues that have the potential to hinder their performance of the planned activity. The purpose of the meeting is to keep the team up to date on team progress as well as to identify any potential delays that need to be proactively addressed. The concurrent, authentic evaluation of this design artifact identified that a daily standup meeting would not be practical for an online course, so a weekly standup meeting via the Zoom Video Conference platform was made standard for the initial video conference meeting.

Design Principle 3: WeeklyTeam Recess

The purpose of the Weekly Team Recess is an attempt to replicate spontaneous communication found in co-located teams and to build team rapport and the connection that co-located teams make organically by way of their close proximity to each other. The intervention in situ found that a weekly organized video conference would not be practical for student teams in an online class so this activity is designed to be a part of the team's regular video conference meetings. The title of this intervention was identified as the WeeklyTeam Recess and was made a part of the regular weekly video team meetings. The teams will spend the first 10 minutes or so of each team video conference meeting focusing on socializing and ice breaker activities, such as those proposed for Intervention #1.

Design Principle 4: Team Collaborative Communication Environment

The purpose of the team collaborative communication environment is to define the tools that the virtual team will use to perform all communications. This includes messaging, phone calls, video conferencing, and document control with change/revision tracking. The Design Stage established the standards for the collaborative communication environment for the online classroom as displayed in Table 1 (above).

Design Principle 5: Project Management Software Tool

This design principle established the project management software that the team will use to track all of the project details, including tasks, schedules, resources, etc. In the initial intervention, the student team recommended the Wrike platform. However, the researchers (aka faculty) noted the need to integrate the project writing/publishing software with external tools, such as Outlook and Google Docs. Working together, the team chose to use the Monday.com platform and established a partnership Monday.com.

Design Principle 6: Team Communication Patterns

The intervention of team communications patterns refers to the methods by which the team members communicate throughout the project. The teams determined that the student teams would be required to conduct all meetings via the standard video conference platform, Zoom, using video so all members can see each other throughout the meeting. This iteration determined that conferences that do not incorporate the video aspect resulted in participants not focusing fully on the conversation at hand. Additionally, this design principle requires the student teams to document each virtual meeting, including the percentage of collaboration by team member and to post this documentation to their team Canvas homepage. This documentation would not only include the minutes of the meeting, but it would also recap all action items and next steps for each team member.

Design Summary

A summation of the complete model of collaborative virtual team project management classroom environment is presented in the table below.

Initial Team Building Video Conference Meeting	Video Conference focused on socializing and building team rapport - Software: Zoom
Weekly Virtual Standup Meeting	Quick weekly meeting focused on what we did, we are we doing, are there any issues - Software: Zoom
WeeklyTeam Recess	First 15 minutes of every meeting focused on social conversation - Software: Zoom
Team Collaborative Communication Environment	Groupme, Canvas Course Platform, Cell Phones, Freeconferencecall.com, Zoom, Google Docs
Project Management Software Tool	Software: Monday.com
Team Communication Patterns	All meetings conducted via Zoom with Video enabled Documentation of all meetings include percentages of collaboration, a recap of action items, next steps

DISCUSSION AND CONCLUSION

Initially, the researchers focused on establishing a collaborative virtual team project management environment in the classroom that would mimic that of the corporate environment. This was accomplished with the help of the graduate students enrolled in MG 670-03 (Leading High-Performance Teams) during the spring 2019 semester, as the majority of them were all currently working in a corporate virtual team project management environment. Student teams were assembled and each team selected an intervention to research in order to identify key performance indicators (KPI's) and Objectives and Key Results (OKR's) for their intervention.

The student created a virtual team environment was then examined in situ with the intention of adapting it to the classroom environment, while keeping the tie to the corporate environment. This was undertaken in order to meet the goal of more fully preparing students for the needs of employers in the digital economy. The eADR research activity leads to the proposed specific design principles for implementation in a future project management course. The complete student assignment is provided in Appendix A and the student team composition is listed in Appendix B.

Lastly, the researchers developed the evaluation criteria utilized to evaluate the resultant model of a collaborative virtual team project management environment, in terms of its effectiveness in improving team cohesion, team commitment, and team collaborative environment. The evaluation of an implemented course will be the topic of future research, as explained in the next section.

The use of the eADR method within the DSR paradigm offered researchers a means to co-design the principles for an improved virtual team project environment in an online course. Doing so offers faculty the opportunity to integrate the resultant design principles into a course with a high(er) expectation of success. This approach can be used in a multitude of course design and avoid the situation where an instructor uses a more common trial-and-error approach to course design or one reliant exclusively on theoretical models.

FUTURE RESEARCH DIRECTION

The above-designed model of a collaborative virtual team project management environment will be implemented in an online classroom and its use will be monitored for its effectiveness in improving team cohesion, team commitment, and team collaborative environment. According to Huang et. al (2003), "Team cohesion refers to the mutual attraction of team members and the closeness that team members feel towards each other" (Huang, Wei, Watson, & Tan, 2003). Accordingly, they state that "team commitment refers to the team spirit, a sense of loyalty and dedication to a team" and "team collaboration climate refers to the degree to which a team can work well together" (Huang et al., 2003). The intention is to administer a survey to the student teams at the beginning of the semester, before introducing the model, and at the end of the semester after they have worked with the model for the entire semester. This survey will be an online survey embedded in their Canvas course. The questions to be used are based on existing measurement instruments adapted from (Huang et al., 2003) and they can be found in Appendix C.

Appendix A

Assignment #1: MG 670-I03 : Virtual Team Environment Research Paper

There are six teams and six interventions, so each team will research one of them. The selection process will be handled on a first come, first served basis, and the selection will be handled based on replies to the Announcement Post. Please do not begin the paper until you receive confirmation that your team was assigned that intervention. Once you have received confirmation, please complete your paper on the intervention of choice by your team.

The team will research and write up a paper meeting the following goals:

- Provide an overview of the intervention
- Provide a detailed description of the intervention and all that it entails
- Discuss the objective of the intervention
- Explain the key performance indicators (KPI's) that should be evaluated for the intervention (the measures by which the effectiveness will be evaluated)
- Explain the specific Objectives and Key Results (OKR's) of the intervention (the characteristics of the intervention)
- Discuss any interventions other than those six listed here that your team feels should be considered in establishing an effective virtual team environment.
- Present specific recommendations for the implementation of their intervention in a future project management course.

Team Intervention #1: Initial Team Building Video Conference Meeting

Note: One of the OKR's should be the video platform for the meeting.

Team Intervention #2: Daily Virtual Standup Meeting

Note: One of the OKR's should be the video platform for the meeting.

Team Intervention #3: Weekly Virtual Video "Coffee Break"

Note: One of the OKRs should be the video platform for the meeting and you should also determine a more suitable name than "Coffee Break"

Team Intervention #4: Team Collaborative Communication Environment

Note: This includes but is not limited to file sharing, chat, communication guidelines for when to use what form of communication (phone call, email, chat, group message, conference call, video meeting, etc). This environment may also include a project management tool.

Team Intervention #5: Project Management Software Tool

Note: This may include a team collaborative environment.

Team Intervention #6: Team Communication Patterns

Note: This should include all team communication guidelines.

Appendix B

Student Team Composition

Team #1:

Shantel Perry, Preston Ray, Erica Sykes, and Melissa Van Otten

Team #2:

Tianna Bateman, Jennifer Bekeny, Kimberly Chambers, and Zoie Hill

Team #3:

Katie Coan, William McEllen, Deborah Taylor, and Kaila Timbs

Team #4:

Aubrey Hudson, Vrinda Pisharody, Kayla Stewart, and Harlan Winston II

Team #5:

Lindsey McKie, Kyle Winn, Kristen Nohrstedt, and Joshua Studdard

Team #6:

Waverly Polak, Matthew Rider, and Siddharth Pandya

Appendix C

Survey Questions adapted from (Huang et al., 2003)

Questions measuring team cohesion (Seashore, 1954)

- Do you feel that you are really a part of your student work team?
  (1) Really a part of my work team.
  (2) Included in most ways.
  (3) Included in some ways, but not in others.
  (4) Don't feel I really belong too much.
  (5) Don't feel I really belong at all.

- If you had a chance to do the same kind of work in another student work team, how would you feel about moving to another team?
  (1) Would want very much to stay where I am.
  (2) Would rather stay where I am than move.
  (3) Would make no difference to me.
  (4) Would rather move than stay where I am.
  (5) Would want very much to move.

- How does your work team compare with other student teams on each of the following points?
  (1) Very much better to (5) Very much worse.

  a. The way people get along together.
  b. The way people work together.
  c. The way people help each other.

Questions measuring team commitment (Larson and LaFasto, 1989)
(1) False to (4) True.

- Achieving our team goal(s) is a higher priority than any individual objective.
- Team members believe that personal success is achieved through the accomplishment of the team goal(s).
- Team members are willing to devote whatever effort is necessary to achieve team success.

Questions measuring collaborative climate (Larson and LaFasto, 1989)
(1) False to (4) True.

- We trust each other sufficiently to accurately share information, perceptions, and feedback.
- We help each other by compensating for individual shortcomings.
- We can trust each other to act completely and responsibly in performing our individual tasks.

REFERENCES

Aga, D. A., Noorderhaven, N., & Vallejo, B. (2016). Transformational leadership and project success: The mediating role of team-building. *International Journal of Project Management, 34*, 806-818. doi:10.1016/j.ijproman.2016.02.012

Bell, B. S., & Kozlowski, S. W. J. (2002). A Typology of Virtual Teams: Implications for Effective Leadership*,* 14.

Elmer-Dewitt, P. (2011). Steve Jobs: The Parable of the Stones. *Fortune*.

Fernandez, D. B., & Jawadi, N. (2015). Virtual R&D Project Teams: From E-Leadership To Performance. *Journal of Applied Business Research, 31*(5), 1693.

Ferrazzi, K. (2012). How to Build Trust in a Virtual Workplace. *Harvard Business Review*.

Ferrazzi, K. (2014). Getting Virtual Teams Right. *Harvard Business Review*(12), 120.

Ferrell, J., & Kline, K. (2018). Facilitating Trust and Communication in Virtual Teams. *People & Strategy, 41*(2), 30-35.

Global Work, Bridging Distance, Culture and Time. (1995). *Behavioral Science*(2), 160.

Hackman, J. R. (1992). Group influences on individuals in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology., Vol. 3, 2nd ed.* (pp. 199-267). Palo Alto, CA: Consulting Psychologists Press.

Handy, C. (1995). Trust and the Virtual Organization. *Harvard Business Review, 73*(3), 40-50.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75-105. doi:10.2307/25148625

Huang, W. W., Wei, K. K., Watson, R. T., & Tan, B. C. (2003). Supporting virtual team-building with a GSS: an empirical investigation*,* 359.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest, 7*(3), 77-124. doi:10.1111/j.1529-1006.2006.00030.x

Liao, C. (2017). Leadership in virtual teams: A multilevel perspective. *Human Resource Management Review, 27*, 648-659. doi:10.1016/j.hrmr.2016.12.010

Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual Teams: What Do We Know and Where Do We Go From Here?*,* 805.

McLaren, T. S., Head, M. M., Yuan, Y., & Chan, Y. E. (2011). A Multilevel Model for Measuring Fit between a Firm's Competitive Strategies and Informatioin Systems Capabilities. *35*, 909-A910.

Mullarkey, M. T., & Hevner, A. R. a. u. e. (2019). An elaborated action design research process model. *European Journal of Information Systems, 28*(1), 6-20.

Paris, Jennifer. (2016). Why You Need a Virtual Water Cooler. *Remote.co.*

Pauleen, D. J. (2003). An Inductively Derived Model of Leader-Initiated Relationship Building with Virtual Team Members. *Journal of Management Information Systems, 20*(3), 227-256.

Pentland, A. (2012). The New Science of Building Great Teams: How researchers at MIT's Human Dynamics Laboratory decoded the secrets of superior group performance*,* 60.

People Matters Editorial Team. (2017). Virtual Coffee Breaks: An espresso shot for remote workers.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. In (Vol. 35, pp. 37-56): MIS Quarterly.

Townsend, A. M., DeMarie, S. M., & Hendrickson, A. R. (2000). Virtual Teams: Technology and the Workplace of the Future*,* 69.

Tuckman, B. W. (1965). Developmental Sequence in Small Groups. *Psychological Bulletin, 63*, 384-399.

DECISION SCIENCES INSTITUTE

The Impact of Mobile Order Ahead Apps on QSR queuing

Thomas R. Robbins
East Carolina University
Email: robbinst@ecu.edu.edu

ABSTRACT

Order ahead mobile apps are becoming increasingly popular at Quick Serve Restaurants. These apps allow customers skip the line when arriving at the store. We investigate the hybrid queuing system created by these applications. Using discrete event simulation, we examine the impact these apps can have on system performance.

KEYWORDS: Queuing Systems, Simulation Analysis, Retail Operations

INTRODUCTION

The availability and use of order ahead mobile applications has become quite common in the quick serve restaurant business. Apps are available from restaurants including McDonald's, Chick-fil-A, Chipotle, Starbucks, Moe's, Noodles and Co., Subway and many others. The Wall Street Journal reports that food and drink apps were downloaded more than 155 million times in 2017; a 35% increase over 2016 (Stevens 2018). My local Starbuck's manager reports 7.5% of her business now comes from mobile orders.

While the functions provided vary from app to app, most allow users to order and pay for their food remotely. When arriving at the restaurant the user is able to skip the line and pick up their pre-made order, often at a special Mobile Order Pick-Up location. These apps create a hybrid queuing system that alters both the real and perceived behavior of the queuing system. In this paper we explore the implications these hybrid queues have from both the customer and service provider's perspective.

The functionality of order ahead applications varies from restaurant to restaurant. Some have built in loyalty programs that allow customers to earn discounts and free items. Most apps have integrated mapping and GPS systems to allow customers to find the nearest restaurant location. Most apps also allow users to reorder previous orders and some allow users to set up favorite items for easy reordering, and nearly all the apps support some level of menu item customization. Virtually all the apps support a pay function that manages the financial transaction as part of the ordering transaction, though some also allow a pay at the store option.

The order processing and fulfillment process also has some variations from app to app, but most follow a basic format. The user picks their restaurant and selects their menu items. The customer can then select the scheduled time for their order to be ready. Most applications have a minimum lead time, but also allow the order to be scheduled further into the future. Minimum times are generally about 10-15 minutes, and most apps allow scheduling pick up an hour or more in the future. Once the order is placed the financial transaction is processed and an order ticket is printed at the store. Some apps, such as McDonalds and Chick-fil-A, do not finalize the order until the customer is at the store and clicks the "I'm here" button. GPS tracking disables the button until the customer is at the store location. This helps prevent accidental orders to the wrong store, but significantly reduces the order ahead time and significantly alters the queuing system behavior. When the customer arrives at the store, they generally skip the line and go to a special pick up area. If their food is ready, they pick it up immediately and the service is complete. If the food is not ready, because the customer arrived early, or the order was late, they wait in a separate queue for their order to be completed.

The remainder of this paper is organized as follows. In the next section we review the relevant literature. We thee describe the Order Ahead Hybrid Queuing System and our simulation approach. The next section presents an analysis of our simulation results for an overall scanning experiment, we then present a separate experiment that isolates the impact of the order ahead percentage. We conclude with a summary of our findings.

LITERATURE REVIEW

Queuing systems are widely studied in the operations literature. Queues may form whenever jobs arrive to a service location; when all servers are busy the jobs wait in queue until a server becomes available. The behavior of the queuing system is governed by several key parameters. The arrival process characterizes the statistical distribution of new jobs presenting themselves for service. The service process characterizes the statistical distribution of the service performed on the job. Performance varies significantly based on the systems utilization, the long-term proportion of time that each server is actively working versus waiting for a new job to service, and the number of servers present in the system. Under several widely accepted assumptions, such as exponentially distributed inter arrival and service time and a first come first served protocol, closed form solutions are available and widely known (Gross 2008). The M/M/N model describes the long run statistical characteristics of a multi-server queue with exponential interarrival and service time. Standard performance metrics for basic queuing systems include average wait time, average total service time, average queue length, and system utilization.

As systems become more complex, and assumptions are relaxed, it becomes increasingly difficult to accurately characterize queuing system behavior. The introduction of balking and/or reneging, customers who arrive but decide not to enter the queue, or who leave the queue before being serviced, leads to the M/M/N/k + M queuing model (Gans, Koole, and Mandelbaum 2003). Customers who balk or renege are said to be impatient, and their behavior can dramatically alter the behavior of the queuing system. This phenomenon has been widely studied in the call center space. As other assumptions are relaxed, for example the service or arrival process distributions, exact closed form solutions become less viable and approximation or discrete event simulation become more popular methods of analysis (Law 2007).

Queues can be physical, virtual, or hybrid combinations of the two. In a physical queue customer stand in line. Traditional retail outlets are served by physical queues. A key characteristic of a physical queue is the visibility of the queue's length and the speed of service. Customers enter the line with full visibility of the number of customers ahead of them and have constant visibility of the queue's length, giving them insight into the speed of service. In a multi-server system, there may be separate queues for each server, but more commonly the queues are pooled into a single queue served by multiple servers. Pooling is more efficient technically but may be perceived as a longer wait due to the longer pooled queue and this may lead to increased balking. In some cases, the choice of queues is based on a customer class, for example priority lines at airport check-in and security, or at hotel check-in.

In a virtual queue there is no physical line for customers to stand in, instead some proxy for a customer is placed in a virtual line. The quintessential example of a virtual queue is the call center, but other examples include restaurant pager queues, ticket queues, or even print queues. Hundreds of papers have been published on call centers; (Gans, Koole, and Mandelbaum 2003) and (Aksin, Armony, and Mehrotra 2007) provide excellent summaries and literature reviews. An important characteristic of the virtual queue is limited visibility, and therefore limited insight into how long the potential wait might be. In a general call center, the caller has no idea how many callers are already in queue, though in some systems a system announcement will give an estimated hold time.

A variation of the physical queue is the ticket queue. In a ticket queue the customer is given a number and they begin service when their number is called. By limiting the visibility of the queue, ticket queues have the potential to significantly alter both the customer's perception of the queue, and their behavior (Kuzu 2015). Ticket queues have been used in environments from the DMV to the deli counter at the local grocery store. In a ticket queue the customer can compare their number to the now being served number and observe the rate at which customers are served to estimate a wait time. Pager queues are often used at full-service restaurants and an estimated wait time is often provided along with the pager, but without special effort on the part of the customer to monitor progress, no updates are available to the initial estimate. While physical queues often implement a first come first serve prioritization mechanism, virtual queues can easily implement alternative service disciplines.

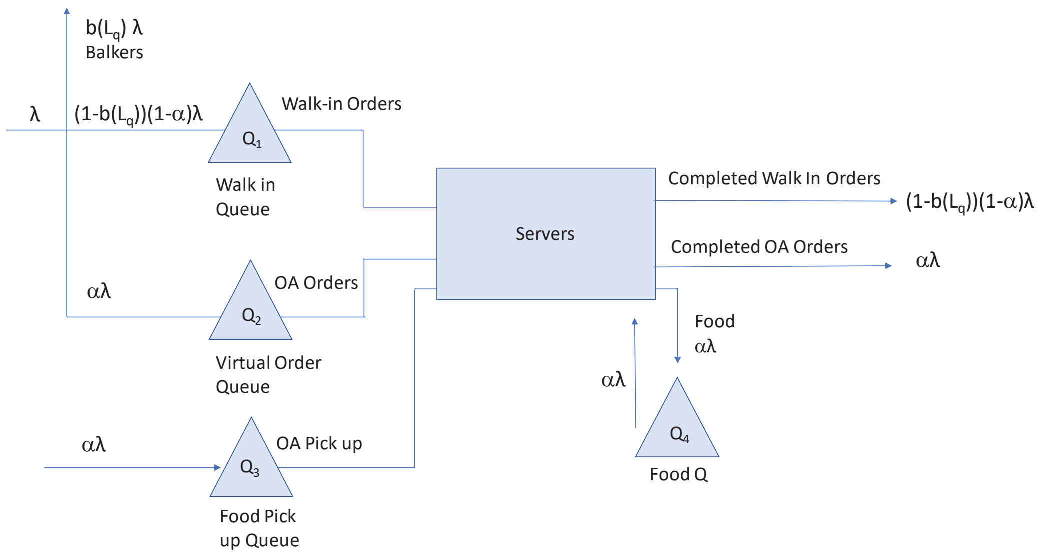
We define a hybrid queue as a linked queuing system that contains at least one physical and one virtual queue. A well-known example is the *Fastpass* system used on popular attractions at Disney theme parks. The hybrid queuing system includes a physical queue for walk up customers along with a virtual queue for Fastpass customers. The Disney Fastpass system is described in (Cope, F. Cope, and E. Davis 2011). A hybrid queuing system for airport security checks is proposed and analyzed via simulation in (de Lange, Samoilovich, and van der Rhee 2013). In both these systems a subset of customers are put in a virtual queue and assigned a reservation time. These customers then join the physical queue at a priority point at their reservation time. A basic version of a hybrid queueing system is used in full-service restaurants that accept reservations. Walk in customers are queued based on arrival time, while customers with reservations jump to the front of the line at their reservation time.

Order ahead apps create a somewhat different type of queuing system; an environment we refer to as the Order Ahead Hybrid Queue (OAHQ). The system contains multiple queues, physical and virtual. The system has the potential to impact the retail establishment in multiple ways. The order ahead capability will clearly impact the service experience of the order ahead customer, and potentially the walk in customer as well. Larson (1987) discusses the issue of social justice in queuing systems that result from what he terms slips and skips. A skip occurs when a customer arrives and is serviced before a current customer, the initial customer having been skipped over. The skipped customer feels victimized, having been treated unfairly. The slipping customer may feel a sense of satisfaction based on their reduced waiting time, but may also feel guilt for having violated a social norm. Stevens (2018) reports on customers who order ahead, but still wait in line to avoid the stigma associated with cutting the line. From a financial perspective the faster service time, along with the potential discounts earned, may lead the order ahead customer to increase their dining frequency. Multiple studies investigate the relationship between faster service and increased revenue (Akilimalissiga, Sukdeo, and Vermeulen 2017, Perdikaki, Kesavan, and Swaminathan 2011, Allon, Federgruen, and Pierson 2011, Lu et al. 2013). Additionally, by removing some customers from the physical queue the order ahead capability may reduce the perceived waiting time for walk in customers and reduce the number of customers who balk.

THE ORDER AHEAD HYBRID QUEUE

The Order Ahead Hybrid Queue system contains a combination of virtual and physical queues. While there are multiple variations for configuring the order ahead process, we develop a model that can represent most configurations. A diagram of the system is outline in Figure 1.

Figure 1 - Queuing System

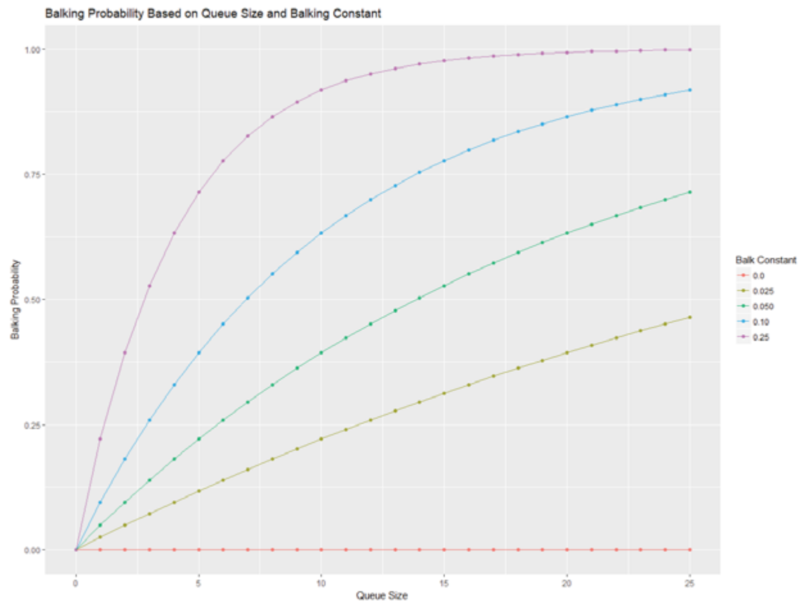


Orders arrive at a rate λ. α percent of the order arrive as virtual order ahead orders, while 1-α percent of the arrivals are physical walk-in customers. Walk in customers observe the physical queue and a proportion of them determine the line is too long and they balk. The decision to balk is based on the length of the queue and the balking constant ($b(L_q)$). The balking decision is based on the number of customers currently in queue. In our model the probability the customer balks is given by

$$p_B = 1 - e^{-k_b n} \qquad (1)$$

Where *n* is the observed number of customers currently in the queue and k_B is the balking constant. If k_B is zero, then customers are infinitely patient, and none balk. As k_B increases the likelihood of balking increases. Figure 2 shows the probability a new customer balks based on queue size and a range of balking constants.

Figure 2 - Balking Probability Graph



The time required to process an order by an average server is a lognormally distributed random variable with mean μ^{-1} and standard deviation σ_{μ}. Walk in customers who decide not to balk enter the walk-in queue (Q1). OA Orders arrive at a rate of $\alpha\lambda$. When the order arrives, it is placed in a virtual queue (Q2). Each order has an associated ready time, which we assume in our model is 20 minutes after the order is placed. Each order is assigned an early start and late start time. In our model we assign the early-start time as 4 times the average service time, and the late start time as 2 times the average service time.

The system includes 3 physical queues and 1 virtual queue.

1. Walk-In Queue: the line that walk-in customers enter to wait for service. Serviced on a first come first serve basis.
2. Order Ahead Virtual Queue: the queue of order ahead orders. Serviced based on scheduled time.
3. Pick Up Queue: a physical queue where order ahead customers wait to pick up their items.
4. Meal Queue: a physical queue where food is kept awaiting customer pickup. Serviced based on order number.

A pool of homogeneous servers service the entire queuing system. Idle servers accept work based on the following prioritization scheme.

1. OA Orders due in sooner than the late start time
2. Walk in Customers
3. OA Orders due sooner than the early start time.

In order to evaluate the performance of the queuing model, we conduct a series of designed experiments. In the first experiment we evaluate the system over a wide variety of input characteristics. Based on the assumptions for our system discussed previously, we define the following set of 6 experimental factors. High and low values are defined to cover a broad range of service scenarios.

Table 1-Experimental Factors

	Factor	**Low**	**High**
1	Proportion ordering ahead (α)	5%	40%
2	Offered System utilization (ρ)	25%	90%
3	Average service time	2 mins	8 mins
4	Balking Constant (*k*)	0	.25
5	Number of Servers (*n*)	1	5
6	Average OA Arrival Shift	-2 mins	+2 mins

The offered utilization represents the average system utilization that would occur if all arrivals were served, in other words without balking. The average OA Arrival Shift represents the mean of the distribution of the physical arrival of an order ahead customer, relative to the scheduled ready time. We assume that the physical arrival time is characterized by a normal distribution with a standard deviation of 2 minutes. So, with a zero valued shift, a customer would on average arrive at the agreed upon time, but half of the customers would arrive early. A negative shift implies a larger proportion arriving before the scheduled time.

Given the relatively large number of experimental factors, a well-designed experimental approach is required to efficiently evaluate the experimental region. A standard approach to designing computer simulation experiments is to employ either a full or fractional factorial design (Law 2007). However, the factorial model only evaluates corner points of the experimental region and implicitly assumes that responses are linear in the design space. However, this system is likely to exhibit non-linear responses, for example queue length will have a decidedly non-linear relationship with utilization. Based on this we chose instead to implement a Space Filling Design based on Latin Hypercube Sampling as discussed in (Santner, Williams, and Notz 2003). Given a set of d experimental factors and a desired sample of n points, the experimental region is divided into n^d cells. A sample of n cells is selected in such a way that the centers of these cells are uniformly spread when projected onto each of the d axes of the design space. We chose our design point as the center of each selected cell. This experimental design allows us to select an arbitrary number of points for any experiment.

Simulation Approach

Our model is evaluated using a straightforward discrete event simulation model at each design point. The purpose of the model is to predict the long term, steady state behavior of the queuing system. The model generates random numbers using a combined multiple recursive generator (CMRG) based on the Mrg32k3a generator described in (L'Ecuyer 1999). Common random numbers are used across design points to reduce output variance. To reduce any start up bias we use a warmup period of 250 customers, after which all statistics are reset. The model is then run until approximately 50,000 customers have been serviced and summary statistics are collected. For each design point we repeat this process for 25 replications and report the average value across replications. Our initial analysis is based on an experiment with 1,000 design points.

The specific process for each replication is as follows. The input factors are chosen based on the experimental design. The average arrival rate is calculated based on the specified service time, number of servers, and offered utilization rate as

$$\lambda = \rho N \mu \tag{2}$$

That arrival rate is then used to generate Poisson arrivals for the replication. Each new customer generated includes an exponentially distributed interarrival time, and a lognormally distributed service time. When a customer arrives a Bernoulli random variable is used to determine if it is an OA or walk in customer. For walk-in customers the probability of balking is determined based on the current walk-in queue length and the balking constant according to equation (1). Another Bernoulli random variable used to evaluate the customer's balking decision. Our model assumes customers that join the walk-in queue remain patient and do no abandon. We also assume that since order ahead customers pay at order time they do not balk.

When the OA Customer arrives, the model determines if the food is ready. If it is ready the customer picks up the food and ends service. Our model assumes a pick-up area, so customers can pick up their prepared food without server involvement. If the customer arrives before their food is ready that are placed in queue 3 and wait until the food is finished. In this scenario services ends when the food is ready, and the total service time is equal to their wait time in Q3.

Over the course of the simulation we collect statistics on the following performance metrics:

1. **Average Queue Size**: the average size of the walk-in physical queue.
2. **Maximum Queue Size**: the maximum size of the walk-in physical queue.
3. **Realized Utilization**: the proportion of time servers are actively working on customer orders.
4. **Average Wait Time**: average time a customer spends waiting; from arrival time until their order is taken.
5. **Proportion Waiting**: the proportion of walk-in customers that must wait prior to their order being taken.
6. **Average Time in System**: the average time a customer spends in the restaurant, from arrival time until their order is filled.
7. **Proportion Balking**: the proportion of walk in line customers who balk.
8. **Number of Times Jumped**: the average number of times a walk-in customer is jumped over by an order ahead customer.
9. **Proportion Jumped:** the proportion of walk-in customers that are jumped by an order-ahead customer.
10. **Average # Jumping Over**: the average number of customers an order ahead customer jumps over.

After all design points have been evaluated, we re-run the experiments. All design factors are repeated with the exception of the OA percentage. In the second run the OA percentage is forced to zero. This allows us to determine the impact the OA customers had on the system. Common random numbers are utilized to minimize the variance and isolate the impact of the OA stream.

SIMULATION ANALYSIS

Baseline Metrics

Before examining the impact of the Order Ahead process, we review the baseline experimental region. Our initial design parameters give us a wide range of scenarios. Key performance metrics for the baseline case are shown in Table 2.

Table 2- Baseline Performance Metrics

	Min	Avg	Median	Max	Range	StdDev	Skewness
Average Queue Size	0.003	0.456	0.292	4.579	4.576	0.525	2.690
Max Queue Size	4.340	9.948	8.770	46.080	41.740	4.488	2.789
Realized Utilization	0.249	0.547	0.559	0.876	0.627	0.163	-0.130
Average Wait	0.302	99.04	50.57	1,265.41	1,265.11	142.19	3.522
Proportion Waiting	0.9%	31.5%	30.1%	80.3%	79.3%	18.8%	0.326
Average Time in System	125.2	399.0	376.7	1,667.0	1,541.8	195.7	1.850
Proportion Balking	0.0%	4.1%	2.4%	21.8%	21.8%	4.3%	1.267
Number of Times Jumped	0.000	0.000	0.000	0.000	0.000	0.000	NA
Proportion Jumped	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	NA
Avg Number Jumping Over	0.00	0.00	0.00	0.00	0.00	0.00	NA

The data in Table 2 shows that our experimental region covers a very wide range of operational scenarios. The key performance metrics cover the range of practical operating scenarios. The distribution of these metrics strongly positively skewed. They also demonstrate the significant volatility present in queuing systems. Figure 3 illustrates that while the median queue size is small, the distribution of avenge queue size is strongly right skewed, with average queue sizes ranging up into the range of 2. While an average queue size of 2 seems small recall that in queues the queue is volatile. Figure 4 shows the distribution of the maximum queue size; demonstrating that queues can balloon up to more than 20 even with a low average queue.

Figure 3 – Baseline Average Queue Size

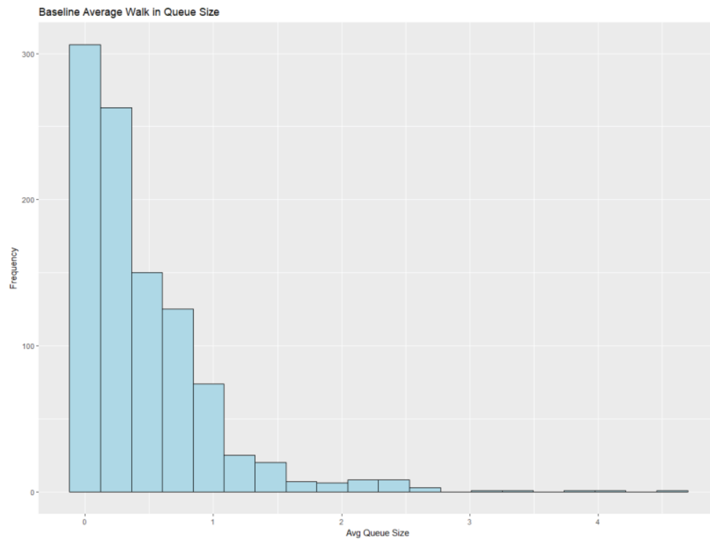
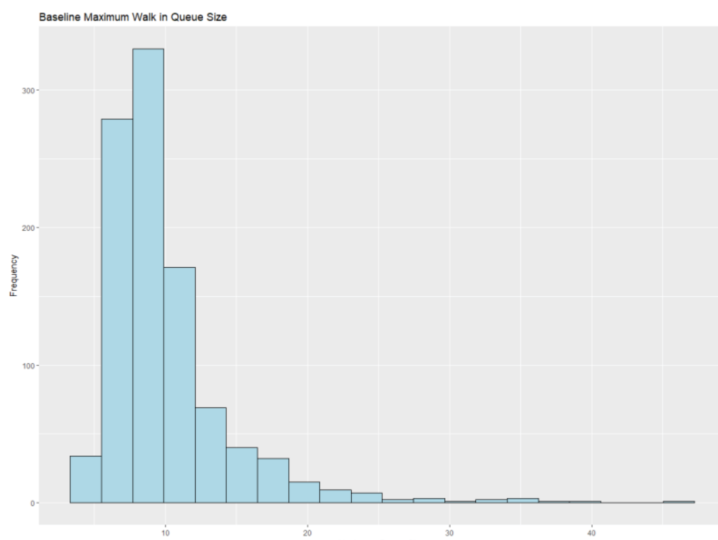


Figure 4-Baseline Maximum Walk in Queue Size



Impact of Order Ahead

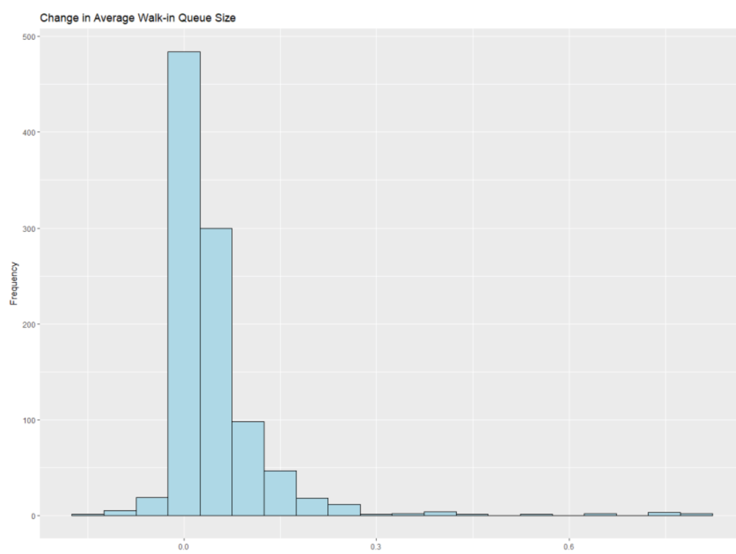
We now turn our attention to the impact Order Ahead customers have on the queuing system. We first examine the impact of OA across the experimental region, Table 3 summarizes the difference that occurs in each of core performance metrics as order ahead is introduced. Recall, that in our experimental design space the proportion of orders that are order ahead varies from 5% to 40%.

Table 3- Change in Key Metrics with OA Capability

	Min	Avg	Median	Max	Range	StdDev	Skewness	pvalue
Average Queue Size	-0.130	0.048	0.024	0.817	0.947	0.085	4.761	2.59E-62
Max Queue Size	-2.320	0.239	0.200	4.320	6.640	0.618	1.173	3.14E-32
Realized Utilization	-0.039	0.000	-0.001	0.085	0.123	0.014	1.192	8.34E-01
Average Wait	0.294	113.78	55.51	1411.59	1411.30	163.01	3.266	3.03E-88
Proportion Waiting	-9.9%	-0.7%	-0.3%	8.0%	17.9%	2.0%	-0.777	5.27E-25
Average Time in System	6.9	66.1	58.6	444.8	437.9	41.1	2.127	6.30E-280
Proportion Balking	-2.5%	0.4%	0.2%	3.6%	6.1%	0.5%	1.545	1.89E-93
Number of times Jumped	-1.990	-0.174	-0.078	0.000	1.989	0.241	-2.888	0.00E+00
Proportion Jumped	-22.2%	-5.1%	-3.5%	0.0%	22.2%	4.6%	-1.143	0.00E+00
Avg Number Jumping Over	-4.65	-0.53	-0.34	0.00	4.65	0.58	-2.299	0.00E+00

Figure 5 shows a histogram of the change in average queue size when order ahead customers are introduced. Note that a positive number indicates a reduction in the observable (walk-in) queue length. Order Ahead reduces the queue length in 94.8% of our design points.

Figure 5- Change in Average Queue Size



In Figure 6 we examine the distribution of the change in the average wait time for walk in customers. This figure shows the data is highly skewed in the opposite direction. In 82.8% of the design space the average wait time increases. Taken together these two graphs show that walk-in customers face shorter but slower moving lines when Order Ahead customers are present.

Figure 6-Change in Average Wait Time

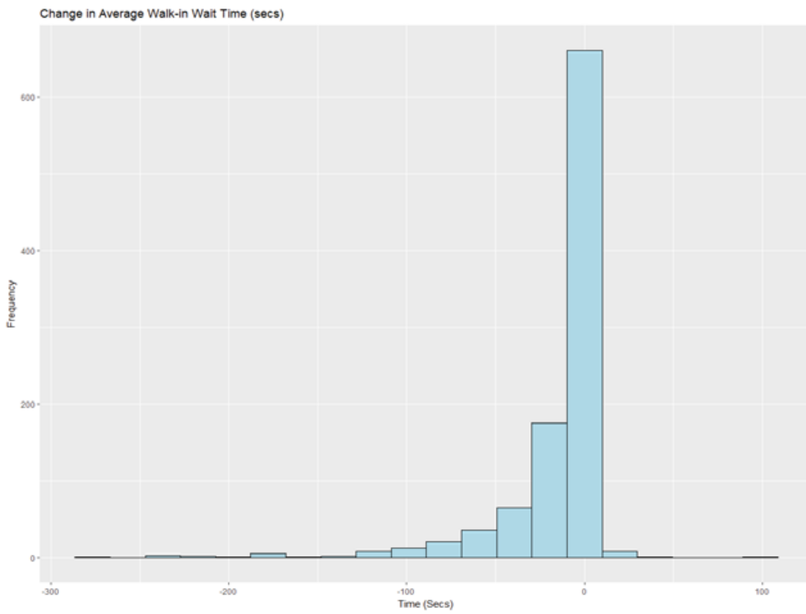
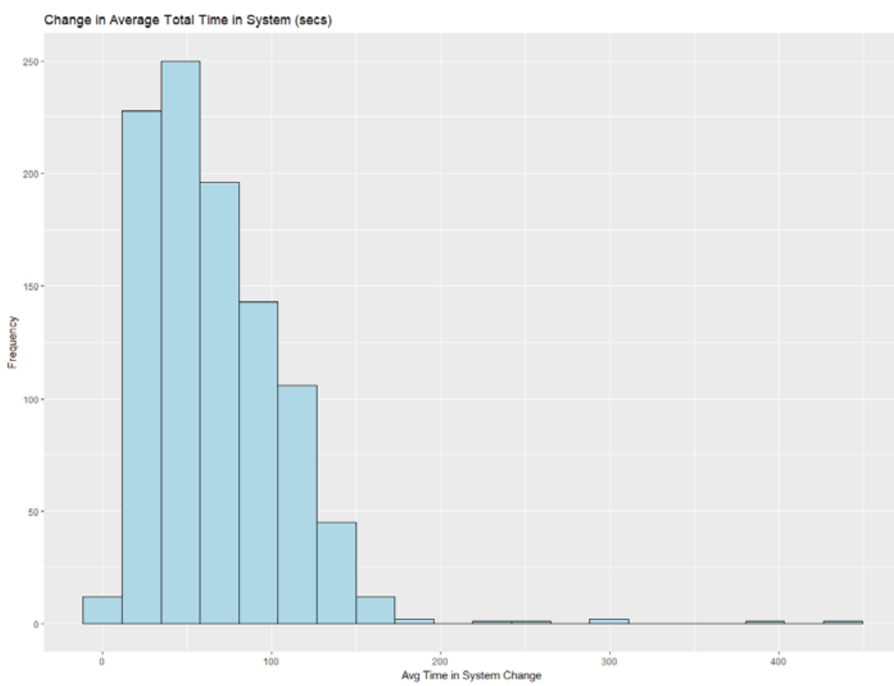


Figure 7 shows a histogram of the change in the Total Time in System for all customers. The overall time in system is reduced in all simulation scenarios.

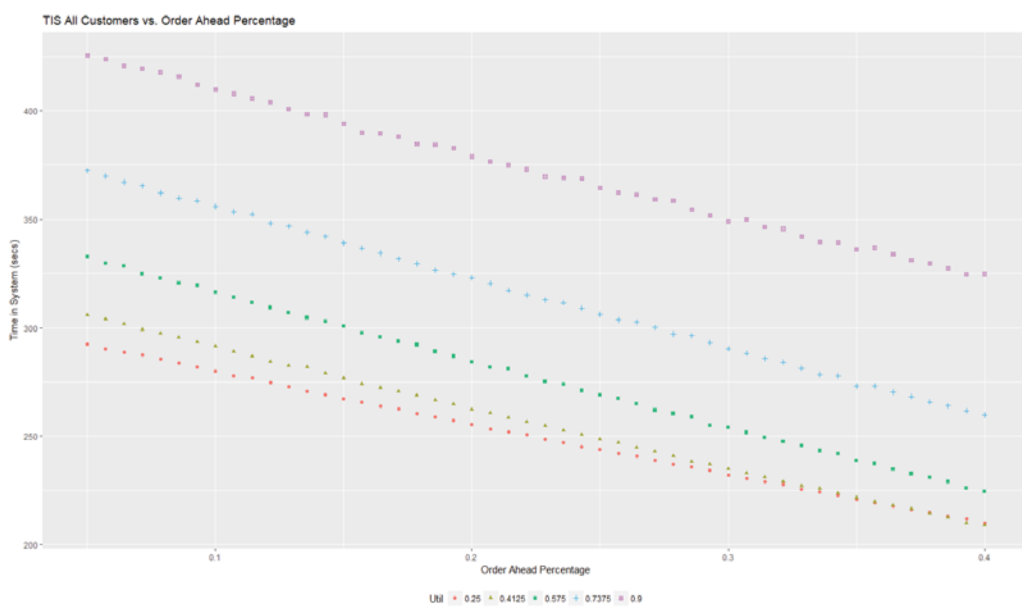
Figure 7-Change in Average TIS



Experiment 2
We now conduct a second statistical experiment. The purpose of this experiment is to isolate the impact that the proportion of customers that order ahead has on overall operations. We use a simple experimental design with two factors, Proportion Ordering Ahead (α) and Offered System Utilization (ρ). We evaluate the model at 5 different utilization levels and 50 different values for the proportion ordering ahead, for a total of 250 design points. This provides us with the ability to analyze how performance metrics vary with order ahead percentage for a range of system utilizations. All the other experimental factors are set at their midpoint, with the discrete parameter Number of Servers (*n*) rounded up to 3.

We first examine the impact on the Time in System. Figure 8 shows the Time in System for all customers, while Figure 9 shows it for Walk In Customers and Figure 10 shows it for Order Ahead customers.
Table 4 shows how the TIS metric changes as the order ahead percentage varies from it's low to high values. The impact of order ahead customers is relatively clear from these graphs. Overall Time in System is reduced as wait time is dramatically reduced for Order Ahead customers. The reduction is effectively linear with the proportion of order ahead customers. As demonstrated in experiment 1, the improvement for overall customers comes in part at the expense of walk in customers. Time in System for walk in customers increases as the proportion of order ahead customers increases.

Figure 8- Time in System All Customers



The increase is very slight, almost trivial for low to moderate levels of utilization. Only when utilization is high is there a meaningful increase. For lower levels of utilization, the order ahead service can be provided during idle time, even when the order ahead percentage is high. Finally, the time in system for order ahead customers, while very short, does increase with order ahead rate. As order ahead rates increase the average wait time increases roughly between 30-50%. Walk-in customer time is barely impacted under low to moderate system utilization levels. In the 2nd highest utilization level total time increases 7%, while in the highest utilization level the time in system increases about 20% as the proportion of order-ahead customers increases from 4% to 40%. The primary effect of the order-ahead capability is to reduce the overall time in system. As the order ahead percentage increase from a low (5%) level to a high (40%) level the overall time in system is reduced by about 30%. This is achieved by shifting food preparation service to the time before a customer arrives, when they are in-transit to the store.

Figure 9-Time in System Walk in Customers

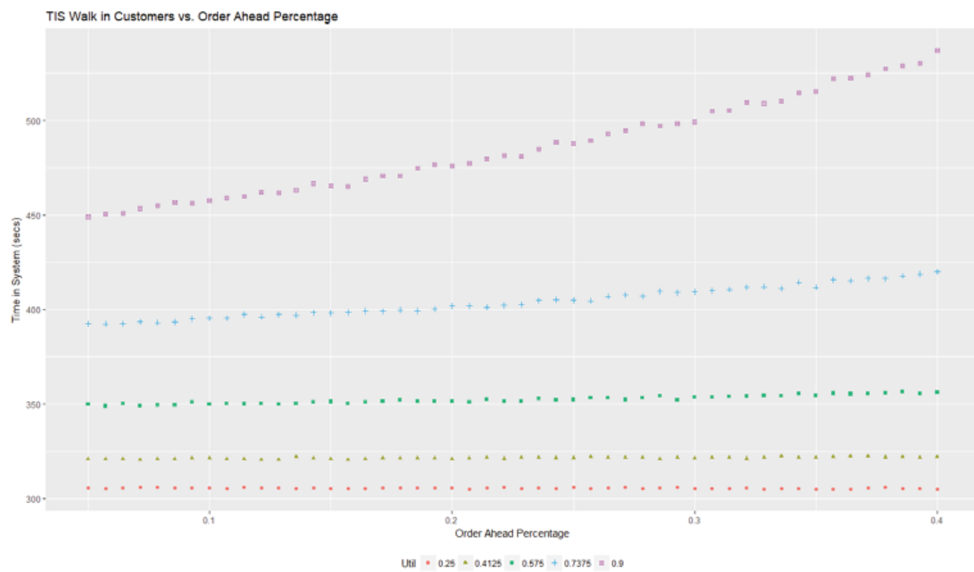


Figure 10-Time in System Order Ahead Customers

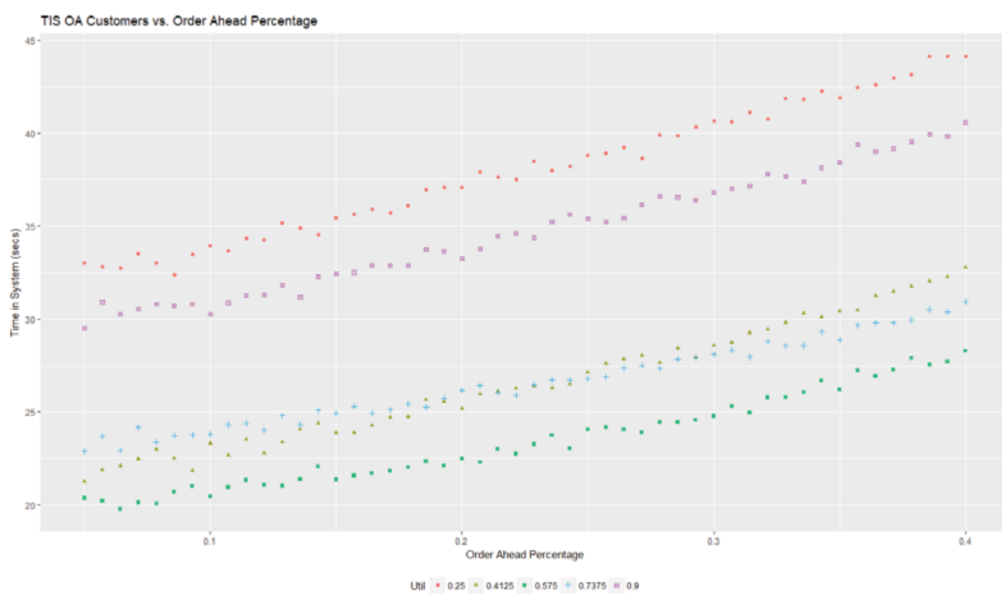


Table 4- TIS Metrics over Range of Order Ahead Proportion

	Time in System											
	All Customers				Order Ahead Customers				Walk In Customers			
Utilization	Low	High	Chg	% Chg	Low	High	Chg	% Chg	Low	High	Chg	% Chg
25.00%	292.4	210.0	-82.4	-28.2%	33.0	44.1	11.1	33.7%	305.3	304.9	-0.4	-0.1%
41.25%	305.9	208.9	-96.9	-31.7%	21.2	32.8	11.6	54.4%	320.8	322.0	1.2	0.4%
57.50%	332.7	224.3	-108.4	-32.6%	20.4	28.3	7.9	39.0%	349.6	355.9	6.3	1.8%
73.75%	372.5	259.8	-112.8	-30.3%	22.9	30.9	8.0	35.2%	392.3	419.9	27.6	7.0%
90.00%	425.4	324.8	-100.6	-23.7%	29.5	40.6	11.1	37.5%	448.8	537.0	88.3	19.7%

SUMMARY AND CONCLUSIONS

Mobile Order Ahead apps are becoming increasingly popular in the quick serve restaurant space. Customers who use these apps see significant benefits. Their orders are prepared prior to their arrival and their wait time is virtually eliminated. These customers go to the front of the line and often get their food almost immediately. Additional benefits often accrue to the customer through loyalty programs and discounts. Benefits also accrue to the restaurant as they capture customer data and potentially increase dining frequency.

Our analysis indicates that overall the mobile application increases the efficiency of the system as a whole. By completing food preparation before the customer arrives the total waiting time for all customers is reduced. The walk-in customer also sees the benefits as overall system efficiency improves although they may feel slighted as customers arrive and skip the line.

This preliminary analysis examines the effect of order-ahead apps over a broad range of operating characteristics. Our analysis is primarily descriptive, analyzing the impact of a reasonable set of operating assumptions. From an operations perspective, further prescriptive analysis is possible to determine optimal operating policies for scheduling walk-in vs. order-ahead service. From a marketing perspective further analysis of the loyalty and purchase frequency benefits that come with order-ahead systems is warranted.

REFERENCES

Akilimalissiga, S., N. Sukdeo, and A. Vermeulen. 2017. "The delivery of service quality to increase customer repurchase behaviour and customer satisfaction at fast food outlets in central johannesburg, South Africa." 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 10-13 Dec. 2017.

Aksin, Zeynep, Mor Armony, and Vijay Mehrotra. 2007. "The Modern Call-Center: A Multi-Disciplinary Perspective on Operations Management Research." *Production and Operations Management* 16 (6):665-688.

Allon, Gad, Awi Federgruen, and Margaret Pierson. 2011. "How Much Is a Reduction of Your Customers' Wait Worth? An Empirical Study of the Fast-Food Drive-Thru Industry Based on Structural Estimation Methods." *Manufacturing & Service Operations Management* 13 (4):489-507. doi: 10.1287/msom.1110.0343.

Cope, Robert, Rachelle F. Cope, and Harold E. Davis. 2011. *Disney's Virtual Queues: A Strategic Opportunity To Co-Brand Services?* Vol. 6.

de Lange, Robert, Ilya Samoilovich, and Bo van der Rhee. 2013. "Virtual queuing at airport security lanes." *European Journal of Operational Research* 225 (1):153-165. doi: https://doi.org/10.1016/j.ejor.2012.09.025.

Gans, Noah , Ger Koole, and Avishai Mandelbaum. 2003. "Telephone call centers: Tutorial, review, and research prospects. ." *Manufacturing & Service Operations Management* 5 (2):79-141.

Gross, Donald. 2008. *Fundamentals of queueing theory*. Edited by Donald Gross. 4th ed. ed. Hoboken, N.J. :: Wiley.

Kuzu, Kaan. 2015. "Comparisons of Perceptions and Behavior in Ticket Queues and Physical Queues." *Service Science* 7 (4):294-314. doi: 10.1287/serv.2015.0116.

L'Ecuyer, Pierre. 1999. "Good Parameters and Implementations for Combined Multiple Recursive Random Number Generators." *Operations Research* 47 (1):159-164.

Larson, Richard C. 1987. "OR Forum—Perspectives on Queues: Social Justice and the Psychology of Queueing." *Operations Research* 35 (6):895-905. doi: 10.1287/opre.35.6.895.

Law, Averill M. 2007. *Simulation modeling and analysis*. 4th ed, *McGraw-Hill series in industrial engineering and management science*. Boston: McGraw-Hill.

Lu, Yina, Andrés Musalem, Marcelo Olivares, and Ariel Schilkrut. 2013. "Measuring the Effect of Queues on Customer Purchases." *Management Science* 59 (8):1743-1763. doi: 10.1287/mnsc.1120.1686.

Perdikaki, Olga, Saravanan Kesavan, and Jayashankar Swaminathan. 2011. *Effect of Traffic on Sales and Conversion Rates of Retail Stores*. Vol. 14.

Santner, Thomas J., Brian J. Williams, and William Notz. 2003. *The design and analysis of computer experiments*, *Springer series in statistics*. New York: Springer.

Stevens, Laura. 2018. "Why Pre-Order Food Then Wait in Line: ‘I Generally Do Not Want to Seem Like an Ass’." *Wall Street Journal*, May 2, 2018.

DECISION SCIENCES INSTITUTE

A Framework to Address Software Backdoor Vulnerabilities

Anil Singh
Robert C. Vackar School of Business and Entrepreneurship
University of Texas at Rio Grande Valley
Email: anil.singh@utrgv.edu

Sandra Henderson
W. P. Carey School of Accountancy
Arizona State University
Email: schenderson@asu.edu

ABSTRACT

Back doors - in the software context - are passwords that – for useful or malicious reasons - bypass normal authentication. Backdoors, therefore, are exploited to hack into IT systems such as the accounting software. Such attacks are pervasive and can lead to enormous costs to the organization including loss of brand value, loyalty to customers and compliance penalties. Although there are efforts to address specific areas of the issue, so far, there are no integrated solution to address this problem. By including legal, ethical, international, policy and technical aspects, we develop a holistic policy level framework to address this issue.

KEYWORDS: Backdoors, Framework, Security, Auditing

INTRODUCTION

Back doors are not only a threat to accounting software, but also to the internal controls of an organization and as such must be addressed by auditors. This issue can only be tackled with multiple aspects ranging from policy issues, technical resolutions to international cooperation. After performing a literature review discussing general and specific Backdoor related threats in AIS, we propose a framework to address threats from Backdoor vulnerabilities.

LITERATURE REVIEW

General IT related threats

While the role of Information Technology in adding organizational value has become increasingly clear (Brynjolfsson & Saunders 2010; Masli et al, 2011; Tambe & Hitt 2012), it has come with its own set of issues. Information technology (IT) weaknesses pose a threat to organizations (Haislip, Peters, & Richardson, 2016). Such weaknesses take longer to fix, are associated with subsequent restatements, less accurate forecasts, higher audit fees, and lower earnings quality, and at times may lead to executives losing their titles (Kim, Richardson, & Watson, 2018). This threat is further exacerbated as Accounting is increasingly becoming dependent on Information Technology (Bansah, 2018). Pervasive use of computers, along with recent compliance requirements, has increased the importance of Information Technology (IT) in auditing services. Consequently, there is growing emphasis on assuring the integrity of internal control related to information systems for complying with the internal control assurance requirements of Section 404 of the Sarbanes-Oxley Act of 2002 (SOX) (U.S. House of Representatives) (Borthick & Bowen, 2008). Therefore, auditors need to place high priority on enhancing the overall effectiveness of auditors' work on internal control, particularly with respect to information systems (O'Donnell & Rechtman, 2005). Vulnerabilities such as backdoors pose a serious threat to AIS. Using a comprehensive integrated approach, this paper proposes a framework to address this threat.

Accounting and Information Technology Security

The role of Information Technology in the financial and accounting routines is becoming more pervasive. Controls are increasingly being embedded in IT systems (Allen, et al. 2006). Accounting information systems however are just as vulnerable as typical IT systems. For example, the Ransomware attack in Ukraine was through a software update coupled to an accounting system (Srinivasan, 2017; Perlroth, Scott, & Frenkel 2017). This threat has been researched all over the world. The research findings (Table 1) suggest that AISs have been found to have weaknesses all over the world. Therefore, it is imperative that these pervasive weaknesses be studied, and addressed. This has led to more cross-discipline research involving accounting and IT (Brandas, Stirbu, & Didraga, 2013; Mihalache, 2011). This paper attempts to do the same in one area of weakness: Backdoors. Problems such as backdoors in accounting software stand at the intersection of IT and Accounting disciplines (Murthy, 2016). The solution framework in the paper may also solve other challenges in IT and AIS.

Table 1: Studies in AIS control issues across the world

AUTHOR(S)	YEAR	COUNTRY	DETAILS
Rajeshwaran & Gunawardana	2009	Sri Lanka	Security controls inadequately implemented.
Abu-Musa	2004	Egypt	Weaknesses in security controls identified.
Abu-Musa	2006	Saudi-Arabia	Financial losses due to internal and external computerized accounting information systems security breaches.
Zweilf	2009	Jordan	Threats to accounting systems including unauthorized access.
Hanini	2012	Jordan	Threats to information systems originate from internal sources that include intentional threats.
Polo & Dima	2013	Kenya	Failure in many cases is due to weak governance and audit monitoring rather than the inadequacy of systems.
Muhrtala & Ogundeji	2013	Nigeria	Organization employees and outsiders' key threats to information assets.
Hood & Yang	1998	China	Information security was found to be inadequate.
Singh & Gupta	2017	India	Results highlight the importance of consistent top management

			support, organizational information security culture and a proper monitoring system for information security effectiveness in organizations.
Riad	2009	United Kingdom	Companies suffer from various security incidents.

Backdoors

Backdoors – often called hardcoded passwords - are userids, passwords or other data, in the source code that are - for practical or malicious purposes - used to bypass regular security protocols. Practical non-malicious backdoors help in case of lost or forgotten passwords. While there are non-malicious practical applications of backdoors, organizations become vulnerable to potential hidden malicious code. By their hidden nature, backdoors in software are difficult to verify and can only be hypothesized. The malicious potential of backdoors is well documented (Table 2). History is replete with examples of hacking attacks through backdoors that consequently cost companies millions of dollars in damages (Kim et al., 2011).

Table 2: Representative list of backdoor attacks in the recent past

YEAR	DETAILS
2016	Some older routers built on the WiMAX technology contain backdoor accounts that appear to have been introduced somewhere along the devices' supply chain.
2016	Wickedly Clever USB Stick Installs a *Backdoor* on Locked PCs
2016	PokemonGo Ransomware installs Backdoor Account and Spreads to other Drives
2017	74 countries hit by WannaCrypt ransomware backdoor
2017	*Backdoor Code* Discovered in Popular Bitcoin Mining Equipment

Over the years, these backdoors have experienced new developments (Klaus, 1997). It could be a programmer writing the original code or an intruder with access to change the code (Landwehr et al, 1994). A malicious coder can hide code to make the software weak or bypass normal authentication checks (Dai et al, 2012). Such hidden code can originate from, or result in spyware, adware, trojans and ransomware (Table 3). The recent cases of ransomware attacks and data breaches have brought to the forefront the importance of addressing backdoors in software (Deepa & Thilagam, 2016). Just like any other IT security situation, it is a tradeoff between security and convenience – or in some cases- as in the Apple case (McLaughlin, 2016) suggests – a tradeoff between security and privacy.

Table 3: Backdoor Variants and related malicious methods.

TYPE	DETAILS
Ransomware	Using backdoors to companies Information Systems.
SQLinjection	Deliberately input Weak code.
Trojans	Can be used to input backdoors using pretext of updating.
Default passwords	People often do not change default password.

Auditor's role in internal control

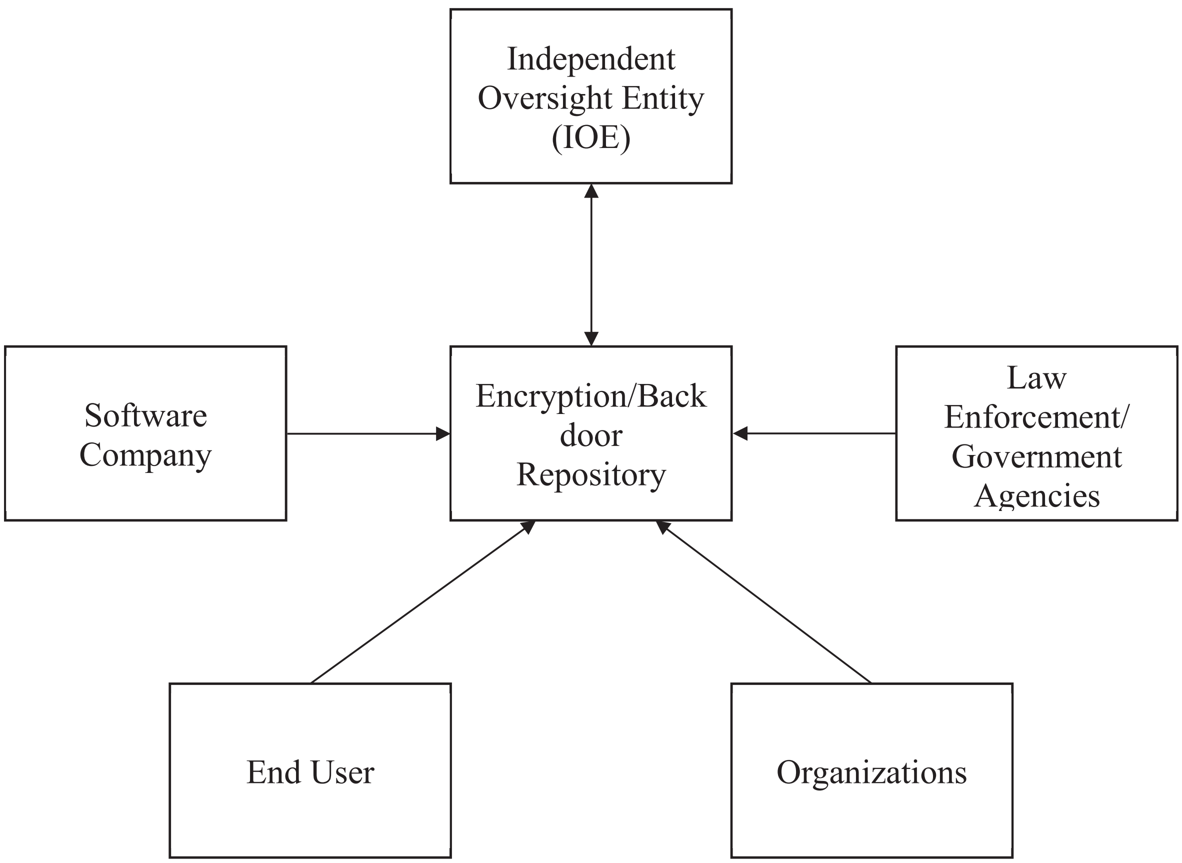
Table 4 provides a representative list of policies, legislation and frameworks addressing source code manipulations. Large public companies are required to have an integrated audit in accordance with SOX. Independent auditors are required to attest to the effectiveness of controls (Huang, Hung, Yen, Chang, & Jiang, 2011). According to SOX Section 302, management must annually certify to the completeness of accuracy of financial reporting by the CEO and the CFO. SOX Section 404 requires management's acknowledgement of its responsibility for establishing and maintaining adequate internal control over financial reporting and auditor attestation of the effectiveness of internal controls over financial reporting (Weidenmier & Ramamoorti, 2006). Information Technology is playing a pivotal role in corporate governance and SOX compliance. However, while IT is an enabler (Weidenmier & Ramamoorti, 2006) in managing risk, it itself becomes a source of risk. IT vulnerabilities open the organization to attacks. Auditors need to address this problem. To do so, a good understanding of the system is vital (Qian, Ward, & Blaskovich, 2012). One of the roles of the IT organization is to inform audit about challenges and ways to mitigate these challenges (Wilkin & Chenhall, 2010). Informing accounting professionals in IT related threats might help them in working with IT teams to identify, assess and mitigate e-business risks (Henderson, Lapke, and Garcia, 2016). To satisfy SOX internal control requirements, companies need to choose and implement suitable IT security frameworks (Wallace, Lin & Cefaratti, 2011).

Table 4: Representative list of policies addressing source code manipulations	
LEGISLATION	
SOX	Sarbanes-Oxley Act -2002
HIPAA	Health Insurance Portability and Accountability Act -1996
GLBA	Gramm-Leach- Bliley Act – 1999
FISMA	Federal Information Security Management Act -2002
FRAMEWORKS	
COBIT	Control Objectives for Information and Related Technologies is a created by ISACA for information technology (IT) management IT governance
COSO-ERM	Committee of Sponsoring Organizations of the Treadway Commission
ITIL	ITIL, formally an acronym for Information Technology Infrastructure Library, is a set of detailed practices for IT service management (ITSM) that focuses on aligning IT services with the needs of business
CMM	The Capability Maturity Model is a used to develop and improve an organization's software development process.
ISO	For IT organizations, ISO standards can reassure management and users that data and processes are safe -- and worth the investment.

PROPOSED FRAMEWORK

There are many cases of backdoors attacks due to technical issues such as hidden malicious code. However not all backdoor solutions are technical. While there are many technical approaches to addressing backdoors (Schuster, Rüster, & Holz, 2013 jul and nov; Pewny, et. al. 2015; Opaska 1986; Miyani, 2016), IT systems cannot be completely secured through technology alone. Behavioral and policy aspects need to be considered too (Sherif, Pitre, & Kamara, 2016). The ethical conundrum in the Apple situation – for example - is not a technical issue but an ethical decision. In many cases, backdoor attacks are not due to technical issues but due to lack of organization policies - for example requiring a change of the password from the initial default password to a more secure one. If not changed, the default password is vulnerable and can be exploited. Therefore, an all-encompassing framework covering all aspects is required. Unlike extant research, this framework brings together all stakeholders. The Backdoor problem encompasses legal, ethical, international, policy and technical issues. We propose a framework to address the problems associated with backdoors. Figure 1 provides the various stakeholders that contribute to the framework.

Figure 1: Proposed Backdoor Governance Framework



Independent Oversight entity

The Independent oversight entity (IOE) will be responsible for deciding the policies rules and regulation regarding use of backdoors. Often, there are ethical issues due to conflict between privacy (one greater good) and protection of life (another greater good). The proposed independent oversight entity will decide by navigating the delicate balance between privacy and protection of life. Such decisions – for example – will allow a backdoor into a system to help enforcement authorities reveal or tap into communications among terrorists. Such decisions can be arrived at by voting among member countries. Releasing a backdoor could be to protect the lives and open fraudulent hidden accounts in a software etc. All software developers will be required to submit a secret backdoor to their products. This backdoor can be accessed only when the IOE agrees to allow access to it.

The IOE can also create an alert system that will alert users about any dangers such as viruses. Just like the International Federation of Reporting Standards (IFRS) pursues standardization, an international code of conduct will be required from software products as well.

IOE Composition

The problem is international in nature with attacks often originating from another country. Therefore, the independent oversight entity needs to be an international entity. The IOE could be a combination of the Governments, Industry consortiums, and software consortia. An international consortium, perhaps a part of the United Nations, in collaboration with industry consortia would be able to oversee and maintain the repository of checksums (explained later) and backdoors for certified software and updates. Industry experts on technical, political, legal and ethical issues would serve as advisors.

Software development companies

The responsibility of member software companies is to upload all checksums (such as MD5 message-digest algorithm) and backdoors of their software - and subsequent updates - to the software repository. MD5 is a one-way cryptographic function is used for authentication. Using such cryptographic functions, organizations can promise privacy through encryption to all. However, in exacerbating circumstances, such as terrorism, fraud -for the greater good - the software can be accessed to protect others. Software owners should promise privacy with exceptions such as terrorism and uncovering of fraud.

Software Repository

The IOE can maintain the repository for checksum and backdoors.

Checksums**-** The software repository will keep all checksums MD5 of the certified software and software updates uploaded by the software companies. These will be publicly available for all to cross-check. This way organizations and individuals can check for any changes to the source code of the software by crosschecking it with the data in the repository. By providing such a facility, auditors can be more confident of the security of their software.
Checksum techniques can verify the integrity of information (Cohen, 1987). We will use MD5 - a widely used checksum algorithm - as an illustration (Klaus, 1997). These have other broader implications too that are of interest in auditors - such as protecting against unintended corruption in files, accidental deletion and modification (Cohen 1992). For example: If you input "Dallas Cowboys," the MD5 created is **5a40b7cccc2a230895881416edda4622**. You cannot get Dallas Cowboys from this value. It there is any change to the string a new value will be created. https://www.md5hashgenerator.com/. This concept applies to entire software code also. Even if there is a small change in the code, the MD5 created will be different. Therefore, software with 200 million lines of code can still be cross-checked for changes. In case of open source software, software development organizations can create a MD5 hash to allow users to check for potential changes to their code. Millions of lines of code cannot be manually checked against

changes. Therefore, an MD5 would allow users to check for changes in their code by comparing their code with the MD5 in the public domain. This can be done for every version and patch as well. For example, Harn et al, (1992) and Okamoto & Masumoto (1990) recommend a signature system to avoid being exploited by unlicensed distributors of the program. Other hash techniques include SHA1, SHA256, SHA384, SHA512, MACTripleDES and RIPEMD160 (Banday, 2011).

Backdoors**-** The software repository will also keep software backdoors uploaded by the software companies. These will not be available in the public domain. Software development companies can continually change backdoors for each version or update. Each of these backdoors will be required to be deposited to the repository. Access to such backdoors can be accessed only through many levels of checks. These checks could be available through litigation, subpoenas, and court orders etc. We discuss this in the next section.

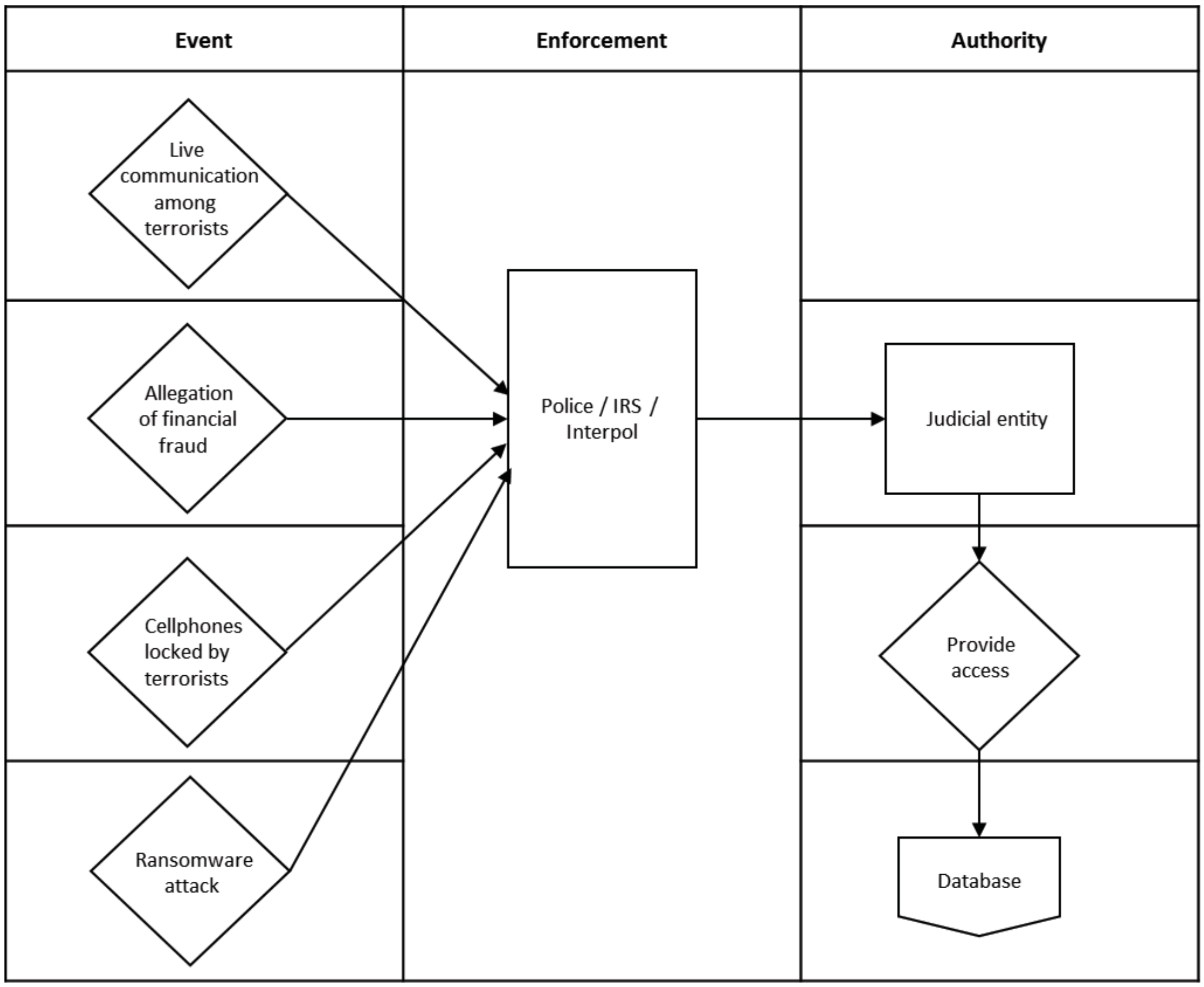
Enforcement Authorities

Enforcement authorities can ask for warrants to obtain the backdoor for individual software entities by petitioning the IOE. By considering the ethical issues and the decision between privacy and safety, IOE may or may not give access to the backdoor. Depending on the issue, the IOE could range from a local judge to international judiciary committee. The flow diagram is shown in Figure 2.

Terrorists can exploit Backdoors (Morris, 2017). Furthermore, encryption technologies protect criminal and terrorist organizations too (Valeri & Knights, 2000; Lee et al, 2016). Encrypted software, systems, communications etc., could be dangerous in the hands of wrong people. Encryption technology allows terrorists to communicate without being monitored by government agencies (Keene, 2011). The risks of having no access to such systems- in other words going dark - are real and endangers nations and therefore lawful access is needed (Corn 2015, 2017). The security and privacy that a software promises, becomes a liability when used by terrorists. In such rare cases, a decision needs to be made by the IOE to figure whether protection of lives overrides privacy.

There are other applications of this framework too. In the accounting context, sometimes enforcement agencies need access to accounting details in accounting software. Once a legal case is reviewed, enforcement authority would be able to get a warrant to get into the software through the backdoor from the repository.

Figure 2: Event Flow Diagram



Individual End-Users

Individual users of accounting software –such as freelance accounting services providers- have critical information about customers. Such users are particularly vulnerable, as they are usually unaware of vulnerabilities and are not protected by organization level controls. This could lead to loss of private customer information, destruction of data and loss of customer trust. To avoid being victims of backdoors attacks, individual users can crosscheck software and their updates from the encryption repository (figure 1). Much like the presidential alert system, end users can register for alerts and repositories can provide alerts to threats.

Organizational End-Users

Policy

For organizational end users, the stakes are higher as the loss of data could be costly due to lawsuits, breach of privacy, and loss of trust by clients. As systems become more robust against attack, only human errors and fallibility remains. Many organizations recognize that their employees, who are often considered the weakest link in information security, can also be great assets in the effort to reduce risk related to information security (Bulgurcu, Cavusoglu, & Benbasat, 2010). Therefore, there can be a policy level enforcement by the organization specifically addressing backdoor attacks.
For organizations, the framework (figure 1) will relieve them of the cost of monitoring every software application. Auditors need to be aware of the compliance requirements too. These include The Sarbanes-Oxley Act of 2002 (SOX), the Health Insurance Portability and Accountability Act of 1996 (HIPAA), and The Payment Card Industry Data Security Standard (PCI DSS). The above-mentioned frameworks can help the auditing departments in ensuring proper internal controls in IT. They can combine these frameworks with established IT control frameworks such as COBIT and COSO.

“User-developed” software and BYOD

Many organizations also use internally developed “User-developed” software that can be prone to malicious intent (Lewis et al., 2009). The problem gets compounded with combinations of insourcing, outsourcing, cloud computing, virtualization, and mobile computing among other environmental challenges (Debreceny, 2013). Auditors need to be cognizant of even end user vulnerabilities as ‘bring your own device’ (BYOD) policies are becoming pervasive (Utter & Rea, 2015). Viruses can find their way through employee devices into the organization systems.

Conclusion and Future Directions

Although IT mitigates risk, it in turn adds risk (Wilkin and Chenhall, 2010). One such risk involves backdoors. Backdoors attacks have cost organizations millions of dollars. The damage can be in the form of breaches of privacy, loss of data, loss of brand value, destroyed data, the loss of customer confidence and even loss of lives.

This attempts to cover all aspects of the backdoor ecosystem by developing policy level framework. Using such a framework could help solve many problems such as misuse of available backdoors, ransomware attacks, financial fraud, viruses, industrial and international espionage, and terrorism. The framework encompasses international, country, company, and individual levels. This framework will help individual users, organizations and countries protect against backdoor vulnerabilities.

Internet of Things (IOT) devices are increasing becoming pervasive in homes as well as in businesses. These too are vulnerable to backdoor attacks (Maras, 2015). In future IOT devices can be included in the above-mentioned framework.

Recommended Policies for individual users

- Use the proposed framework
- Update administrative passwords as soon as possible
- Do not install or update software from non-validated sites
- Periodically cross check changes to software and its updates

Recommended Policies for auditors

- Use the proposed framework
- Update administrative passwords as soon as possible
- Restrict employees from installing or updating software
- Periodically cross check changes to software and its updates
- Forbid vendors from direct access to install or change code without administrative monitoring
- Use software - such as Assembla and Bitbucket - to monitor software
- Manage BYOD policies
- Use only software from genuine certified vendors
- Monitor and maintain software audit logs
- Accept code with digital signatures only

Limitations of Framework

One challenge is to include employee developed software into the framework. This is particularly important as software developers are also employees. Employees often build small programs such as in Excel, Access and Visual Basic. Employees with malicious intent may have a hidden backdoor to the application they develop. Additionally, these are small programs that are at times, amateurishly built and do not have proper security checks (Cale, 1994). The risk magnifies if the employee developer has malicious intentions.

References

Abu-Musa, A.A. (2004) “Investigating the security controls of CAIS in an emerging economy: An empirical study on Egyptian banking industry”, Managerial Auditing Journal, vol. 19(2), 272-302

Abu-Musa, A.A. (2006) “Exploring perceived threats of CAIS in developing countries: the case of Saudi Arabia”, Managerial Auditing Journal, vol. 21 (4), 387-407

Allen, R. D., Hermanson, D. R., Kozloski, T. M., & Ramsay, R. J. (2006). Auditor risk assessment: Insights from the academic literature. Accounting Horizons, 20(2), 157-177.

Apostolou, B., Apostolou, N., & Schaupp, L. C. (2018). Assessing and Responding to Cyber Risk: The Energy Industry as Example. Journal of Forensic & Investigative Accounting, 10(1).

Banday, M. T. (2011). Easing PAIN with digital signatures. *International Journal of Computer Applications*, *975*, 8887.

Bansah, E. A. (2018). The threats of using computerized accounting information systems in the banking industry. Journal of Accounting and Management Information Systems, 17(3), 440-461.

Borthick, A. F., & Bowen, P. L. (2008). Auditing System Development: Constructing the Meaning of “Systematic and Rational” in the Context of Legacy Code Migration for Vendor Incentives. Journal of Information Systems, 22(1), 47-62.

Brandas, C., Stirbu, D., & Didraga, O. (2013). Integrated approach model of risk, control and auditing of accounting information systems. Informatica Economica, 17(4), 87.

Brynjolfsson, E., & Saunders, A. (2010). Wired for Innovation: How Information Technology is Reshaping Economy,–Cambridge (Massachusetts).

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness. MIS quarterly, 34(3), 523-548.

Cale Jr, E. G. (1994). Quality issues for end-user developed software. Journal of Systems Management, 45(1), 36.

Cereola, S. J., & Cereola, R. J. (2011). Breach of data at TJX: An instructional case used to study COSO and COBIT, with a focus on computer controls, data security, and privacy legislation. Issues in Accounting Education, 26(3), 521-545.

Cohen, F. (1987). A cryptographic checksum for integrity protection. Computers & Security, 6(6), 505-510.

Cohen, F. B. (1992). Defense-in-depth against computer viruses. Computers & Security, 11(6), 563-579.

Corn, G.S. 2015. Averting the Inherent Dangers of “Going Dark”: Why Congress Must Require a Locked Front Door to Encrypted Data. Washington and Lee Law Review 72(3): 1433.

Corn, G. S. (2017). Encryptions, Asymmetric Warfare, and the Need for Lawful Access. Wm. & Mary Bill Rts. J., 26, 337.

Dai, S., Wei, T., Zhang, C., Wang, T., Ding, Y., Liang, Z., & Zou, W. (2012, May). A framework to eliminate backdoors from response-computable authentication. In Security and Privacy (SP), 2012 IEEE Symposium on (pp. 3-17). IEEE.

De, R. 2009. Economic Impact of Free and Open Source Software: A Study in India. Interop, Mumbai, October: 7-9.

Deepa, G., & Thilagam, P. S. (2016). Securing web applications from injection and logic vulnerabilities: Approaches and challenges. Information and Software Technology, 74, 160-180.

Debreceny, R. S. (2013). Research on IT governance, risk, and value: Challenges and opportunities. Journal of Information Systems, 27(1), 129-135.

Haislip, J. Z., Peters, G. F., & Richardson, V. J. (2016). The effect of auditor IT expertise on internal controls. International Journal of Accounting Information Systems, 20, 1-15.

Hanini, E. (2012) “The risks of using computerized accounting information systems in the Jordanian banks; their reasons and ways of prevention”, European Journal of Business and Management, vol. 4(20): 53-63

Harn, L., Lin, H. Y., & Yang, S. (1992). A software authentication system for information integrity. Computers & Security, 11(8), 747-752.

Henderson, D., Lapke, M., & Garcia, C. (2016). SQL Injection: A Demonstration and Implications for Accounting Students. AIS Educator Journal, 11(1), 1-8.

Hood, K.L. & Yang, J.W. (1998) “Impact of banking information systems security on banking in China: The case of large state-owned banks in Shenzhen Economic Special zone – an introduction”, Journal of Global Information Management, vol. 6 (3): 5-15

Huang, S. M., Hung, W. H., Yen, D. C., Chang, I. C., & Jiang, D. (2011). Building the evaluation model of the IT general control for CPAs under enterprise risk management. Decision Support Systems, 50(4), 692-701.

Keene, S. D. (2011). Terrorism and the internet: a double-edged sword. Journal of Money Laundering Control, 14(4), 359-370.

Kim, W., Jeong, O. R., Kim, C., & So, J. (2011). The dark side of the Internet: Attacks, costs and responses. Information systems, 36(3), 675-705.

Kim, G., Richardson, V. J., & Watson, M. W. (2018). IT does matter: The folly of ignoring IT material weaknesses. Accounting Horrizons

Klaus, C. (1997). An Introduction to Backdoors. EDPACS: The EDP Audit, Control, and Security Newsletter, 25(6), 1-9.

Landwehr, C., A. Bull, J. Mcdermott, and W. Choi. 1994. A Taxonomy of Computer-Program Security Flaws. ACM Computing Surveys 26(3): 211-54.

Lee, S., Beyer, J. L., Karmali, N., Isberto, C., & Oh, H. (2016). JSIS Cybersecurity Report: How Should the Tech Industry Address Terrorist Use of Its Products? .

Lewis Jr, S. X., Magruder, J. S., Burks, E. J., & Smolinski, C. (2009). Business Risks: When IS Fails to Detect Circumvention Activities. Journal of International Technology and Information Management, 18(1), 7.

Li, C., Peters, G. F., Richardson, V. J., & Watson, M. W. (2012). The consequences of information technology control weaknesses on management information systems: The case of Sarbanes-Oxley internal control reports. Mis Quarterly, 179-203.

Maras, M. H. (2015). Internet of Things: security and privacy implications. International Data Privacy Law, 5(2), 99.

Masli, A., Richardson, V. J., Sanchez, J. M., & Smith, R. E. (2011). The business value of IT: A synthesis and framework of archival research. Journal of Information Systems, 25(2), 81-116.

Mihalache, A. S. (2011). Risk Analysis of Accounting Information System Infrastructure.

Miyani, D. (2016). BinPro: A Tool for Binary Backdoor Accountability in Code Audits (Doctoral dissertation, University of Toronto).

Morris, S. A. (2017). The Misuse of Encryption and the Risks Posed to National Security (Doctoral dissertation, Utica College).

Muhrtala, T. O., & Ogundeji, M. (2013). Computerized Accounting Information Systems and Perceived Security Threats in Developing Economies: The Nigerian Case. Universal Journal of Accounting and Finance, 1(1), 9-18.

Murthy, U. S. (2016). Researching at the Intersection of Accounting and Information Technology: A Call for Action. Journal of Information Systems, 30(2), 159-167.

O'Donnell, J. B., & Rechtman, Y. (2005). Navigating the standards for information technology controls. The CPA Journal, 75(7), 64.

Okamoto, E., & Masumoto, H. (1990). ID-based authentication system for computer virus detection. Electronics Letters, 26(15), 1169-1170.

Opaska, Walter P. "Closing the VAX Default Password “Backdoor”." EDPACS: The EDP Audit, Control, and Security Newsletter 14, no. 3 (1986): 6-9.

Perlroth, N., Scott, M., & Frenkel, S. (2017). Cyberattack Hits Ukraine Then Spreads Internationally. The New York Times.

Pewny, J., B. Garmany, R. Gawlik, C. Rossow, and T. Holz. 2015. Cross-Architecture Bug Search in Binary Executables. IEEE Symposium on Security and Privacy.

Polo, J. & Oima, D. (2013) “Effects of computerized accounting systems on audit risk management in public enterprises: A case of Kisumu county, Kenya, International Journal of Education and Research, vol. 1(5): 1-10

Qian, J., Ward, K., & Blaskovich, J. (2012). Integrating IT frameworks into the AIS course. AIS Educator Journal, 7(1), 1-26.

Rajeshwaran, N., & Gunawardana, K. (2009). An Empirical Investigation of the Security Controls of Computerized Accounting Information Systems (CAIS) in the Selected Listed Companies in Sri Lanka.

Riad, N. I. (2009). Security of accounting information systems: A cross-sector study of UK companies. Cardiff University (United Kingdom).

Schuster, F., & Holz, T. (2013, November). Towards reducing the attack surface of software backdoors. In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (pp. 851-862). ACM.

Schuster, F., Rüster, S., & Holz, T. (2013, July). Preventing backdoors in server applications with a separated software architecture. In International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (pp. 197-206). Springer, Berlin, Heidelberg.

Sherif, K., Pitre, R., & Kamara, M. (2016). Why do information system controls fail to prevent unethical behavior?. VINE Journal of Information and Knowledge Management Systems, 46(2), 251-266.

Singh, A. N., & Gupta, M. P. (2017). Information Security Management Practices: Case Studies from India. Global Business Review, 0972150917721836.

Singleton, T., & Singleton, A. (2004). Cyberterrorism: Are you at risk?. Journal of Corporate Accounting & Finance, 15(5), 3-12.

Srinivasan, C. R. (2017). Hobby hackers to billion-dollar industry: the evolution of ransomware. Computer Fraud & Security, 2017(11), 7-9.

Tambe, P., & Hitt, L. M. (2012). The productivity of information technology investments: New evidence from IT labor data. Information Systems Research, 23(3-part-1), 599-617.

Utter, C. J., & Rea, A. (2015). The" Bring your own device" conundrum for organizations and investigators: An examination of the policy and legal concerns in light of investigatory challenges. Journal of Digital Forensics, Security and Law, 10(2), 4.

Valeri, L., & Knights, M. (2000). Affecting trust: terrorism, internet and offensive information warfare. Terrorism and Political Violence, 12(1), 15-36.

Wang, T., Lu, K., Lu, L., Chung, S. P., & Lee, W. (2013, August). Jekyll on iOS: When Benign Apps Become Evil. In USENIX Security Symposium (Vol. 78).
Wallace, L., Lin, H., & Cefaratti, M. A. (2011). Information security and sarbanes-oxley compliance: An exploratory study. Journal of Information Systems, 25(1), 185-211.

Weidenmier, M. L., & Ramamoorti, S. (2006). Research opportunities in information technology and internal auditing. Journal of Information Systems, 20(1), 205-219.

Wilkin, C. L., & Chenhall, R. H. (2010). A review of IT governance: A taxonomy to inform accounting information systems. Journal of Information Systems, 24(2), 107-146.

Zweilf, A. (2009) “The Nature of the threats of the electronic accounting information systems: An application study of the Jordanian insurance companies Arabian”, Journal of Accounting, 46-64.

DECISION SCIENCES INSTITUTE

Dockless Shared-Use Mobility: A Pattern Mining Approach

Ali Rahim Taleqani
North Dakota State University
Email: ali.rahimtaleqani@ndsu.edu

Ameneh Forouzandeh Shahraki
North Dakota State University
Email: ameneh.forouzandehsh@ndsu.edu

Peyman Tavallali
Independent Researcher
Email: tavallali@gmail.com

Jill Hough
Upper Great Plains Transportation Institute
Email: jill.hough@ndsu.edu

ABSTRACT

Human mobility pattern mining is invaluable to transportation decision-making because it enables public and private entities to reduce operational cost, improve system performance, and promote shared mobility modes of transportation. We offer two data mining approaches to a dockless shared mobility service in Washington, D.C. The authors used data collected from two different operators. We outline the geohashing technique combined with data mining methods to provide a promising alternative to the existing model. We intend to provide a better understanding of commuter movements at different levels of spatiotemporal resolution.

KEYWORDS: Big data, Business intelligence, Shared-use mobility, Pattern mining

BACKGROUND

City officials promote active transportation to enhance the sustainability of transportation systems. Micromobility options provide affordable and convenient alternatives within these systems. Bicycle and scooter-sharing schemes are two typical modes of this service. Unlike the traditional dock-based model, free-floating, non-station based, or “dockless” sharing systems offer the flexibility to pick up and leave the shared vehicle (bike or scooter) in any accessible locations across the operational area. This system is meant to supplement the dock-based model. However, it has the potential to compete with the traditional dock-based model as it has lower upfront operating costs, more flexibility, and may accommodate people with disabilities (Rahim Taleqani et al., 2019). The traditional dock-based service fails to cover the gap between the docking stations and points of arrival/departure. Dockless systems allow the geolocation of every single bike or scooter. The geo-localized data collected by operators contains valuable mobility information about users’ commuting patterns such as trip length, time, visited locations, etc.

In the transportation field, human mobility is much more complex and stochastic as it entails interpersonal variability, including gender, age, income, daily activity patterns, destinations, routes, and modes of transportation. Data collected on dockless mobility systems could reveal many implicit and explicit patterns. In general, mobility pattern mining could be at the micro or macro level and provide insight into individual or group mobility. Mobility pattern mining and trajectory prediction are the focus of many studies in transportation, mobile computing, infectious disease modeling, and other areas (Lu et al., 2012; Shahraki et al., 2017). Spatial and temporal data of individual entities are generally captured in the form of global positioning system (GPS) data points (Liu et al., 2015).

Abundant research is focused on modeling human mobility and forecasting movement trajectories with various methods, including Markov chain (MC), Artificial Neural Network, Bayesian network, and finite automaton. In general, most existing literature falls into two broad categories: 1) spatiotemporal prediction, 2) next location prediction.

Individual mobility can be captured with devices such as GPS receivers or smartphones. The latitude, longitude, and timestamp are collected at various frequency rates, depending upon the device's capability. The result is a sequence of locations with a different time for each location.

Various approaches and techniques have been applied to spatiotemporal data points. Spatiotemporal modeling aims to discover prominent daily temporal habits as well as predict future individual activities (Scellato et al., 2011; Zheng and Ni, 2012). Spatial movement modeling facilitates next location prediction (Bayir et al., 2010; Lv et al., 2014), residence time prediction (Baumann et al., 2013; Chon et al., 2012), life pattern mining (Bayir et al., 2010; Ye et al., 2009), and identification of departures from a mobility routine (Mcinerney et al., 2013; Smith et al., 2014). Lu and his team implemented a series of Markov Chains to analyze the travel patterns of 500,000 individuals during the post-civil war period in Cote d'Ivoire. Individual trajectories are collected by tracing mobile phone call data. They improved the predictability to as high as 88% by considering the uncertainties of movements using entropy, the frequencies, and the temporal correlations of individual trajectories. They concluded that human mobility behavior is far from random and that historical behavior greatly influences individuals' movements (Lu et al., 2013).

In wireless communication, the quality-of-service guarantee can be achieved if enough resources are reserved before a new cell arrives. Hence, the system has to predict the exact number of the next-crossing cell. Shiang-Chun et al. proposed a neural network prediction scheme to predict the location of the mobile host. Their work efficiently improves the accuracy of mobile hosts trajectory prediction, increases the success probability of resource reservation, and enhances bandwidth utilization(Liou and Lu, 2003). The University of Colorado's adaptive control of home environments project developed a smart house that observed the lifestyle of the inhabitants and learned to infer and accommodate their needs such as adjusting the thermostat, turning on a particular configuration of lights. Motion detectors track occupants, and a neural network approach is used to predict the next room the person will enter and the activities in which that person will be engaged (Mozer, 1998). In response to a similar problem, Akoush and Sameh presented a novel hybrid Bayesian neural network model to address the uncertainty for predicting locations on cellular networks. Unlike the conventional approach, which is based on dynamic tracking, their method is based on topology-independent user tracking. This method reduced the cost of tracking significantly (Akoush and Sameh, 2007). They concluded that Bayesian learning for neural networks predicts location better than standard neural network techniques because it uses a well-founded probability model to represent uncertainty about the relationships being learned. Instead of using Markov chains, Bayesian networks, and neural networks, Petzold et al. proposed using global and local state context prediction to address human mobility with context information. Their work originated in branch prediction techniques which are used in CPU design but was modified to handle context prediction. They modeled two-level two-state predictors with local and global first-level histories. The model is used by simulating the predictors with behavior patterns of people walking through a building as a workload. Findings show excellent performance, but exhibit differences in training and retraining speed and in their ability to learn intricate patterns (Petzold and Bagci, 2003).

One of the challenges with the previous methods is the absence of a reliable data collection at consistent spatial and temporal resolution. The contribution of this research is a geohash-based data mining approach to detect underlying patterns in daily commutes. The ability to look at different resolutions could address a range of problems from urban design related issues such as blocked entry/exit building access, old city infrastructure to inter/intra city transportation management issues, including the lack of comprehensive enforcement on reckless commuters.

The remainder of the article is structured as follows. In the next section, we introduce the details of pattern mining and clustering methods. We then discuss the findings from each method and a possible way for visualization. We conclude the article with a synopsis of our research contributions and a discussion of opportunities for future research.

METHODOLOGY

Pattern Mining

The scooter/bicycle movement could be seen as a **shopping basket analysis** problem. Techniques used in this field seek to understand influences on the purchase of other items. The findings help retailers enhance sales and revenue by cross-selling or up-selling. With some modification, the same approach could be applied to a shared-use mobility system. Because of privacy, users are anonymous, so we focus on the movement of the vehicle and the importance of prior location on traveling to the next location. We describe the transaction as a sequence of bike/scooter movements during the day of the week for a specific user. The series could be either a sequence of check-out or check-in locations. Each transaction is composed of a set of items. An Item is defined as a region of interest during a time frame (period). The goal is to understand a group of spatiotemporal items that occur together frequently. For example, as described in Table 1, “A_6”, “B_6” as region “A” and “B” during 6 a.m.

Apriori algorithm is a classical algorithm for frequent mining of itemsets and relevant association rules. There are three metrics to measure association:

Support indicates the popularity of an itemset, as measured by the proportion of transactions in which an itemset appears. As shown in Table 1, the support of {C_5} is 4 out of 8 or 50% or the support of {C_5, B_6, C_7} is 2 out of 8, or 25%.

Confidence indicates the probability of item Y (consequent) as a destination when item X (antecedent) is the prior location as an origin, expressed as {X -> Y}. It is measured by the proportion of transactions with item X, in which item Y also appears. In Table 1, the confidence of {C_5 -> B_6} is 3 out of 4, or 75%.

$$\textbf{Confidence}\ \{\mathbf{C_5} \rightarrow \mathbf{B_6}\} = \frac{\textbf{Support}\ \{\mathbf{C_5}, \mathbf{B_6}\}}{\textbf{Support}\{\mathbf{C_5}\}} \quad (1)$$

Lift indicates the likelihood of item Y as a destination when item X is the origin while controlling for how popular item Y is. Similarly, the lift of {C_5 -> B_6} is 1, implying that there is no association between items. A lift value higher than 1 implies that item Y is likely to be selected if item X is the origin, while a value less than 1 means that item Y is unlikely to be chosen if item X is the origin.

$$\boldsymbol{Lift}\{\boldsymbol{C_5} \rightarrow \boldsymbol{B_6}\} = \frac{\boldsymbol{Support}\{\boldsymbol{C_5}, \boldsymbol{B_6}\}}{\boldsymbol{Support}\{\boldsymbol{C_5}\} \times \boldsymbol{Support}\{\boldsymbol{B_6}\}} \quad (2)$$

Table 1: Transactions vs. Items

Transactions	Items
1	A_6, B_6, C_7, H_12
2	A_6, B_6, C_7
3	A_6, B_6
4	A_6, D_11
5	C_5, B_6, C_7, H_12
6	C_5, B_6, C_7
7	C_5, B_6
8	C_5, D_11

Clustering

Feature Extraction with Restricted Boltzmann Machine (RBM)

The Boltzmann machine is energy-based unsupervised nonlinear feature learner. It consists of hidden and visible neurons (units) in a fully-connected bipartite graph. Previous research using RBM as feature extraction tools was conducted by Hinton (Tieleman, 2008). Because the RBM uses the hidden units to model the correlation of the data set, the expected value of the hidden units of a learned RBM can be viewed as the extracted features of visible units. Empirical research shows that using these extracted features in place of the data itself yields significant improvements in several tasks (Cai et al., 2012; Zhang et al., 2018). One of the significant advantages of RBM is that the extracted features of a learned RBM can be used as input for training another RBM to capture much higher-order abstract features, and this procedure could be repeated as many times as needed to build a deeper network.

T-distributed stochastic neighbor embedding

t-distributed stochastic neighbor embedding (tSNE) (LJPvd and Hinton, 2008) is a method to reduce the dimensionality of the dataset while preserving neighborhood structure to visualize the structure of the data before clustering. Unlike Principal Component Analysis (PCA), it is a non-linear, probabilistic method with a random walk on neighborhoods graphs to find structures within the data. It converts the similarities of data points to a joint probability distribution to minimize the divergence between two distributions: distribution of the original data points (high dimensional dataset) and distribution of dimensionality reduced dataset.

Density-based spatial clustering of applications with noise (DBSCAN)

DBSCAN locates clusters as areas of high density by counting the number of points in the region of a specified radius. Data points with a density above a specified threshold are constructed as clusters. The DBSCAN algorithm is able to discover clusters with an arbitrary shape such as linear, concave, oval, etc. Additionally, unlike some clustering algorithms, it does not require the predetermination of the number of clusters. DBSCAN has been proven in its ability to process massive databases (Ester et al., n.d.; Taleqani et al., 2019). DBSCAN basically needs two parameters: 1) the minimum distance (eps) between two points to be considered as a neighbor, 2) the minimum number of points (min_samples) to form a dense region. Together, these two parameters define density. Higher min_samples or lower eps indicate a higher density necessary to build a cluster. The parameters estimation is a critical DBSCAN. A too-small value of eps results in a large number of outliers, while very large numbers render few large clusters. The eps should be chosen based on the k-nearest neighbors' distance graph. The value of k corresponds to min_samples. Min_sample, as a general rule of thumb, can be derived from the number of dimensions (minPoints ≥ D + 1). The larger the min_samples, the better for datasets with noise.

EXPERIMENTS AND RESULTS

Case Study

For this research, the District of Columbia (D.C.) is selected because it is one of the rare cases in which the dockless operators publicly share real-time data. As of September 2017, the District Department of Transportation (DDOT) conducted a dockless demonstration to evaluate the pros and cons of this type of shared mobility. From September 2017 through June 2018, seven dockless operators participated in a D.C., pilot demonstration. Mobike, Lime, Spin, and Ofo operated traditional pedal bicycles. JUMP provided motorized or electric-assist bikes. Skip Bird, and Lime operated electric scooters. During the period, more than 625,000 dockless trips were made by approximately 233,700 unique users using the seven companies.

Data for six active dockless programs are available through the public application programming interface (API). These programs include Jump (Neon Red), Lime (Green), Spin (Orange), Skip (Yellow), Bird (Black), Lyft (Pink). However, only two of them, Lime and Lyft, provide the correct availability status of the vehicles in a given time. To be consistent with our analysis, we used only the data from these two companies. To use the data for the other four companies, we could approximate the speed of the vehicle by computing movement and timer intervals. However, this is not a very reliable method because we are not able to identify if a vehicle is moving because a user is using it or because the bikes are on a truck being moved around the city for rebalancing.

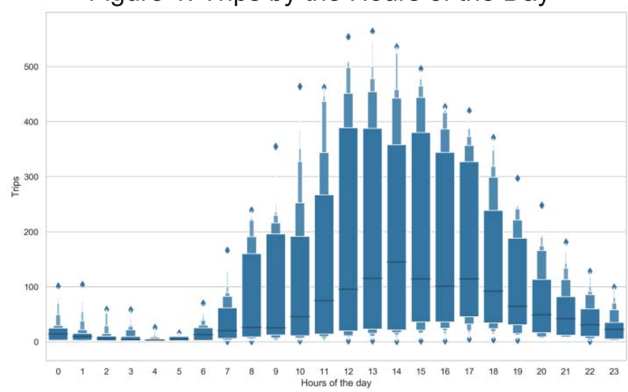
Data were live-streamed through their APIs for a period of one month, from March 15 to April 15. Data were retrieved every minute for a higher degree of time resolution and an in-depth understanding of mobility. The dataset before preprocessing has almost 33 million data points. These data include the bike/scooter ID, geographic coordinates (latitude and longitude), and timestamp of the available bicycle (at one-minute temporal resolution). Further steps were necessary to convert the bike/scooter availability data into trips. Since users are anonymous, we focus only on the movement of an individual vehicle and report the number of trips made by a vehicle.

Trips summary

Temporal patterns

The mean duration of a trip for both operators was approximately 15 minutes while the mean for a Lime trip is 16 minutes, and Lyft trips take 7.5 minutes on average. The tendency towards longer trips by Lime users is significant and may be partially due to the design of the vehicle or the popularity due to the availability of Lime vehicles across the region. However, the longer time does not mean longer traveled distance as a user might travel a short distance but put the scooters/bikes on hold for further use. The temporal popularities of the two shared mobility services are shown in Figure 1. The trip start times are aggregated to the nearest hour of a day. We see an expected daily pattern, with the majority of trips taking place during daylight hours, especially the midday period, for both operators. However, fewer rides in early morning imply that this mode is not popular for the early morning commute.

Figure 1: Trips by the Hours of the Day



In Figure 2, there is no visible monthly cyclic pattern, which might be because of inclement weather during March and April and corresponding interruptions in services. For example, on March 21, unlike Sundays, the low number of trips was due to high precipitation, and on April 14 because high winds were reported. In terms of days of the week, we expect a higher number of trips during the weekend, as shown in Figure 3.

Figure 2: Trips by the Days of the Month

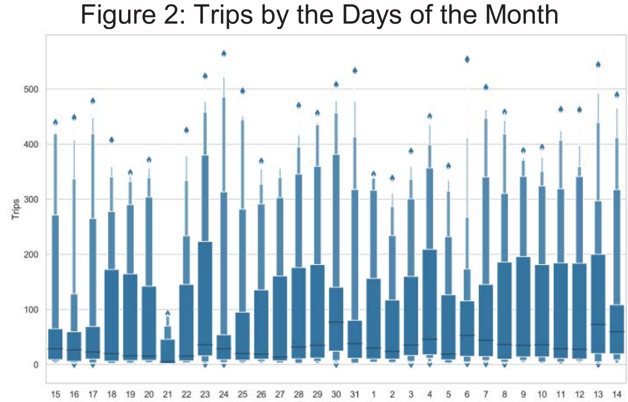
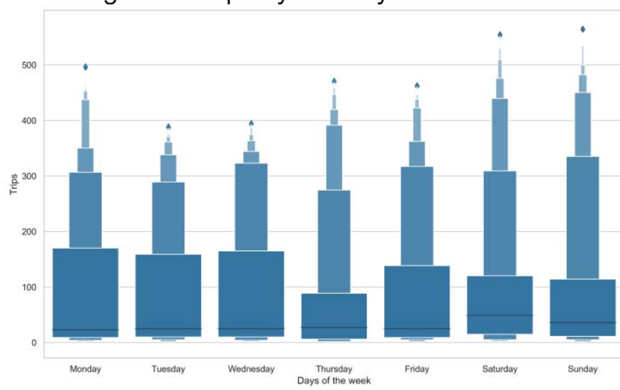


Figure 3: Trips by the Days of the Week



Spatial Patterns

The spatial pattern of mobility is critical to both city planners and operators. For this purpose, we used the concept of geohash tessellation to partition Washington, D.C. into equal size polygons. The advantage of using geohash over transportation analysis zone parcel data is that we can use different resolution (granularity level) to discover any movement pattern as small as 1 foot by 1 foot. Figures 4 and 5 illustrate the net flow of trips by region based on geohash 6. The dark red rectangles indicate a higher number of outgoing trips, and grays rectangles imply a higher number of incoming trips. Trips made through Lyft had sporadic patterns and extended to Arlington and Reagan National Airport. In comparison, trips made through Lime mostly happened in D.C. areas with three imaginary layers of outgoing, incoming, then another layer of outgoing trips from the center to surroundings.

Figure 4: Net Flow of Trips for Lime

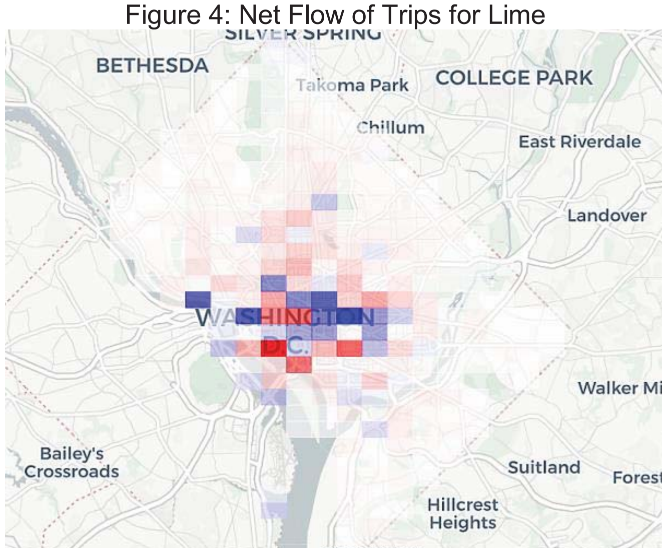
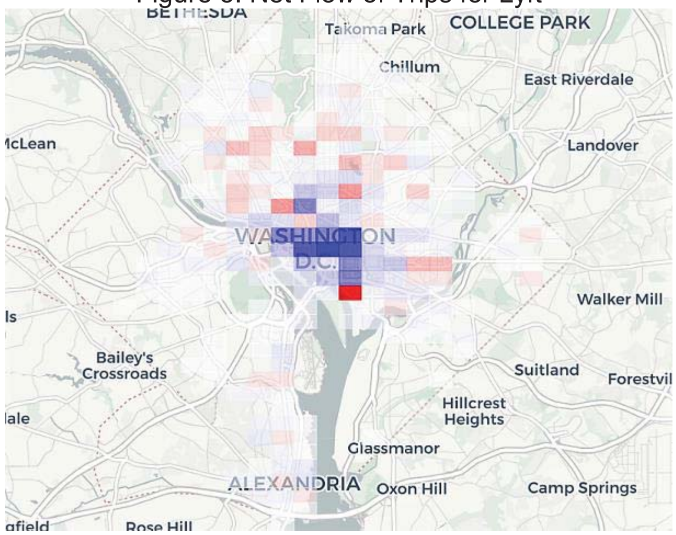
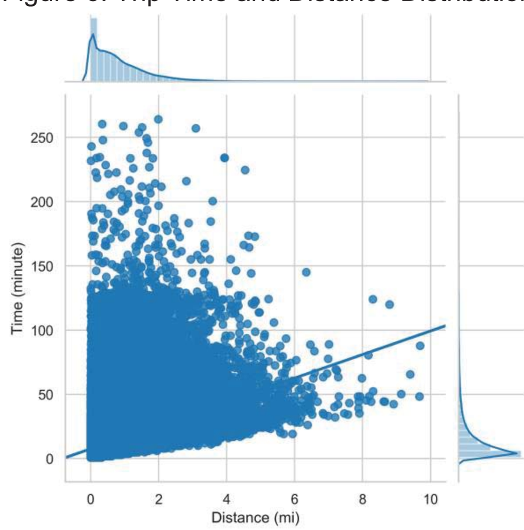


Figure 5: Net Flow of Trips for Lyft



Dockless time offers a relaxed version of shared mobility transportation options, so we expect a very stochastic spatial pattern from the traditional model. Figure 6 shows the trip distance and time distribution. A Manhattan (city block) distance method is used to calculate the distance between the origin and destination of a trip. As expected, there is a linear relationship between time and distance as shown by a tendency toward longer distances in less time. This could be because of the distance approximation method. Both time and distance are heavily right-skewed (with mean of 0.82 miles and 15 minutes) with some outliers. In terms of speed, the maximum is less than 20 mph, which is the limit with a mean of 4.5 mph.

Figure 6: Trip Time and Distance Distribution



From an operational perspective, there were only three bicycles available, which might either indicate the status as a pilot project or the unpopularity of bikes under dockless programs (Table 2). Surprisingly, Lime has fewer scooters available but made the higher turnover per scooter with almost 651 unique scooters active each day on average, with each making 5.8 trips. A low market share might be due to the fact that Lyft had recently debuted the service in Washington, D.C., and had achieved less popularity. The other key factor is that when a user has some credit on his account, they are highly unlikely to switch to other operators. Consequently, we are expecting Lyft operation statistics to improve over time.

Table 2: Trip Statistics

Operator	Lime		Lyft
Type	Bicycle	Scooter	Scooter
Total number of unique vehicles that made a trip	3	2064	2596
The average number of vehicles that made a trip per day	1.2	650.94	279.29
The average number of trips per vehicle per day	1.8	5.8	1.6
The average time in service per vehicle per day (minutes)	35.3	93.89	12.3
The average distance traveled per vehicle per day (miles)	1.12	0.98	0.04

One critical aspect of dockless programs is their higher degree of sprawl across a region. This may result in vehicles being parked close to personally preferred locations rather than in areas with high public demand. Therefore, we expect some locations to be considered as sink points when the vehicles are unused for most of the day while some other locations are considered hot spots with little idle time. The average idle times are 1.36 and 1.15 minutes for Lime and Lyft, respectively. However, some geohases experience the larger value of idle time. Figures 7 and 8 demonstrate the average daily percentage of the time that a bike is inactive (unused). For instance, regions close to the National Mall, the Tidal Basin, and the Arboretum (as tourist attractions) experience lesser idle time per vehicle as vehicle turnover is fairly high. On the hand, some regions like metro stations, parks and residential have larger idle time per vehicle as vehicles might stay longer until next check-out.

Figure 7: Idle Time of Lime Vehicles in Minutes

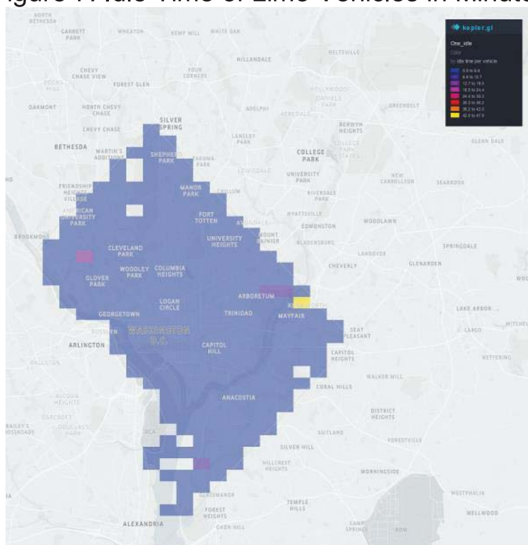
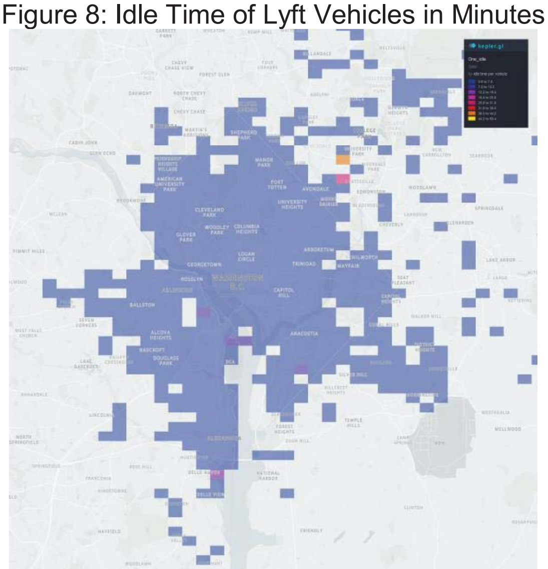


Figure 8: Idle Time of Lyft Vehicles in Minutes



Visitors are likely to travel to tourist attractions and then explore the area on foot to visit points of interest which are close by. Because visitors travel in groups, finding the right number of vehicles to travel together may not be possible, which leads to walking on foot instead of riding bikes/scooters. The other surprising fact is that very close to the sink regions, there is a region with less idle time (high turnover) implying that the area works as a virtual hub or meeting point for users to find a scooter when they are getting out of the sink regions. As explained earlier, the difference between Lime and Lyft could be that Lime was well-established before Lyft and has a higher market share. These findings will help operators rebalance their vehicles across the city to improve asset utilization and increase revenue. This information also could be used for customized pricing to offer some incentive to users to travel to regions with higher idle times.

Data-driven pattern mining

In this section, because of limited computational power available, we only used a subset of data which included one week from March 25 to March 31.

Sequence Mining

For the first step, we used a sequence mining technique to find a set of frequent geohashes (regions) visited at a different time of the day. We understand that, because there are no fixed docking stations under the dockless system, movement patterns are highly sprawled and stochastic. This leads to a lower value of support and confidence of the Apriori algorithm, respectively. Unlike in dock-based models, vehicle movements happen between fixed locations, so the support and confidence values are higher. The findings can be categorized into three groups: 1) groups of visiting points making one-leg trips (Figure 9), 2) groups of three-visiting points making two-leg trips (Figure 10), and 3) a group of four visiting points making three-leg trips. We focus only on the three-leg trips, as shown in Figure 11. The starting leg is illustrated in a light color and last leg in dark blue. There are two significant series of visiting points (numbered locations in Figure 11). One series started with the first leg from Franklin square to a point near the White House. The second leg takes longer, ending at a point near Capitol Hill. The last leg ended at Capitol Hill and the Library of Congress. The second important series of locations started at Madame Tussaud's museum area close to many shops, restaurants, museums, and galleries and ended at the Smithsonian Museum on the National Mall. The second leg extended to the Washington Monument and finally finished on the side of the National Mall next to the Lincoln Memorial. These two patterns imply that most commuters are probably tourists exploring the D.C attractions, Access to bike/scooter data offers an opportunity for a company such as Lime to rebalance its inventory while considering the location of its assets as well those of its competitors

Given that dockless scooters/bikes can be left virtually anywhere, we can assume that vehicles are most often parked at the most convenient locations for their users. This information could be used daily to rebalance vehicles across the city to match demand. This information can also be used by competitors like Capital Bikeshare (with fixed docking stations) to assess the optimality of current docking station locations.

Figure 9: Frequent 1-leg Trips

Figure 10: Frequent 2-leg Trips

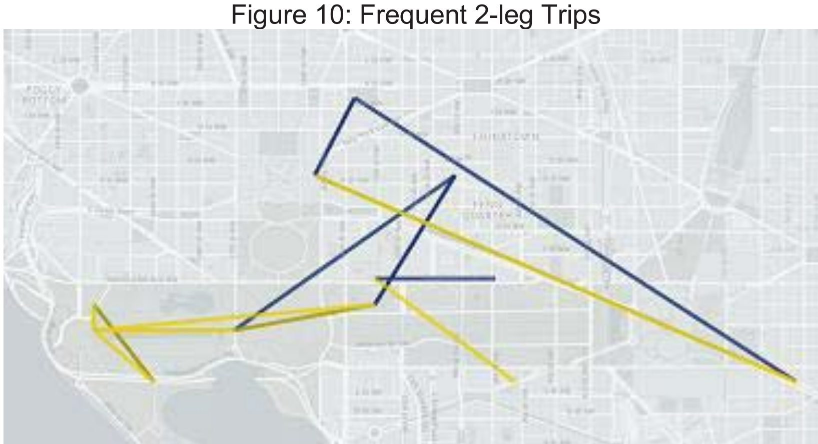
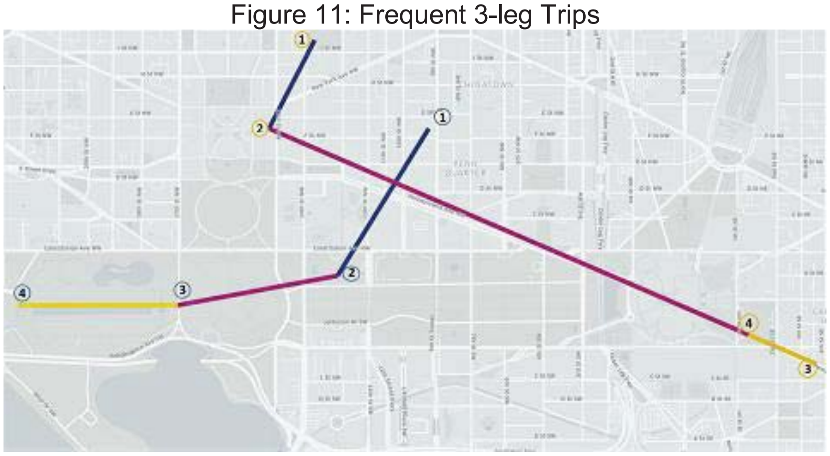


Figure 11: Frequent 3-leg Trips



Clustering Similar Groups

In the second step, we tried to find groups of movement with some kind of similarities. Hence we do the clustering based on DBSCAN as described earlier in the methodology section. Feature extraction is done by the RBM method to extract 20 latent features in the dataset. For determining eps and minimum points for the DBSCAN algorithm, we calculated the k-nearest neighbor distance. The value of k is set to 22 and corresponds to minimum points. Next, these k-distances are plotted in ascending order. The aim is to determine the “knee,” which corresponds to the optimal eps parameter. A knee corresponds to a threshold where a sharp change occurs along the k-distance curve.

We employed the t-SNE algorithm to reduce the dimensionality of the dataset and visualize the structure in 2D space. The resulting plot is shown in Figure 12. The plot shows the presence of dense regions with arbitrary shapes surrounded with very tight empty space.

We also checked the silhouette score as an extra validation method for evaluating the consistency within the clusters of data. The coefficient ranges from -1 to +1. Values near 0 indicate that the object is on the border of two clusters. The silhouette score is based on the assumption of convex clusters, so their expected values in a non-convex case are generally smaller. However, we want to maximize silhouette scores ($\boldsymbol{silhouette} \rightarrow \mathbf{1}$) and at the same time avoid a large number of clusters.

We instantiated the DBSCAN model with eps = 0.003 and minimum samples = 22. As DBSCAN marks noisy points with the label -1, the resulting plot is shown in Figure 13. The total number of clustering, excluding noise, is 48. Moreover, eps = 0.003 is associated with the highest silhouette scores equal to 0.93, which indicates that the overall clustering is reasonably correct. We can also observe two fundamental results: the noise points (marked with a cross) are not isolated in the t-SNE representation, and some clusters are partially split. This is not a failure of the algorithm, but a direct consequence of the dimensionality reduction. In the original space, all noise points are not density-connected to any point in the sample but can be overlapping or close to some points in the t-SNE visualization. A large number of clusters is expected for the context of dockless shared mobility as movement patterns are individually shaped. For the purpose of this paper, we analyzed the five most significant clusters.

Figure 12: Visualization of Reduced Dimensions by t-SNE

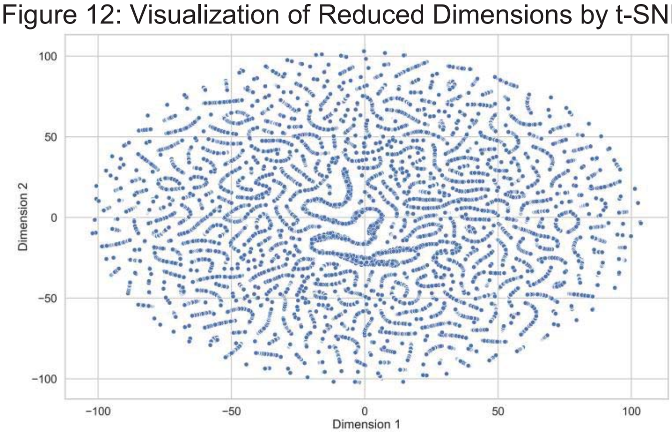
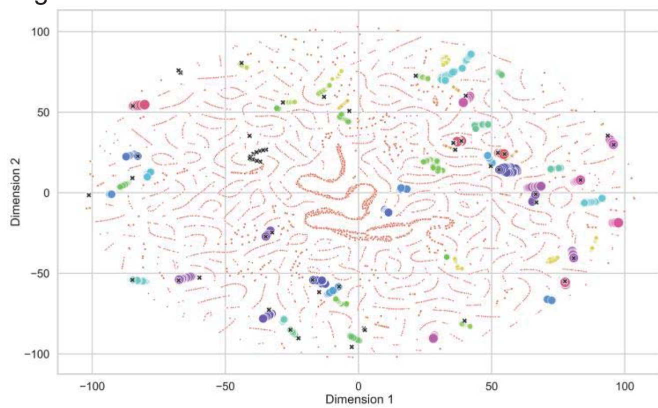


Figure 13: Visualization of Clusters in Reduced Format



As shown in Table 3, there are two majors patterns in the clusters labeled 0 and 1. Both have an unbalanced number of check-ins and check-outs. The former has a large number of check-ins, implying geohashes as destinations, and the latter has a higher number of check-outs, representing origins. Clusters 7, 14, and 24 have an equal amount of check-ins and check-outs but include only data points between 4 -10 p.m. Each of the three represents a separate day: 7, 14, and 24 for Sunday, Monday, Tuesday respectively. The number of features is limited in this study. For future studies, income level and the number and the type of business in each geohash could give more details of subgroup mobility patterns.

Table 3: Summary of Clustering

Cluster Number	Number of data points
-1 (noise)	353
0	21971
1	7056
7	228
14	208
24	193

CONCLUSIONS

We have argued that pattern mining is essential to transportation professionals and decision-makers for improving customer satisfaction and asset utilization. We used the example of the Washington, D.C., dockless shared mobility data to show that knowledge from publicly available data could benefit public and private sectors to realize the strengths and weaknesses of shared mobility schemes and improve operational performance.

Contribution to Research

First, we extend the literature on mobility pattern mining by applying a new approach to understand hidden user commuting patterns. Both sequence mining and clustering demonstrate promising results for extracting meaningful and practical implications from users' mobility data.

Second, unlike the common approach of using transportation analysis zones (TAZ), zip codes, or parcel data, we tried the geohash concept to find patterns in different mobility resolution. This approach could benefit, not only shared-use mobility operators but also urban designers to understand parking patterns relative to building entry/exit doors, sidewalks, streets, etc.

Contribution to Practice

Our findings offer several insights for practitioners. This work investigates the spatial and temporal dimensions of dockless scooter/bike-share services in Washington, D.C. First, we suggest that time-dependent geohash would be the best option to find frequent users' mobility patterns. The results may lead to a cost-effective approach for rebalancing bike/scooters by operators. Different pricing strategies based on the findings would also help operators shift some of the rebalancing work to users. This would dramatically decrease labor for collecting vehicles scattered all-around blocks or at intersections.

Second, we emphasize the importance of users with similar mobility patterns by clustering them into various groups. Operators could apply different types of incentives to frequent users, or record number of accidents for each cluster and find out the groups with a high risk of injuries. City planners could impose traffic rules for different clusters to manage the interactions between shared mobility users and larger vehicles such as cars and transit buses.

Implications for future research

Though much of this analysis is limited to only two active dockless programs, the findings suggest that there are great opportunities to understand the mobility patterns of commuters by using machine learning techniques. Both sequence mining and clustering show promise as approaches for better analysis of hidden profiles in commuters' mobility. However, several improvements could yield better findings:

1. More cities with different spatial patterns could make mobility data available to researchers, transportation engineers, and city planners. The findings of these analyses may influence maintenance planning, infrastructure budgeting, and new development opportunities.
2. Trip surveys of commuters could help realize the true characteristics of clusters we found which may be beneficial to companies for determining pricing schemes, demand analysis, etc.
3. Given the sheer volume of data generated every second from these shared mobility systems, neural networks, especially recurrent neural network, could be used to build a model for predicting the next locations visited.
4. The results presented in this paper are not comprehensive enough because, as explained earlier, access to the spatial and temporal resolution of dockless shared mobility data in the United States is limited. Analyzing more data over a more extended period will provide additional insight. Future work will examine the impact of socioeconomic factors, compare these patterns to supplement or complement modes to dockless shared mobility alternatives.

ACKNOWLEDGMENT

We express our appreciation to Kenny Durell and the Coord Company for allowing us to download the initial dataset through their API for free.

REFERENCES

Akoush, S., Sameh, A., 2007. Movement prediction using bayesian learning for neural networks, in: 2007 Second International Conference on Systems and Networks Communications (ICSNC 2007). IEEE, p. 6.

Baumann, P., Kleiminger, W., Santini, S., 2013. How long are you staying?: predicting residence time from human mobility traces, in: Proceedings of the 19th Annual International Conference on Mobile Computing & Networking. Citeseer, pp. 231–234.

Bayir, M.A., Demirbas, M., Eagle, N., 2010. Mobility profiler: A framework for discovering mobility profiles of cell phone users. Pervasive Mob. Comput. 6, 435–454.

Cai, X., Hu, S., Lin, X., 2012. Feature extraction using restricted Boltzmann machine for stock price prediction. CSAE 2012 - Proceedings, 2012 IEEE Int. Conf. Comput. Sci. Autom. Eng. 3, 80–83. https://doi.org/10.1109/CSAE.2012.6272913

Chon, Y., Shin, H., Talipov, E., Cha, H., 2012. Evaluating mobility models for temporal prediction with high-granularity mobility data, in: 2012 IEEE International Conference on Pervasive Computing and Communications. IEEE, pp. 206–212.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., n.d. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.

Liou, S.-C., Lu, H.-C., 2003. Applied neural network for location prediction and resources reservation scheme in wireless networks, in: International Conference on Communication Technology Proceedings, 2003. ICCT 2003. IEEE, pp. 958–961.

Liu, J., Zhang, S., Kato, N., Ujikawa, H., Suzuki, K., 2015. Device-to-device communications for enhancing quality of experience in software defined multi-tier LTE-A networks. IEEE Netw. 29, 46–52.

LJPvd, M., Hinton, G.E., 2008. Visualizing high-dimensional data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605.

Lu, X., Bengtsson, L., Holme, P., 2012. Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. 109, 11576–11581.

Lu, X., Wetter, E., Bharti, N., Tatem, A.J., Bengtsson, L., 2013. Approaching the limit of predictability in human mobility. Sci. Rep. 3, 1–9. https://doi.org/10.1038/srep02923

Lv, Q., Mei, Z., Qiao, Y., Zhong, Y., Lei, Z., 2014. Hidden Markov model based user mobility analysis in LTE network, in: 2014 International Symposium on Wireless Personal Multimedia Communications (WPMC). IEEE, pp. 379–384.

Mcinerney, J., Stein, S., Rogers, A., Jennings, N.R., 2013. Breaking the habit: Measuring and predicting departures from routine in individual human mobility. Pervasive Mob. Comput. 9, 808–822.

Mozer, M.C., 1998. The neural network house: An environment that adapts to its inhabitants. Proc. AAAI Spring Symp. Intell. Environ. 110–114. https://doi.org/SS-98-02/SS98-02-017

Petzold, J., Bagci, F., 2003. Global and local state context prediction. Artif. Intell. ….

Rahim Taleqani, A., Hough, J., Nygard, K.E., 2019. Public Opinion on Dockless Bike Sharing: A Machine Learning Approach. Transp. Res. Rec. https://doi.org/10.1177/0361198119838982

Scellato, S., Musolesi, M., Mascolo, C., Latora, V., Campbell, A.T., 2011. Nextplace: a spatio-temporal prediction framework for pervasive systems, in: International Conference on Pervasive Computing. Springer, pp. 152–169.

Shahraki, F.F., Yazdanpanah, A.P., Regentova, E.E., Muthukumar, V., 2017. A trajectory based method of automatic counting of cyclist in traffic video data. Int. J. Artif. Intell. Tools 26, 1750015.

Smith, G., Wieser, R., Goulding, J., Barrack, D., 2014. A refined limit on the predictability of human mobility, in: 2014 IEEE International Conference on Pervasive Computing and Communications (PerCom). IEEE, pp. 88–94.
Taleqani, A.R., Bridgelall, R., Hough, J., Nygard, K., 2019. Data Driven Analytics of Road Quality. Proc. 34th Int. Confer 58, 454–463.
Tieleman, T., 2008. Training restricted Boltzmann machines using approximations to the likelihood gradient, in: Proceedings of the 25th International Conference on Machine Learning. ACM, pp. 1064–1071.
Ye, Y., Zheng, Y., Chen, Y., Feng, J., Xie, X., 2009. Mining individual life pattern based on location history, in: 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware. IEEE, pp. 1–10.
Zhang, N., Ding, S., Zhang, J., Xue, Y., 2018. An overview on Restricted Boltzmann Machines. Neurocomputing 275, 1186–1199. https://doi.org/10.1016/j.neucom.2017.09.065
Zheng, J., Ni, L.M., 2012. An unsupervised framework for sensing individual and cluster behavior patterns from human mobile data, in: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM, pp. 153–162.

Decision Sciences Institute

Software Languages for Analytics in Research

Ceyhun Ozgur
Valparaiso University
Email: ceyhun.ozgur@valpo.edu

Pervaiz Alam
Kent State University
Email: palam@kent.edu

David Booth
Kent State University
Email: dvdbooth8@gmail.com

Abstract

This paper discusses how R Python SPSS and SAS can be used to predict bank failures and the benefit that it would have to company's futures. By using stable forecasting, investors can predict the future bank failures. By using these software, organizations can govern the life cycles of their inquiries and be strategic in their in their future financial stability in the bank. We provide an example of how companies can calculate internal rate of return (IRR) and use this to determine future bank failures. This paper also discusses how each software can be used in different manners to determine future financial stability.

Keywords: Statistical Software Languages, Banking Research R, Python, Excel, SPSS, SAS

Introduction

R, Excel, SAS, SPSS and Python are five software programs designed to run statistical analyses and output graphics, can be used for banking research. R can run on any operating system, is open-source, and reflects many of the changing field preferences. It is also highly standardized. SAS is a paid software system that provides high performance analytics for banking research. Organizations can identify, investigate, and govern the life cycle of their inquiries. When it comes to data science one of the most common points of debate is R vs SAS vs Python vs Excel vs SPSS. It is a well-known fact that R, Python, Excel, SPSS and SAS are the most important five languages to be learned for data analysis.

When it comes to data science one of the most common points of debate is R vs SAS vs SPSS vs. Python. We should also include Excel, because, It does all of the functions that the other four programing languages provide. It is a well-known fact that R Python SPSS and SAS are the most important 4 programming languages to be learned for data analysis by data scientists. However, we can also utilize Excel, to perform most of the operations, especially the simple ones. Excel can be purchased when you purchase or download from Microsoft windows website.

In banking research, there are many software packages are being used including R, Python, SPSS.and SAS. Among these packages Python and R open source, free of charge software programming languages. However SAS and SPSS are paid for by licensing companies or universities. (C. Ozgur et.al 2017)

SAS Programming Language

It is a programming language where the input language common spreadsheets as input mechanisms to generate output based on the results of statistical analysis in the form of tables and graphs as RTF PDF, HTML documents. It is an expensive language that is not affordable by most data scientists. Unless the organization helps individuals with financial aid, they will not be able to use SAS. However, both R and Python are free of charge, they could be accessed by anyone, anywhere. This is a biggest advantage of to all data scientists for using both R and Python free of charge.

Classroom and Online Training for Certifications *https://www.edupristine.com/*

SPSS Programming Language

It is also a programming language where the input language common spreadsheets as input mechanisms to generate output based on the results of statistical analysis but the difference from SAS is that the programming language for SPSS is easier to learn and more similar to Excel. However, just like SAS, it is not free of charge. However, we are able to solve large scale problems. *https://www.edupristine.com/*

Python Programming Language

It is another programming language free of charge like R. It is another open source programming language, free to access by everyone. Its code is easier to learn than R or SAS. Python is easier programming language to learn than both. *https://www.edupristine.com/*

R Programming Language

R It is an open source programming language, free to access and pen to all to perform data analysis tasks. It is supported by the R Foundation for basically Statistical Computing. The R language is widely used among data miners for developing statistical software and data analysis. The source code for the R software environment is written primarily in C, FORTRAN, and R. Language R is freely available under the GNU General Public License and is pre-compiled binary versions that are provided for various operating systems. While R has a command line interface, there are several graphical front-ends also available.

Classroom and Online Training for Certifications *https://www.edupristine.com/*

Excel

It also can be used to solve simple problems that other's programs can solve. We can also utilize Excel, to perform most of the operations, especially the simple ones. Excel can be purchased when you purchase or download from Microsoft windows website.(Microsoft, 2019)

R in Finance

The software grew in popularity in response to the increasing overall demands on big data analytics and the need for programs which could handle massive data files. This changed how computational finance was done; the availability of a new and versatile tool is apt to reframe at least a portion of the discussions around calculation and application. Rickert (2014) notes that there were at least 70 separate finance packages available in 2014.

R's flexibility as a software means that it is fully able to perform the necessary tasks for financial analytics. There have been several programs written to facilitate the usage of R in such a setting; one of the most popular is jrvFinance (Trajanov, 2017). This package, alongside the standard tools of R and the bundle package Rmetrics (Trajanov, 2017), allows users to perform the necessary standard financial computations in the software. The largest advantage of this program is that it is set up to mimic the Excel format, allowing familiarity with the system. As Varma (2016) notes in his overview of the jrvFinance package, there are several specific commands to ensure that the correct operations are being performed.

JrvFinance specifically allows for the Internal Rate of Return (IRR) operations within R. Once loading R's syntax, the next step would be to load the package in the current R session (as packages are not stored from session to session). The syntax would be written as (where cf is the cash flow):

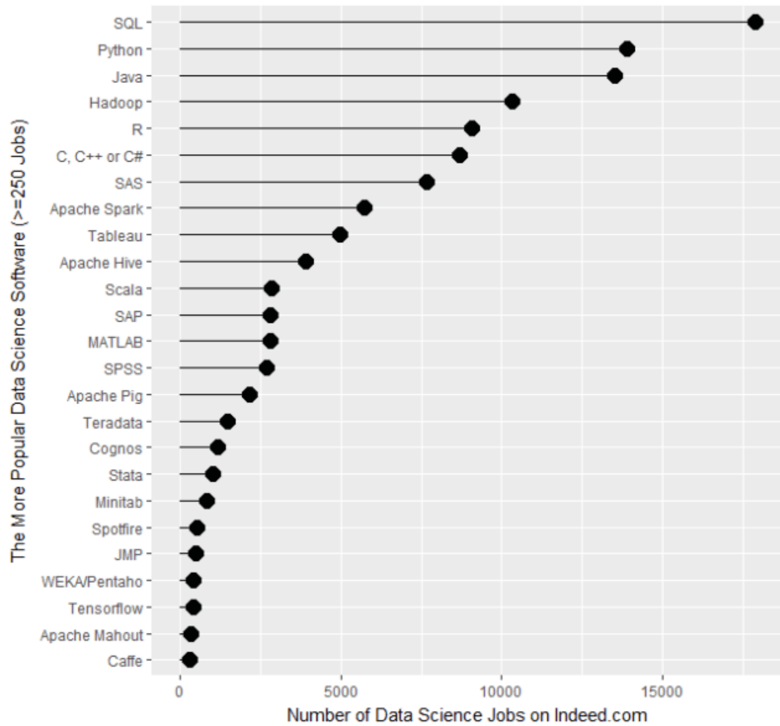
Internal Rate of Return: *irr(cf)*
(1)
Internal Rate of Return example: *irr(cf=c(-10, 3, 4, 5, 6))*
(2)

The example would compute an IRR of 0.2488813 for the cash flow specified in the function. The presented example is a fairly simple string of code to learn.

R's Growth and Comparison to Other Software

R's growth in the data analytics field has come to surpass SAS, SPSS, Stata, and MatLab, leading to an increase in job postings where applicants must know how to use R. The overall trend of R is observable in Figure 1, stated from Muenchen (2017).

Figure 1. NUMBER OF DATA SCIENCE JOBS POSTED ON A JOB SEARCH WEBSITE WITH OVER 250 JOB HITS



1,

SAS & R for banking and finance

Continuous innovation(https://www.sas.com en_ph/news/press-releases/2017/january/2016-financials.html)

Analysts named SAS a leader in predictive and advanced analytics, customer intelligence, data management, data integration and data quality. According to IDC, SAS holds a 31.6 percent share of the advanced and predictive analytics market.[2] SAS has also been recognized by industry analysts as a leader in fraud detection, risk and retail analytics. (Cary, 2017).

Maintaining this leadership is heavily dependent upon innovation. Year after year, SAS reinvests about twice the average of major technology firms into R&D – 26 percent in 2016. This unwavering commitment to innovation is behind the ground-breaking new SAS Viya technology – dubbed by analysts as changing the industry. SAS continues to introduce new innovation around this open and cloud-ready high-performance analytics and visualization platform – most recently with SAS Visual Investigator, which marries advanced analytics with dynamic and interactive visual workspaces. With it, organizations can identify, investigate and govern the entire life cycle of an investigation, search or inquiry. SAS plans to introduce even more to the SAS Viya family in the first quarter of 2017. (Cary, 2017).

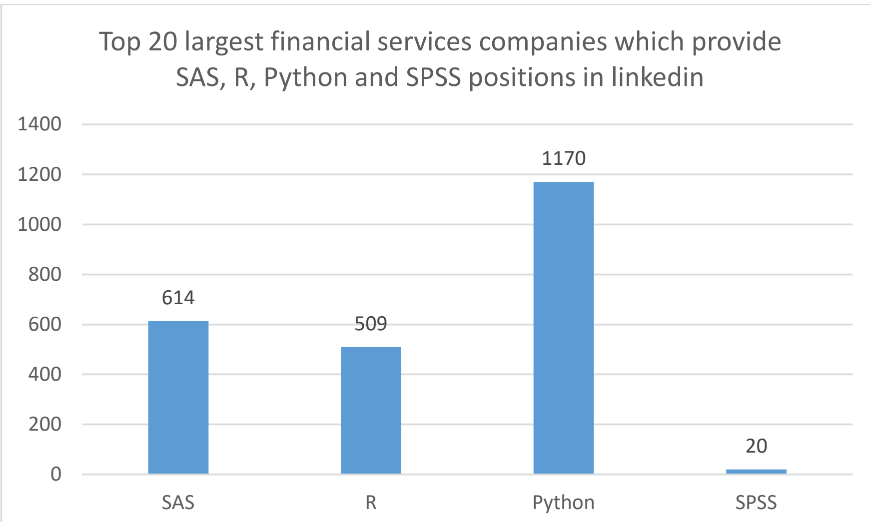
Looking further ahead in 2017, SAS will continue innovation in its core focus areas, including analytics, visualization, data management, customer intelligence, risk and fraud. Additionally, SAS Viya, artificial intelligence, cloud and IoT will be strong investment areas for SAS. (Cary, 2017).

2,

Table 1. SAS vs R vs Python vs SPSS

SAS	R	Python	SPSS
614	509	1170	20

Figure 2. Top 20 Largest Financial; Services Companies With SAS, Pythom, R & SPSS



Top 20 largest financial services companies which provide SAS, R, Python and SPSS positions in linkedin

(https://www.forbes.com/global2000/list/

https://en.wikipedia.org/wiki/List_of_largest_financial_services_companies_by_revenue)

Deposit insurance is a key element in modem banking, it guarantees the financial safety and stability of the depository financial institutions. If an insured depository institution fails to fulfill its obligations to its depositors, the insuring agency will step in to honor the principal and accrued interests. (Dar-YehHwang Cheng F.Lee K.Thomas Liaw)

Table 2. Word's Largest 20 Banking & Insurance Companies

Rank	Company	Industry	Headquarters
1	Berkshire Hathaway	Conglomerate	United States

2	AXA	Insurance	France
3	Allianz	Insurance	Germany
4	ICBC	Banking	China
5	Fannie Mae	Investment Services	United States
6	BNP Paribas	Banking	France
7	Generali Group	Insurance	Italy
8	China Construction Bank	Banking	China
9	Banco Santander	Banking	Spain
10	JP Morgan Chase	Banking	United States
11	Société Générale	Banking	France
12	HSBC	Banking	United Kingdom
13	Agricultural Bank of China	Banking	China
14	Bank of America	Banking	United States
15	Bank of China	Banking	China
16	Wells Fargo	Banking	United States
17	Citigroup	Banking	United States
18	Prudential	Insurance	United Kingdom
19	Munich Re	Insurance	Germany
20	Prudential Financial	Insurance	United States

Forecasting Bank Failures

There has been major concern throughout economists that there will be a recession by 2020 in the United States. Specifically, a survey from J.P. Morgan found that 75 percent of ultra-high net worth investors predict that the US economy will fall into a recession by that time (Turak, 2018). Although there is concern regarding ultra-net investors, there is not as much concern in the realm of the banking industry. For instance, the International Monetary Fund recently upped its U.S. growth forecast for 2018 to 2,9 percent. This increase was caused by Tax Cuts and Job Acts that enlarged the federal spending package. Although this has present benefits, in the long run it is likely that the banks will not benefit from the tax cuts. (Turak, 2018).

References

Classroom and Online Training for Certifications *https://www.edupristine.com/ Januay 2016*

Cary, N. C. (2017). SAS continues growth with 2016 revenue of US $3.2 billion. *SAS.* Retrieved from https://www.sas.com/en_ph/news/press-releases/2017/january/2016-financials.html

C. Ozgur, T. Colliau, G. Rogers, Z.Hughes & B. Myer-Tyson "MatLab vs. Python vs. R" Journal of Data Science Vol. 15 No. 3 pp. 355-372 July 2017

Dar-YehHwang Cheng F.Lee K.ThomasLiaw. International Review of Economics & Finance. Retrieved from https://www.sciencedirect.com/science/article/pii/S1059056097900411

Forbes. (2017). *The world's biggest public companies* [Data file]. Retrieved from https://www.forbes.com/global2000/list/

Microsoft for Excel, PowerPoint, Word, Outlook and Access. Received from: https://products.office.com/en-US/compare-all-microsoft-office-products?&icid=CNavSoftwareOffice&activetab=tab%3aprimaryr1

Muenchen, B. (2017, March 1). *Data Science Job Report 2017: R Passes SAS, But Python Leaves them Both Behind*. Retrieved from: https://www.r-bloggers.com/data-science-job-report-2017-r-passes-sas-but-python-leaves-them-both-behind/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+RBloggers+%28R+bloggers%29

Rickert, J. (2014, May 8). R and Finance. Retrieved from: http://blog.revolutionanalytics.com/2014/05/r-and-finance.html

Trajanov, D. (2017, February 27). R software for statistical computing: useful in finance? [Web log comment]. Retrieved from: https://www.researchgate.net/post/R_software_for_statistical_computing_useful_in_finance#view=58bd86c196b7e4950c77810e

Turak, N. (2018, April 19). 75% of the ultra-rich forecast a US recession in the next two years, survey finds. Retrieved from https://www.cnbc.com/2018/04/19/75-percent-of-the-ultra-rich-forecast-a-us-recession-in-the-next-two-years-survey-finds.html

Varma, J. (2016 August 29). Package ‘jrvFinance’. Retrieved from: https://mran.microsoft.com/web/packages/jrvFinance/jrvFinance.pdf

Wikipedia. (2017). *List of largest financial services companies by revenue* [Data file]. Retrieved from https://en.wikipedia.org/wiki/List_of_largest_financial_services_companies_by_revenue

DECISION SCIENCES INSTITUTE

An Obvious Basis for MTABS Analysis of 2x2 Cross-Tabulations that Doesn't Work

John R. Dickinson
University of Windsor
Email: MExperiences@bell.net

ABSTRACT

The chi-square test of independence attending cross-tabulations is among the most popular types of analysis in the social sciences. Recently, a system for measuring the reliability of that test, MTABS, has been developed. In construing MTABS a seemingly obvious basis for its progression–it is a form of sensitivity analysis–is directly apparent. That basis, however, is invalid. This study explains and demonstrates that invalidity.

KEYWORDS: MTABS, Cross-tabulation, Chi-square test of independence

INTRODUCTION

Cross-tabulation, often accompanied by the chi-square test of independence, is a mainstay of marketing research and research in the social sciences generally.

- "A cross-tabulation...is easily the most widely used data analysis technique in marketing research..." (Iacobucci & Churchill 2010, p. 362)
- "Perhaps the most common type of bivariate analysis in practice is the cross tabulation of two nominal variables." (Kinnear & Taylor 1991, p. 577)
- "Perhaps the most common of all uses of the χ^2 text is the test of whether an observed breakdown of frequencies in a 2 X 2 contingency table could have occurred under H_0. (Siegel 1956, p. 107)
- Of 20 types of statistical analysis published in major advertising journals from 1980 through 2010, cross-tabulation was the fourth most frequently used. (Yoo *et al.* 2015, p. 554)

MTABS (Moves To and Beyond Significance, Dickinson 2016, 2017) analysis is a complement to the usual chi-square hypothesis test of independence in a cross-tabulation. As a form of sensitivity analysis, its purpose is to provide a measure of the reliability of that statistical test.

From the original cross-tabulation, MTABS successively moves single observations between pairs of cells comprising the cross-tabulation. Moving TO significance, MTABS identifies the two cells where moving a single observation from one to the other achieves the greatest decrease in the p-value. With that move made, MTABS repeats the process for the revised cross-tabulation. The progression is the same for moving the p-value BEYOND significance, only the pair of cells identified is the one that achieves the greatest increase in the p-value.

From the p-value of the original cross-tabulation MTABS determines the number of observations that must be moved to achieve a p-value less than, say, 0.05 and the number of observations that must be moved to achieve a p-value greater than or equal to, say, 0.10.

α=0.05 and α=0.10 are traditional benchmarks for statistical hypothesis tests, though any levels specified by the researcher may be used.

SUM ELEMENT BASE

With that objective in mind, one might be inclined to think that the MTABS "to significance" progression would be made by shifting a single observation from the cell making the lowest contribution to the chi-square statistic to the cell making the highest contribution; it's obvious. The chi-square calculation is a sum and this "obvious" progression can be phrased as moving a single observation from the cell having the smallest sum element to the cell having the highest sum element. This "should" provide the maximum increase in the chi-square statistic value and, correspondingly, provide the maximum decrease in the p-value, i.e., the "to significance" direction.

But this "sum element" basis doesn't work. The reason is that with the movement of a single observation not only does the "from" cell count decrease and the "to" cell count increase, but also the row and column counts of both of those respective cells change. With this, the expected values in all cells in those rows and column change, not just the expected values in the "from" and "to" cells. So it is not necessarily the case that the "sum element" basis results in a regular (i.e., monotonic) increase (decrease) in the chi-square statistic value.

The purpose of this study is to recognize and demonstrate this invalidity for MTABS.

ACCOMPLISHING THE INTENDED MTABS PURPOSE

The sum element basis *may* work monotonically for both decreasing the p-value TO significance and increasing the p-value BEYOND significance. Such is the case, for example, for Bush & Leigh (1984, Table 3, p. 37, Prime Time National~Local x Traditional~Cable Networks) with original p-value 0.11001. And also the case for Murphy, Kangun, & Locander (1978, Exhibit 3, p. 64, Soft Drinks) having original p-value 0.00056. Substantially different original p-values, yet the sum elements basis succeeds in changing the p-value as intended.

SAME DATA BASE, OPPOSITE INTENDED p-VALUE CHANGE

As recorded above, the sum elements basis for moving observations *may* monotonically decrease the p-value TO significance and also *may* monotonically increase the p-value to BEYOND significance. This is in keeping with the intended purpose of MTABS. Such is the case with Lynn (1981, Table 2, p. 15, Marital Status) that relates newspaper readership to marital status.

Relating newspaper readership to education using that same data base (Table 4, p. 15, Education), however, in both TO and BEYOND the sum elements basis monotonically changes the p-value in the direction *opposite* that intended. In the latter cross-tabulation, the sum elements basis completely opposes the purpose of MTABS.

NONMONOTONICITY

The sum element basis toward significance moves observations from the cell with the lowest chi-square sum element to the cell with the highest element. Or toward insignificance from the cell having the highest sum element to the cell having the lowest. In neither direction is the effect necessarily monotonic. For two published cross-tabulations, Table 1 provides examples of this nonmonotonicity.

In Table 1 the direction of change in the p-value after each of 10 observation moves is denoted by ▼ for a decrease and by ▲ for an increase. As may be seen, in the "to"intended direction of change in the p-value, the p-value increases with the eighth and tenth observation moves. In the "beyond" intended direction, the first two observation moves using the sum element basis decrease, not increase, the p-value.

Table 1: Nonmonotonic Changes in p-Value

Intended Direction	Source	Original p-Value	Move 1-2-3-4-5-6-7-8-9-10
TO	Golan & Wanta (2001), Table 3, p. 254, Cut Taxes x Poll~Media	0.25858	▼-▼-▼-▼-▼-▼-▼-▲-▼-▲
BEYOND	Golan & Wanta (2001), Table 3, p. 254, Trust x Poll~Positive	0.02360	▼-▼-▲-▲-▲-▲-▲-▲-▲-▲

EFFICIENCY

Using the sum elements basis for Golan & Wanta (2001, Table 3, p. 254, Best Chance to Win x Poll~Positive Coverage) finds monotonically decreasing p-values for the TO significance direction and also monotonically increasing p-values for the BEYOND significance direction. For this particular application, the sum elements basis validly serves the purpose of MTABS.

In the TO significance direction, the sum elements basis takes 10 observation moves to achieve a p-value of 0.00504, still not less than 0.001. The MTABS approach, however, achieves a p-value of 0.00079 with the movement of just five observations. In the BEYOND significance direction, the sum elements basis takes 4 observation moves to achieve a p-value of 0.10061, just above one traditional benchmark for significance of $\alpha=0.10$. The MTABS approach exceeds that benchmark in only two moves ($p=0.115507$).

Beyond its demonstrable invalidity, it may be said, then, that the MTABS basis is more efficient than the sum elements basis.

THE STANDARDIZED RESIDUAL

The sum elements that are the focus of the present study do serve a purpose. Reynolds (1984, p. 22) suggests examining, "...each component of the chi square statistic...[to] indicate which categories of variables are most closely related." More prominent in the literature, though, is the standardized residual which is the square root of a given sum element. "The contribution of each cell to the Pearson chi-square statistic is printed in the table of **standardized residuals**. Chi-square is the sum of the squares of these." (SPSS 1999, p. 77, bold theirs) With this, the two statistics are perfectly nonlinearly correlated. The standardized residuals, then, suffer the same invalidities as the sum elements.

MTABS SUMMARY

As stated at the outset of this study, MTABS was developed to assess the reliability of the chi-square test of independence in cross-tabulations. The prescribed resultant of an MTABS analysis is to report:

- the chi-square p-value for the original table
- MTABS moves TO α=0.05 % (n)
- MTABS moves BEYOND α=0.10 % (n)

where n is the number of moved observations to decrease the p-value to less than 0.05 or to increase the p-value to greater than or equal to 0.10 and % is the number of such moved observations as a percent of the total number of observations. For the cross-tabulations cited herein, MTABS results are reported in Table 2.

Table 2: MTABS Summary for Cited Cross-Tabulations

Cross-Tabulation	**Original p-Value**	**To p-Value <0.05**	**To p-Value ≥0.10**
Bush & Leigh (1984, Table 3, p. 37, Prime Time National~Local x Traditional~Cable Networks)	0.110	0.385% (2)	n.a.
Golan & Wanta (2001, Table 3, p. 254, Best Chance to Win x Poll~Positive Coverage)	0.037	n.a.	1.389% (2)
Golan & Wanta (2001, Table 3, p. 254, Cut Taxes x Poll~Media)	0.259	1.843% (4)	n.a.
Golan & Wanta (2001, Table 3, p. 254, Trust x Poll~Positive Coverage)	0.024	n.a.	2.439% (2)
Lynn (1981, Table 2, p. 15, Marital Status)	0.014	n.a.	0.538% (8)
Lynn (1981, Table 2, p. 15, Education)	< 0.001	n.a.	3.146% (46)
Murphy, Kangun, & Locander (1978, Exhibit 3, p. 64, Soft Drinks)	0.001	n.a.	3.955% (7)

REFERENCES

Bush, A. J. & Leigh, J. H. (1984). Advertising on cable versus traditional networks. *Journal of Advertising Research*, *24*(2), April/May, 33-38.

Dickinson, J. R. (2017). An anomaly in MTABS analysis of cross-tabulations" (abstract). *Proceedings*, Decision Sciences Association 48th Annual Conference, p. 1289154-1-

Dickinson, J. R. (2016). Two observations between 'truth' and meh: a complement to cross-tabulation chi-square hypothesis testing (abstract). *Proceedings*, 2016 Decision Sciences Institute 47th Annual Conference.

Golan, G. & Wanta, W. (2001). Second-level agenda setting in the New Hampshire primary: a comparison of coverage in three newspapers and public perceptions of candidates. *Journalism and Mass Communication Quarterly*, *78*(2), Summer, 247-259.

Iacobucci, D. & Churchill, G. A., Jr. (2010). *Marketing research: methodological foundations*, Tenth Edition. Mason, OH: South-Western Cengage Learning.

Kinnear, T. C. & Taylor, J. R. (1991). *Marketing research: an applied approach*, Fourth Edition. New York: McGraw-Hill, Inc. ISBN: 0-07-034757-3

Lynn, J. R. (1981). Newspaper ad impact in nonmetropolitan markets. *Journal of Advertising Research*, *21*(4), August,, 13-19.

Murphy, P. E., Kangun, N., & Locander, W. B. (1978). Environmentally concerned consumers–racial variations. *Journal of Marketing*, *42*(4), October, 61-66.

Reynolds, H. T. (1984). *Analysis of nominal data*, Second Edition. Beverly Hills, CA: Sage Publications. ISBN: 0-8039-0653-6

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill Book Company. ISBN: 07-057348-4

SPSS (1999). *SPSS base 10.0 applications guide*. Chicago: SPSS Inc. ISBN: 0-13-017901-9

Yoo, K., Joo, E., Choi, H., Reid, L., & Kim, J. (2015). Trends in the use of statistics in major advertising journals over four decades. *International Journal of Advertising*, *34*(3) , May, 549-572.

DECISION SCIENCES INSTITUTE

A Preliminary Study in Text Mining of Kickstarter Crowdfunding Projects

Xiaochen Zhu
University of Massachusetts Boston
Email: Xiaochen.Zhu001@umb.edu

Josephine Namayanja
University of Massachusetts Boston
Email: Josephine.Namayanja@umb.edu

ABSTRACT

Crowdfunding platforms such as Kickstarter improve financial resource allocation efficiency and act as an essential part of financial innovation. Online crowdfunding platforms provide financing opportunities to a large group of individuals as well as organizations. In this study, we examine the effect of textual descriptions on each project's success rate. Specifically, we generate text-based models that utilize weighted and un-weighted term frequency measures respectively and evaluate their performance using multiple classifiers. We conduct experiments using real-world crowdfunding cases from Kickstarter.

KEYWORDS: Crowdfunding, Text analytics, Classification

INTRODUCTION

Crowdfunding is derived from the concept of crowdsourcing, which is defined as a way to harness the creative solutions of a distributed network of individuals. The first successful example of crowdfunding happened in 1997, when a British rock band took online donations to pay for a tour. Due to the development of online crowdfunding platforms and an increase in the types of crowdfunding projects, there has been an exponential increase in researchers' attention. Crowdfunding differs from traditional financing methods such as banking and venture capital investments mainly because it is designed to support specific and, in most cases, little projects or ideas by raising money from a crowd of people. Crowdfunding provides a revolutionary way to support creative ideas or projects in many industries, such as technology, art, crafts, publishing. With the current trend of online platforms, crowdfunding platforms lend themselves to the development of internet finance, which is a new form of financial activity that takes place primarily over the internet or any other internet protocol network. It is therefore an integration of modern finance and technology (Yang et al., 2018). Internet finance is characterized as the intersection of information technology, big data, and social networks. It improves the financial resource allocation efficiency by creating new avenues for finance in the network era. The combination with technology increases information transparency in the financial market as it relies on information technology to provide transaction information, as well as to coordinate and monitor the transactions. Big data also plays an important role by accelerating the competition and provide more transparent information which reduces the information asymmetry and uncertainty in investment returns.

Driven by trends of sharing and collaborative economy, and fueled by the internet and social media networks, new forms of fundraising and finance activities, based on collective intelligence, are growing. Crowdfunding is an open-market activity to gather financial resources from the public, which mostly takes place on internet-based platforms without standard financial intermediaries. It enables fundraisers to evade complicated regulatory requirements and reduce transaction costs by not requiring standard financial intermediaries (Mollick, 2014). Crowdfunding has many benefits due to its simpler and faster process compared to traditional means. It also serves as a platform for feedback and constructive ideas from the public, as well as a marketing strategy based on high accessibility to an extensive crowd of people.

A number of studies have been geared towards identifying factors that determine the success of crowdfunding projects (Etter et al., 2013; Greenberg et al., 2013; Mollick, 2014; Yuan et al., 2016). The focus has primarily been on structured attributes of projects such as, goal amount, duration of project, category of project, among others. However, the evaluation of unstructured attributes such as the 'blurb', a textual description of the project, is still an open challenge. The textual description is a pitch that presents the purpose and highlight of a project. We argue that unstructured attributes, in this case, the textual description of a project, contain messages pertinent to the success of a project. This research therefore aims to examine the predictive power of textual descriptions of each project's blurb on the project's success rate using text mining. For this study, we examine real-world cases scraped from Kickstarter. Kickstarter started in 2009 and is one of the biggest and most popular crowdfunding platforms on the internet. Similar to many other crowdfunding platforms, it aims to support individuals with small projects and creative ideas, in art, film, food, games, and others. The mechanism on Kickstarter is reward-based and All-or-Nothing. This means if a campaign does not achieve its' goal by the pre-determined deadline, it will not collect any funds. Therefore, the successful fulfillment of a funding goal is imperative for all projects. Our study extracts terms from the textual description and determine their frequency based on the following; (i) term frequency which focuses on high-frequency terms and (ii) td-idf which focuses on the reversed frequency of each term by identifying distinctive terms. In each case, we generate a feature-set in the form of a term matrix, which is used to predict the success rate of project.

Given only less than half of all launched campaigns reach their funding goal on the Kickstarter (Kickstarter), it is extremely important for fund seekers to know the factor(s) that potentially impact project outcomes. The proposed research approach poses potential benefit to entrepreneurs to identify the most influential textual features embedded in project descriptions in order to promote their projects better and improve chances of raising sufficient funds. This research makes the following contributions;

- We present a novel approach of analyzing unstructured features to predict crowdfunding success rate. Our proposed approach poses potential insights into the predictive power of using text analysis in crowdfunding success.
- We evaluate our proposed approach using multiple classification techniques in order to gain an understanding of suitable models for prediction.
- We conduct experiments using real-world crowdfunding cases from Kickstarter to demonstrate the effectiveness of our methods.

RELATED WORK

Crowdfunding Drivers: Crowdfunding is widely studied in many streams. Cumming et al. (2012) describe a French music crowdfunding start-up which is among the first studies of the crowdfunding. Mollick (2014) defines crowdfunding in an entrepreneurial context as a venture by drawing on small contributions from a large group of individuals using the internet characterized by an attempt to disrupt standard financial intermediation. Belleflamme et al. (2015) discuss various crowdfunding sectors and present a range of statistics on existing crowdfunding platforms. They also discuss the functioning of the crowdfunding platforms and the asymmetric information problems in the market. Duarte et al. (2012) examine the trust issue based on a P2P lending site built upon existing financial theories. Michels (2012) investigates the credibility issue in the crowdfunding market and finds that unverifiable disclosures influence the cost and activities in the lending. A group of scholars address the mechanisms in crowdfunding activities regarding possible factors that could influence the fundraising activities. In Ahlers et al. (2015), they focus on the effectiveness of signaling effects of some factors in the context of equity crowdfunding. Their study provides some empirical evidence that some information related to risks could be effective signals and can therefore strongly impact the probability of funding success. Colombo et al. (2015) study the early contributions on Kickstarter and find that internal social capital developed inside the crowdfunding community facilitates a self-reinforcing mechanism and fully mediated by the capital and backers collected in the campaign's early days.

Pattern Detection in Crowdfunding: Interestingly, there is a large body of literature describing the utilization of data mining techniques to analyze the behaviors and project characteristics on different platforms. Although crowdfunding websites such as Kickstarter and Indiegogo behave as online intermediary agents that allow project founders to quickly reach a large number of individual investors with minimal costs, initiators are not always guaranteed to reach their initial fundraising goals (Yuan et al., 2016). Studies about crowdfunding take on different perspectives. While some investigate from a participant or donor angle to check their motivation in participating in online crowdfunding activities (Gerber & Hui, 2013), others focus on project related characteristics that could lead to the success or failure of each fundraising project. In both cases, factors of most interest include number of backers, duration of the fundraising activity, or category of the project that requires funding among other structured features. On the other hand, this study addresses the gap in literature by utilizing data mining and focuses on the potential of unstructured features to characterize the success of projects.

DATA AND PRE-PROCESSING

For this study, we utilize data from Web Robots (Web Robots), which is a public data source. It provides scraped data from several sources including Kickstarter. Detailed data on projects is available on a monthly basis. The sample data used in this study is taken from a large pool of data that was last updated on March 14, 2019. In this data collection, all projects that are still alive on Kickstarter at the time the data is scraped are included. We download all posted projects on Kickstarter specific to the last updated timestamp. For purposes of this study, we randomly select several data files and end up with 37872 observations. According to the stats from Kickstarter (Kickstarter), there are more than 450000 projects launched so far. The raw scraped dataset includes 37 features. Many features are about identification information of scraped projects, for example, URLs about sources, creators' information, links of video, among others, which are not relevant to this study and we remove them. Our final set consists of five features. In table 1, we provide a summary of the features selected for this study.

The target feature or class label is "state", which is the status of the project at the time of data collection. Given that this study is targeted at investigating successful and failed projects, we eliminate any projects whose state is described as "canceled", "suspended" and "live". There are also several features with unstructured values generated as a result of raw data directly scraped from website. In order to preserve the identifying information of projects, we extract additional features such as category, category id, location, location id and project id. However, only the blurb and state are used during the prediction phase.

Table 1. Summary of Features and Descriptions

Feature	Description	Data Type
state	Outcome of project. Five states: successful, failed, canceled, live, suspended.	String
category	Project category in detail (e.g. Web, product design, accessories)	String
goal	Goal in dollars of project	Integer
blurb	Text description of project	String
location	Location of project (e.g. Los Angeles, London, New York)	String

Further data cleaning is applied to check for missing values and duplicate data. Following data cleaning, the selected sample consists of 34160 observations. In figure 1 we provide an overview of the rate per state for projects - successful versus failed.

Figure 1. Rate of Successful versus Failed Projects

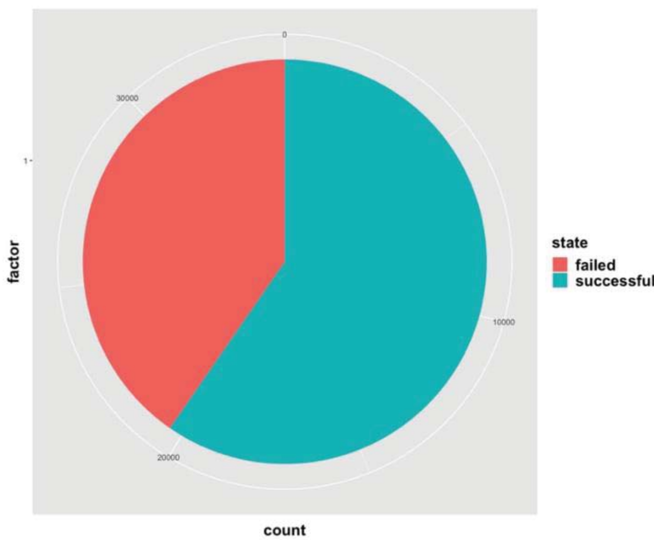


Figure 1 clearly shows that there are more successful cases in the selected sample. This can be attributed to the fact that most of the cases were still live. Nevertheless, it can also be argued that there is a rapid growth in successful projects in recent years given that Kickstarter has been around for 10 years.

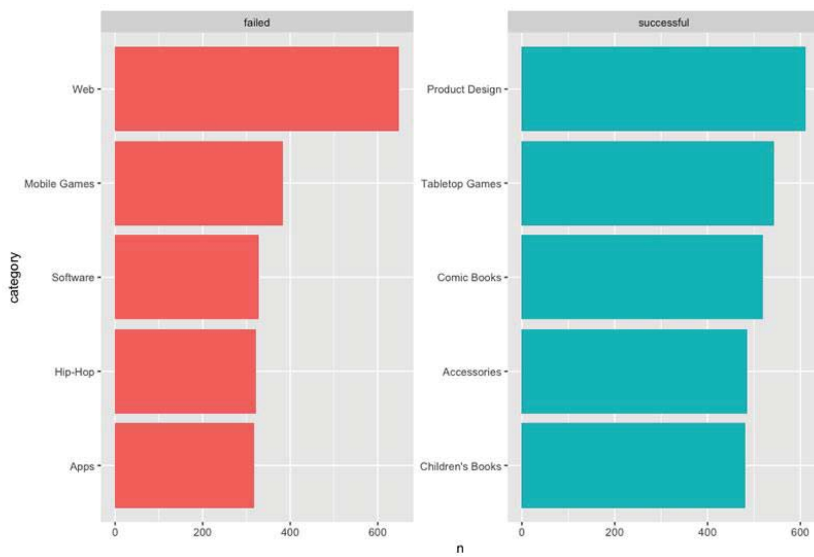
By drilling down into the various projects, the most frequent crowdfunding projects are web-based, followed by product design, accessories, and others. Additionally, most projects are located in Los Angeles, California. In figure 2. we provide a summary of the project counts based on the category and location.

Figure 2. Project Counts by Category and Location

```
        category_name          location_name
Web           :  777   Los Angeles  : 1569
Product Design:  612   London       : 1274
Accessories   :  556   New York     : 1225
Tabletop Games:  543   Chicago      :  636
Comic Books   :  520   San Francisco:  577
Apparel       :  499   Brooklyn     :  540
(Other)       :30653   (Other)      :28339
```

We extend our descriptive statistics to understand the distribution of successful and failed projects by category. As shown in figure 3, web-based, mobile games and software are predominant as the most failed categories in projects. While the most successful cases belong to product design, tabletop games, comic books, among others.

Figure 3. Distribution of Project State per Category

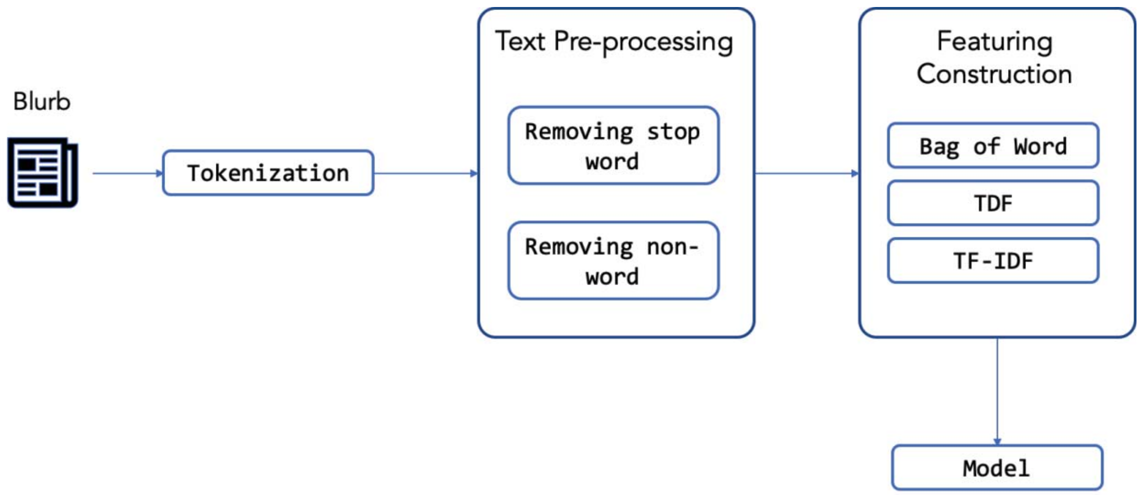


PREDICTION MODEL

Our focus in this research is the textual description, particularly the project blurb and trying to determine its impact on the funding goal. In order to conduct text analysis, we generate a text-based feature set from all project blurbs and then apply classification using state as the target feature. Feature extraction is an essential process for textual analysis (Ravisankar et al., 2011; Tsai & Hsiao, 2010; Yuan et al., 2016). In figure 4 we provide an overview of our proposed text mining approach.

First, we consider the blurb of each project as a document, then apply word segmentation to divide each English term. For this study, we only keep those descriptions in English and remove non-English (e.g. Japanese, French) descriptions due to limitations in the text handling process. We also remove stop words, numbers or other symbols. From that we construct a Document Term Matrix (DTM) using two methods: (1) term-document frequency (tdf) which is based on the frequency of each word in each document; (2) term frequency-inverse document frequency (tf-idf) which is the probability that combines within document-term frequency with the inverse of the expected document frequency in the collection (Amati & Rijsbergen, 2002). Simply put, tf-idf assigns a higher weight to those distinct words which appear less frequently across documents. Using a DTM generates a high dimensional space which often results into high sparsity. Therefore, we also remove sparse terms by limiting the sparsity level at 99%. Our objective is to compare the two matrix outcomes and their predictability on the success rate of projects.

Figure 4. Overview of Text Mining Process



Finally, we apply classification on the term feature sets respectively and using target label - state defined as successful or failed. Due to computational complexity, we only include up to 3000 observations during classification. We also apply data balancing between classes where successful and failed cases assigned a 55/45 ratio. This proportion is in line with the total sample scale as shown in figure 1. We apply k cross-validation using k = 5. In this study, we compare three classifiers, specifically, decision tree (C4.5), K-nearest neighbors and naïve Bayes, all of which are widely used in classification tasks. However, other classifiers are also applicable. For each classifier, we test both term matrices to evaluate their performance.

RESULTS

For model evaluation, we apply the several widely used metrics in classification tasks, specifically: Accuracy which measures the number of correctly classified instances, and precision and recall which measure the rate of true positive cases as a percentage of false positive and false negative instances respectively. In table 2, we provide a summary of the evaluation results for the selected classifiers based on with two different sample sizes, that is using 2000 and 3000 observations.

Table 2. Classification Metrics

	C4.5_TDF	C4.5_TF-IDF	KNN_TDF	KNN_TF-IDF	NB_TDF	NB_TF-IDF
Sample size=2000						
Accuracy	0.5555	0.5775	0.4625	0.5730	0.45	0.45
Precision=TP/(TP+FP)	0.5930	0.5734	1.0000	0.5670	NA	NA
Recall=TP/(TP+FN)	0.6145	0.9018	0.0236	0.9382	0	0
Sample size=3000						
Accuracy	0.5952	0.6032	0.4695	0.5732	0.45	0.45
Precision	0.6206	0.5918	0.9130	0.5724	NA	NA
Recall	0.6782	0.8964	0.0382	0.8836	0	0

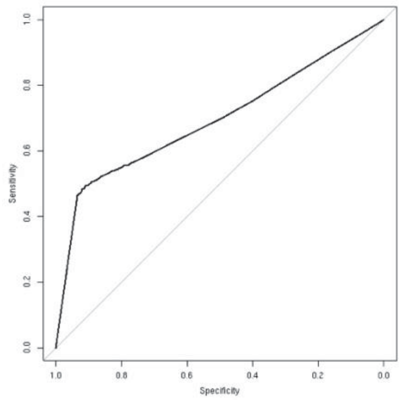
Our findings indicate that tf-idf outperforms tdf, particularly for C4.5 and KNN as indicated by slightly higher accuracy, and significantly higher recall values. The higher recall for tf-idf compared to tdf for both C4.5 and KNN indicates that measuring distinct terms using tf-idf generates a lower rate of false negatives. The overall performance of naïve Bayes is low across all evaluation metrics compared to other classifiers.

On the other hand, while accuracy and recall levels for C4.5 and KNN are higher in tf-idf cases, it is observed that precision levels are higher with tdf for both classifiers particularly for KNN with precision level is > 90% for tdf and < 60% for tf-idf in both sample sizes.

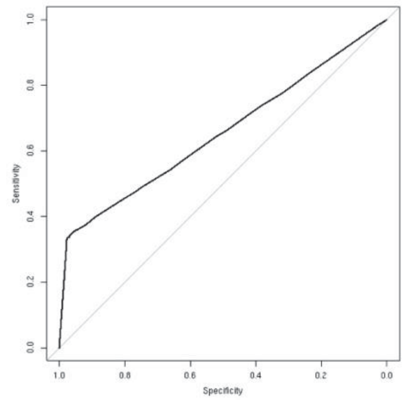
We also observe similar patterns in the receiver operating characteristic curves – area under the curve (ROC-AUC) showing a trade-off of the true positive rate (sensitivity) versus the false positive rate (specificity) at different thresholds. In figure 5a and 5b, our findings show an AUC of slightly more than 50% for all C4.5 and KNN classification except for KNN with tdf matrix.

Based on the prediction outcomes, we drill down to understand the terms utilized in the feature sets. As shown in figure 6, we present a word cloud generated for successful projects. Overall, it is observed that a lot of frequent words are related to the category of projects. Therefore, for future studies, we plan to apply controls on the inclusive terms in order to try to improve the purity of the text.

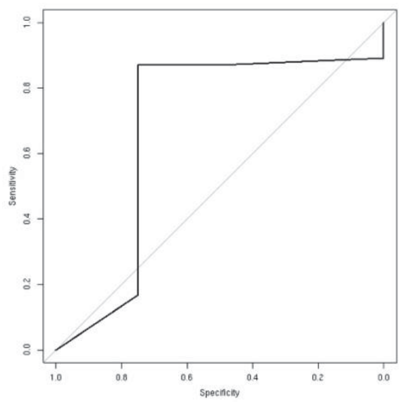
Figure 5a. ROC Curves for Sample Size of 2000



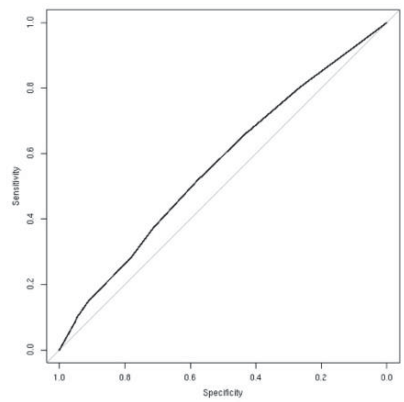
C4.5 TDF



C4.5 TF-IDF



KNN TDF



KNN TF-IDF

Figure 5b. ROC Curves for Sample Size of 3000

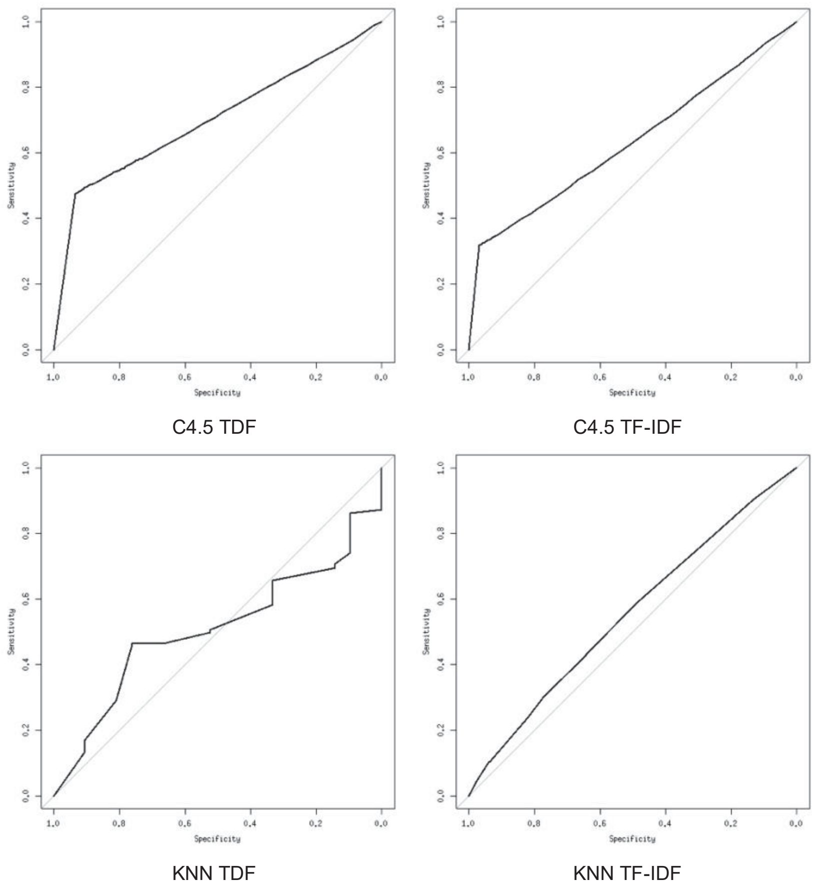


Figure 6. Word Cloud for Successful Projects

DISCUSSION AND FUTURE DIRECTIONS

This research is a preliminary study to utilize text-based descriptions to predict the success or failure of crowdfunding projects. Our findings show that text features predict the success of projects approximately 60% of the time. Furthermore, by comparing different classifiers and feature sets, we detect some interesting information pertaining to the model.

Given the limitations associated with this study such as focus on English terms, we plan to apply advanced text pre-processing methods to translate other languages and symbols. Our study has only been limited to 3000 instances. Therefore, we plan to apply distributed processing so can extend our methods to larger data samples. Additionally, we plan to combine structured and unstructured features in predicting project success in order to improve the robustness of the results in future studies.

Furthermore, given that the rate of fundraising for a project is driven by the will of the funder's feelings, it would be interesting to understand the sentiment included in the textual descriptions to determine if strong sentiments are expressed by the fundraisers and how this potentially influences project outcomes. On the other hand, a case-by-case analysis can be applied to understand why some cases supersede the goal. Other factors influencing the project outcomes in terms of different amounts could also help future fundraisers derive improved strategies on the crowdfunding platform. In the future, we also plan to explore ensemble learning in order to account for the diversity in how text is utilized.

REFERENCES

Ahlers, G. K. C., Cumming, D., Gunther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship: Theory and Practice, 39*(4), p 955+.

Amati, G., & Rijsbergen, C. J. V. (2002). Probabilistic models of information retrieval based on measuring the divergence from randomness. *ACM Trans. Inf. Syst., 20*(4), pp. 357-389.

Belleflamme, P., Omrani, N., & Peitz, M. (2015). The economics of crowdfunding platforms. *Information Economics and Policy, 33*, pp. 11-28.

Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal Social Capital and the Attraction of Early Contributions in Crowdfunding. *39*(1), pp. 75-100.

Cumming, D., Schwienbacher, A., & Larralde, B. (2012). Alternative Types Of Entrepreneurial Finance: Oxford University Press.

Duarte, J., Siegel, S., & Young, L. (2012). Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *The Review of Financial Studies, 25*(8), pp. 2455-2484.

Etter, V., Grossglauser, M., & Thiran, P. (2013) *Launch Hard or Go Home! Predicting the Success of Kickstarter Campaigns*. Paper presented at the COSN'13, Boston, Massachusetts: ACM, pp. 177-182.

Gerber, E. M., & Hui, J. (2013). Crowdfunding: Motivations and deterrents for participation. *ACM Trans. Comput.-Hum. Interact., 20*(6), pp. 1-32.

Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013) *Crowdfunding support tools: predicting success & failure*. Paper presented at the CHI '13 Extended Abstracts on Human Factors in Computing Systems, Paris, France : ACM, pp. 1815-1820.

Kiskstarter. Stats. Retrieved from https://www.kickstarter.com/help/stats on May 30, 2019.

Michels, J. (2012). Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending. *The Accounting Review, 87*(4), pp. 1385-1413.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing, 29*(1), pp. 1-16.

Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems, 50*(2), pp. 491-500.

Tsai, C.-F., & Hsiao, Y.-C. (2010). Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems, 50*(1), pp. 258-269.

Web Robots. Kickstarter Datasets. Retrieved from https://webrobots.io/kickstarter-datasets/ on Mar 20, 2019.

Yang, D., Chen, P., Shi, F., & Wen, C. (2018). Internet Finance: Its Uncertain Legal Foundations and the Role of Big Data in Its Development. *Emerging Markets Finance and Trade, 54*(4), pp. 721-732.

Yuan, H., Lau, R. Y. K., & Xu, W. (2016). The determinants of crowdfunding success: A semantic text analytics approach. *Decision Support Systems, 91*, pp. 67-76.

DECISION SCIENCES INSTITUTE

Intellectual Property Theft
What are the Chinese stealing? How are they stealing it? What can we do about it?

Anita Reed
Texas A & M University – Corpus Christi
Email: anita.reed@tamucc.edu

William Harding
Texas A & M University – Corpus Christi (Retired)
Email: bharding_42@yahoo.com

ABSTRACT

We began this project by interviewing twelve Fortune 500 Companies. Most of these companies were very reluctant to directly share their experiences with IP theft. From their reactions, however, it was quite evident that all of these companies did have direct experience with some intellectual property (IP) theft issues. One company (referred to as Company X) proved to be very helpful in this study. Two other companies (referred to as Company Y and Z) offered some information. The information that we were able to collect is presented. First, in our paper, we define what IP theft is, and how it occurs. Next we discuss several examples of IP theft and the Chinese government's continuing support of it. Next we look at what some companies are doing to better protect themselves from IP theft. Finally, we look at what impact the U.S. trade tariffs and the trade talks are having, or not having, relative to Chinese IP theft.

Key Words: Intellectual property (IP), IP theft, forced IP transfers, copyrights and trade secrets, counterfeit goods, cloning, cybersecurity, proprietary knowledge, reverse engineering, trade tariffs, IP protection, No-spy agreements.

INTRODUCTION

Intellectual property (IP) is essentially any unique product of the human intellect that contains commercial value. Examples of IP are books, songs, movies, paintings, inventions, chemical formulas, computer software, etc. In the U.S. we have IP protection through: Trade secrets, trademarks, patents, and copyright laws. Trade secrets do not expire, but other companies can conceivably reverse engineer the formula or process. The advantage of a patent is that the U.S. government gives the patent owner the exclusive right to the IP, but the patent expires after 20 years. Patents are considered an unreliable way of protecting IP for software because the U.S. Patent Office has given out many bad/unreliable software patents that cannot hold up in court. Therefore, many times software is better protected in the U.S. through the Copyright laws. In the U.S. Trademarks can be registered and offer some protection from IP theft.

RESEARCH QUESTIONS AND METHODOLOGY

We are attempting to identify what kinds of IP are being stolen, and specifically what are the Chinese stealing. Additionally, identify what methods are commonly employed to steal IP, and what companies are doing to prevent IP theft. Our methodologies include: Interviewing 12 Fortune 500 companies, literature review, and current events in the news.

LITERATURE REVIEW

The Interviews

We interviewed twelve Fortune 500 Companies. They were very reluctant to directly identify specific cases of IP theft within their respective companies. Some of these companies were willing however, to talk about things that they were doing to help prevent IP Theft. We will discuss these examples later in this paper.

Annually, the US loses from $225 billion to $600 billion due to Chinese intellectual property theft. While not the only nation that steals intellectual property from the US, it is the main one. Some companies were infiltrated by Chinese spies who brought the information back to China, or they bribed insider employees. Theft is most prevalent; however, it is not the only way the Chinese are creating and selling similar products to those in the US. As much as 87% of seized counterfeit goods imported into the US were produced using reverse engineering of trademarked, patented, and copyrighted materials from many foreign companies. (Saady, 2018).

China's theft of intellectual property is encouraged by official Chinese policy and often involve active participation of government personnel. The theft of intellectual property includes fashion designs, pirated movies / video games, patent infringement, theft of proprietary software, and much more (Blair, 2017).

Since intellectual property helps the Chinese economy and thus, its government, the Chinese government, currently, does not plan on cracking down on most of these cases. In fact, intellectual property theft in China isn't illegal. When a US company wants to go into the Chinese market, they are faced with laws that force them to collaborate with Chinese companies in several key industries. This forces many companies to give their intellectual property to the Chinese. The Chinese realize that they need foreign technology companies, so they've created a means to legally obtain the information they want (Oviti, 2018). As of the writing of this paper, the U.S. government and the Chinese government are participating in Tariff/Trade talks. Hopefully, as a result of these talks, this situation can be improved. The Chinese do recognize the need for IP protection, especially for themselves.

Examples of Stealing/Cloning Techniques

The Beijing company Ninebot decided to market its own version of the two-wheeled self-balanced single-person vehicle after the debut of the pioneering technology by Segway in the US. The design is uncanny to that of the Segway which was marketed and sold in the US alongside its competitor. So Segway pursued an import ban on Ninebot for alleged stolen patents and design. It was thereafter that Ninebot outright bought Segway for a confidential amount (Ninebot, 2015). This is seen as a way of "quelling" the competition and silencing them after exploiting them. A second example happened when DuPont had its method of making a shimmering white color to its products, a chemical engineering problem that it invested a substantial amount of time and money. It was so successful that China sought after the process but for a quicker way to do it with less effort compared to the DuPont originator. So, they used a nationalized American citizen from their country to coerce former DuPont engineers to divulge their secret process. Eventually one of the DuPont engineers cracked, made a deal with China, and helped establish a factory with them that would take the color with DuPont's knowledge in tow. DuPont won their court case after seeking reparations, but ultimately lost the war as its trade secret it kept a lid on for making white was forever exposed to the world, ruining any market advantage (Lesperance, 2018).

A Chinese firm, Sinovel, "contracted with American Superconductor Corp. (AMSC) for more than $800 million in products and services to be used in wind turbines it manufactured" (Raymond, 2018). However, according to the same Reuters article, "Sinovel recruited Dejan Karabasevic, an employee of an AMSC unit, to join the Chinese company and to secretly copy information from AMSC's computer system, including the source code for the PM3000, part of its wind turbine control system". The code they stole was used in wind turbines they commissioned in Massachusetts. Following this, the American company AMSC lost over $1 billion and had to cut 700 jobs.

Several companies have suffered because their ideas have left their headquarters, boarded a plane and showed up in a different market, before the company could even decide on a launch date. One of the most recent IP Thefts was caught before it left the country. An ex-Apple employee, Xiaolang Zhang, was arrested at the airport by the FBI before attempting to board a flight to China. Zhang had worked as an engineer on a project developing self-driving cars. He's accused of stealing this information with plans of using it as a Chinese startup company called X Motors (Saady, 2018).

China has been thriving off stolen ideas for decades; this is nothing new. Outside of the tech industry the knock-off brand market in China is very well known and thriving to this day. China has stepped up and put in motion different tariffs and laws to better amend these detrimental blows to foreign companies. In 2017, China started a nationwide campaign to protect foreign firms' international property rights. The Ministry of Commerce says China is a developing country and doesn't have a perfect system to protect IP, acknowledging that there's much work to do…" (Bloomberg, 2018).

Huawei Technologies is a Chinese multinational conglomerate which specializes in telecommunications equipment, consumer electronics, and technology-based services and products. Huawei's smartphones tablets, etc. can be purchased from Sam's Club, Walmart, Amazon and many others. In 2014, T-Mobile filed a civil case against Huawei. Huawei was accused of stealing intellectual property related to a robot that T-Mobile used to diagnose quality control issues in mobile phones. A jury in Seattle found Huawei guilty in May 2017. In December 2018, Huawei CFO, Meng Wanzhou (the daughter of Huawei's founder), was arrested in Canada for extradition to the U.S. related to alleged violations of Iranian sanctions. (Wakabayashi, 2018). Canada now blames U.S. for the Huawei CFO arrest backlash that has left two Canadian citizens in Chinese prisons and a third Canadian citizen on death row. On December 11, 2018, President Trump told Reuters that he might intervene in Meng's case if it would help his trade war with China. "If I think it's good for the country, if I think it's good for what will be certainly the largest trade deal ever made – which is a very important thing – what's good for national security – I would certainly intervene if I thought it was necessary," the president said. In late January the U.S. Justice Department charged Meng and Huawei with conspiring to violate U.S. sanctions on Iran. On Friday, March 1, 2019 the department of Justice Canada officials issued an authority to proceed, formally commencing an extradition process in the case Ms. Meng Wanzhou. China, whose relations with Canada have deteriorated badly over the affair, denounced the decision and repeated previous demands for

Meng's release. Legal experts had predicted the liberal government of Prime Minister Justin Trudeau would give the go-ahead for extradition proceedings, given the close judicial relationship between Canada and the United States. It could be years though before Meng is ever sent to the United States, since Canada's slow-moving justice system allows many decisions to be appealed. Professor Wesley Wark of the University of Ottawa's Graduate School of Public and International Affairs said "the Canadians will take a beating throughout this whole process" from China. "I suspect the Trudeau government is desperately hoping that the Americans reach a deal with the Chinese," he said. Brock University professor Charles Burton, a former Canadian diplomat who had served two postings in China, said Beijing was likely to retaliate further. "They're not going to take this lying down… one shudders to think what the consequences could be," he told the Canadian Broadcasting Corp., saying Beijing might crack down on Canadian canola shipments or stop Chinese students from going to Canada. (Ljunggren, 2019)

Other news around the world: Huawei and fellow Chinese company ZTE were both banned from supplying 5G equipment to Australia's wireless network in August, 2018 on national security grounds. China's Huawei offered Berlin a "no-spy agreement" to address security concerns over the Chinese company's involvement in building Germany's next-generation 5G mobile infrastructure, a German magazine said. "We (Huawei) talked to the German Interior Ministry and said that we were ready to sign a no-spy agreement with the German government and to promise that Huawei will not install any backdoors in the networks," Wirtschaftswoche (of Germany) quoted Huawei Chief Executive Ren Zhengdei as saying. He called on the Chinese government to sign a similar no-spy-agreement and to adhere to European Union data protection laws. Germany last month set tougher criteria for vendors supplying network equipment, stopping short of singling out Huawei for special treatment and instead saying the same rules should apply to all vendors. (Alkousaa, 2019)

What some Companies are doing to prevent IP Theft

Intellectual property protection is crucial to any business. Complete secrecy and information block-outs seem to be the first line of defense. However, IT can play an important support role in protection. Cybersecurity measures must be implemented to protect proprietary knowledge stored in computer servers. Servers should be separated from outside networks and access privileges should be limited. The information should be encrypted, and network traffic should be monitored. Digital watermarking and time stamping can lace ownership rights into digital materials to help prove ownership in cases of litigation. Along with these typical cybersecurity measures, new technologies have led to new problems and new solutions. A new problem in copyright infringement is the availability of 3D printers. CAD files are shared online with printing information for patented products and designs. It is hard to track down the printers of these files, but companies should use web crawlers to find these CAD files to make sure their products do not end up being easily counterfeited (Wessing, 2017).

Corporations can only do so much to protect their copyrights and trade secrets. Educating the public to be conscious of the difference between a genuine product and a fake one, and the economic consequences of purchasing a fake, is one way that companies can combat this kind of theft. Most governments have had some type of intellectual property law on the books for years, which allows companies to use fear of prosecution as a deterrent. Governments around the world have begun writing laws specific to new circumstances that have arisen in the cyber-age. When state sanctioned theft occurs across nations there is not much a single company can do to stop a foreign company from profiting from the property they stole. As these cases have become more prevalent, the US Government is now taking a stand to back US companies in the fight again international intellectual property theft.

The following involves an interview with one manager at one Fortune 500 Company (who did not want to be quoted), which we will call Co. X: In a personal interview with manager James, of Co. X, He explained how they use a Hierarchy system to enact the principle of least privileges. Co. X uses an employee number to identify the privileges you have, the higher the number - the more privileges an employee possesses. They also require that all physical documents to be shredded. In addition, Co X utilizes a segmented network, not keeping all of their data in one location. In the event of a data breach, the compromised network is the only network accessible by the hacker thus minimizing the severity of damage to the company. James also explained that Co. X monitors their logs regularly. They monitor information being accessed as well as by whom. Co. X also requires that input devices such as USBs be restricted throughout all of their buildings. The interview with James was very informative on how a corporation handles the threat of intellectual property theft. Co. X has appropriate safe-guards in place for the biggest threat to a company's intellectual property, an insider. With least privilege access, persistent auditing, and encryption. This is one of the multitudes of contributing factors as to why Co. X has only a very few cases of IP theft.

Company "Y" uses several controls to limit their threat risk. They have 24-hour security officers and their security systems are monitored 24 hours a day, 365 days a year for signs of tampering or unauthorized logins. They also use encryption, a VPN, vulnerability testing as well as limiting their own employee's access to their systems.

Company "Z" requires all their employees and users sign off on an Intellectual Property Notice. The notice basically says that all the content, information and materials on their site, such as text, graphics, logos, images, data, software and other material is owned or licensed by Company "Z" is proprietary and protected by copyright, trademark, patent, and/or other intellectual property rights. Company "Z" grants only a limited use, and the user may not copy, reproduce, modify, remove, delete, add to, publish, transmit, transfer or sell any of this material.

What Impact are the Tariffs having?

As of March of 2018, the president of the United States—President Donald Trump—imposed strict tariffs on China in order to protect U.S. technology (Trump, 2018). While the tariffs were made as a case for fair trade and "to defend our industry and create a level playing field for the American worker", the tariffs are largely a response to Chinese firms and businesses that were stealing intellectual property from United States persons or businesses.

On CNBC the other day, Kevin O'Leary, a CNBC Contributor, present a "drain cleaning rod" that was invented and being sold by an American company. The device has been cloned and is being sold by other companies from foreign countries, but because of the Tariffs on Chinese imports, it is too expensive to buy the clones from China. So, this is one very small victory supporting U.S. Tariffs on China. However, our farmers and live-stock producers, that were selling many of their products to the Chinese, continue to lose billions of dollars in lost sales every day with the tariffs still in place. What effect this tariff war will eventually have on our trade policy with China, is anybody's guess. We are hopeful that as an outgrowth of these trade talks, our U.S. government will then be able to actively assist U.S. companies (large and small) to receive no-spy, no IP theft, and no invasion of privacy agreements with all other countries. It is the enforcement side of these potential agreements (especially with China) that U.S. companies will need the most help.

More Recent Developments

Recent developments have occurred that indicate the issue has not been resolved and continues to be in issue:

May 15: The Trump administration (via executive order) added the Chinese Company Huawei to the U.S. department of Commerce's "Entity List" (blacklist). This has had a large impact on U.S. companies such as Google, Intel and Qualcomm that were using Huawei's products. Most U.S. companies have publicly stated that they will obey the administration's order.

May 20: The Trump administration's Commerce Department issued a temporary license that will allow Huawei to maintain its current products (for existing customers). The license expires August 19, 2019.

May 23: President Trump has called Huawei "very dangerous," but said the U.S. is open to including the company as part of a future trade agreement between the U.S. and China.

May 28: Huawei filed a legal motion claiming the ban on their company working with other U.S.-based companies violates the U.S. Constitution.

May 31: China threatens to create its own "Entity List" (blacklist) to include American firms. Huawei ordered employees to cancel technical meetings with American contacts. Huawei also reportedly sent back American citizens who worked in research and development roles.

June 6: Huawei will build a 5G network for Russia's largest carrier.

June 7: Facebook will no longer allow Huawei to pre-install any of its apps on the company's smartphones. These apps include Facebook, Instagram, and WhatsApp, three of the most popular apps in the world. The ban only applies to phones that have not yet left the factory.

June 29: On the sidelines of the G20 summit in Osaka, Trump said "U.S. companies can sell their equipment to Huawei," without going into detail. "We're talking about equipment where there's no great national security problem with it," Trump continued. It's not clear what this means for now.

July 16: A bipartisan group of senators has introduced a bill, "The Defending America's 5G Future Act" that would effectively set the original Huawei blacklisting in stone. It would "codify" the executive order forbidding sales of telecom equipment to customers posing national security risks, and bar the removal of Huawei from the Commerce Department Entity List without an act of Congress.

August 15: Are the Hong Kong protests related to the US/China trade war? Maybe? In 1997, the Crown Colony of Hong Kong officially reverted to Chinese sovereignty, ending 156 years of British rule. Since then, Hong Kong has continued to enjoy more open trade commerce than mainland China. However, other basic freedoms and rights have been changed to conform to mainland China. In recent years the Hong Kong protests by its citizens continue. If the trade war is not settled, Hong Kong could also lose some of its open trade agreements with the U.S. As of the paper submission for publication date: The Trade War between the U.S. and China escalates from one retaliation to another, but some helpful tweets and comments also come about. The two might continue their trade talks in September. Huawei CFO Meng Wanzhou still resides in Canada and is still fighting extradition to the U.S.

RESEARCH CONCLUSIONS

IP theft is rampant and world-wide. Most IP theft is done through cloning, copying, and reengineering. In the past, the Chinese have not wasted their time or money on research and development. Mostly, they have found ways to just steal it, and with their government support. All twelve of the Fortune 500 companies that we interviewed have experienced IP Theft. Most of these companies were reluctant to share their experiences, but they are doing what they can to protect themselves. Many companies have learned that doing business with the Chinese is risky. Small American companies (especially) could use American government support in their fight against IP theft.

REFERENCES

Alkousaa, Riham and Tassilo Hummel. “News about Huawei CEO.” Bing.com/news. 15 Apr. 2019.

Blair, Dennis C., and Keith Alexander. “China's Intellectual Property Theft Must Stop.” *The New York Times*, The New York Times, 15 Aug. 2017, www.nytimes.com/2017/08/15/opinion/china-us-intellectual-property-trump.html.

Bloomberg.com, Bloomberg, www.bloomberg.com/news/articles/2018-03-22/what-s-intellectual-property-and-does-china-steal-it-quicktake.

Lesperance, Jeffery. “Steps to Protect American IP from China.” *IndustryWeek*, 19 Sept. 2018, www.industryweek.com/technology-and-iiot/steps-protect-american-ip-china.

Ljunggren, David. “Canadian Extradition Proceedings Against Meng Wanzhou.” Ottawa (Reuters), 01 Mar. 2019.

Ninebot. “Chinese Rival Firm Ninebot Buys Segway.” *BBC News*, BBC, 15 Apr. 2015,www.bbc.com/news/business-32316865.

Oviti, Thomas. “China is still stealing intellectual property — but that's not the biggest problem,” 25 Sep. 2018. https://www.abc.net.au/news/2018-09-26/china-is-after-intellectual-property-not-always-illegally/10302424

Raymond, Nate. “China’s Sinovel Fined in U.S. Trade Secrets Theft Case,” Reuters, 6 Jul 2018. https://www.reuters.com/article/us-sinovel-wind-gro-usa-court/chinas-sinovel-fined-in-u-s-trade-secrets-theft-case-idUSKBN1JW2RI

Saady, Brian. “The Unreal Scope of China's Intellectual Property Theft.” *The American Conservative*, July 23, 2018. www.theamericanconservative.com/articles/the-unreal-scope-of-chinas-intellectual-property-theft/.

Trump, Donald. “President Donald J. Trump is Standing Up for American Innovation.” Whitehouse Briefing, 22 Mar 2018, https://www.whitehouse.gov/briefings-statements/president-donald-j-trump-standing-american-innovation/

Wakabayashi, Daisuke, and Alan Rappeport. “Huawei C.F.O. is arrested I Canada for Extradition to the U.S.” *The New York Times*, 5 Dec 2018. https://www.nytimes.com/.../huawei-cfo-arrest-canasa...

Wessing, Taylor. “Unauthorized 3D printing – what use are intellectual property rights?” *Media and Tech Law*, March, 2017. https://www.taylorwessing.com/download/article-unauthorised-3d-printing.html.

DECISION SCIENCES INSTITUTE
Enhancing resilience in cyber security

Alessandro Annarelli
Sapienza University of Rome
Department of Computer, Control, and Management Engineering
Via Ariosto 25, 00185, Roma, Italy
Email: alessandro.annarelli@uniroma1.it

Fabio Nonino
Sapienza University of Rome
Department of Computer, Control, and Management Engineering
Via Ariosto 25, 00185, Roma, Italy
Email: alessandro.annarelli@uniroma1.it

Giulia Palombi
Sapienza University of Rome
Department of Computer, Control, and Management Engineering
Via Ariosto 25, 00185, Roma, Italy
Email: alessandro.annarelli@uniroma1.it

ABSTRACT

The organizational ability to react to cyber attacks recalls the concept of resilience. In order to provide a deep insight on how companies effectively introduce and implement cyber systems to enhance cyber resilience, a multiple case study analysis revealing interesting managerial implications is conducted.

KEYWORDS: Resilience, Cyber Security, Multiple Case Study

INTRODUCTION

Nowadays, according to complexity and level of digitalization of its business and to value offered to social and economic context, every organization has an IT system and the necessity of protecting its information and digitalized processes from fraud attempts and act of vandalism. These crimes, called cyber attacks, can be fronted by the implementation of a cyber resilient system.

The ability to react to these attacks recalls the concept of resilience, known in physics as the assumption of sustaining crashes without breaking. In management sciences, resilience is defined as the ability of an organization to adapt to unexpected disruptive changes or a trouble, for being able to continue the necessary operations both in expected condition and unexpected conditions (Wieland & Wallenburg, 2013).

Two features belong to a resilient system: (1) robustness against the predictable attacks and (2) ability to come back to a safe state if an attack has been successful, without regressing (Rogers *et al.*, 2016). This is the ability of the system to adapt, to reduce the probability of having to deal with sudden disturbances and having to resist their spread, maintaining control over structures and functions, and being ready to recover and respond with immediate and effective reactive plans to overwhelm the disturbance and restore to a robust state of operations.

The paper is organized as follows. The first section contains the literature review presenting the concepts of cyber security and resilience and the main stream of research for the cyber resilience: risk management process and resilience management process; further this section highlights the emerging research gap and defines the research question.
The third section (methods) presents the research methodology: a case studies analysis built thanks to the interviewing of cyber security managers: in this section the research protocol is detailed. The fourth section (results) contains the descriptive analysis of the cases and the last one (discussion and conclusion) discuss our results and report our conclusions, summarizing contributions and limitations of the research, together with possible future research directions.

LITERATURE REVIEW

Cyber Resilience

In the interconnected digital world, cyber resilience is a condition for continuous existence and competitive advantage, so the trend is towards adopting resilient strategies and operations to ensure success when operating in a hyper-connected system. Attention paid by companies to resilience is not only vital for the sustainability and growth of their business models but also a source of competitive advantage. In order to survive and guarantee longevity, companies must be constantly evolving, being able to adapt quickly, reacting to the ever-changing environment and, above all, they should have the ability to recover quickly from unforeseen events. Resilience of cyber-systems or cyber-physical systems is essential to maintain trust and privacy by mitigating security risks, protecting customers and providing them with better and safer services. Two features belong to a resilient system: (1) robustness against the predictable attacks and (2) ability to come back to a safe state if an attack has been successful, without regressing (Rogers *et al.*, 2016). This is the ability of the system to adapt, to reduce the probability of having to deal with sudden disturbances and having to resist their spread, maintaining control over structures and functions, and being ready to recover and respond with immediate and effective reactive plans to overwhelm the disturbance and restore to a robust state of operations.
Roege *et al.* (2017) propose a distinction between safety and resilience: the first one focusses on the protection of systems from threats or events, while the second one is the ability to prepare and adapt to changing conditions, to resist and to recover quickly from interruptions. In detail, resilience includes the ability to resist and recover from intentional attacks, accidents, threats and natural disasters.

Models for Cyber Resilience

Cyber resilience requires the integration of the risk management process and the resilience management process (Haimes, 2009; Collier *et al.*, 2014). Risk management selects and prioritizes potential measures to prevent or mitigate impacts, whereas resilience seeks to improve the system's inherent ability to respond to inevitable changes, both long and short, thus adopting a time perspective. “Improving a system’s resilience offers significant advantages in managing risk; improving the resilience of a system constitutes an integral part of the risk management process“ (Haimes, 2009: 500). “Successful implementation of these mechanisms will depend on the integration of risk- and resilience-based management in an adaptive framework” (Collier *et al*., 2014: 70). Indeed, a vision based on risk management focuses on prevention or protection against intrusions (avoiding risk) or on mitigating the negative consequences of an event. Differently, an approach based on resilience is concerned with ensuring the continuity of functions and critical services.

Several authors have proposed over the years methods for creating resilient systems for cyber security using a risk management process or a resilient management process.

Risk Management Process for Cyber Resilience

Katsumata *et al.* (2010) propose a proactive approach that builds protective and resilient mechanisms to provide greater assurance to the organizational information system. This approach is called Cyber Security Risk Management (CSRM). CSRM starts with the NIST methodology, which includes three processes: risk assessment, risk mitigation and monitoring/control, extending to a four-step process that includes an initial phase of risk management planning. The evaluation process establishes the risk baseline; the mitigation process uses a cost-benefit analysis to evaluate potential additional countermeasures that reduce risk; the CSRM risk monitoring and control process provides a method of "monitoring of identified risks, monitoring of residual risks, identification of new risks, execution of risk mitigation plans and assessment of their effectiveness". It also includes feedback cycles to allow a continuous process of risk assessment, mitigation and monitoring.
Steen and Aven (2011) in order to enhance resilience engineering, suggest the use of a risk perspective different from the most popular one. In this case probability is replaced by uncertainty. The (A, C; U) risk perspective in fact takes in consideration event A, consequence C and uncertainty U. Collier *et al*. (2013) in their risk-based cybersecurity framework show how risk-based standards can move beyond risk assessment to create systems that are more resilient to dynamic threats. According to Jensen (2015), cyber-resilience for the efficient functioning of the maritime industry should improve the risk management approach implemented by the introduction of the following guidelines: 1) informational campaigns directed at the maritime companies in terms of the cyber-risks they face; 2) pressure from customers who are made increasingly aware of the risk to their cargo in cases where maritime companies lack cyber-defenses; and finally 3) "cyber premiums" on insurance policies that reflect the degree to which maritime companies adhere to the voluntary guidelines.

Resilience Management Process for Cyber Resilience

Burstein *et al.* (2012) propose a model called STRATUS: strategic and tactical resiliency against threats to ubiquitous systems. This model uses an overload of computational resources to diagnose an attack, to quickly switch to calculated backup actions and predict downstream events by the attackers to make the functions resilient to those attacks. It uses an ontological model containing missions, components, vulnerability, reliability levels for hardware and software resources provided with attack tests, schemes of the attack plan for multi-stage attacks, host and cluster organization to think about how physical or network proximity can contribute to the possibility of vulnerability.
Linkov *et al.* (2013) designed the resilient matrix framework. In this matrix, the column indexes are the four stage life-cycle by the US National Academy of Sciences (NAS, 2012) i.e. plan/prepare, absorb, recover and adapt and the rows represent the four domains taken from Network-Centric Operations doctrine (Alberts, 2002) i.e. physical, information, cognitive and social. Cells in this matrix provide recommendation for resilience metrics. Linkov *et al.* (2013) provided these metrics.
Ferdinand (2015) studied how to build and to maintain a state of cyber resilience in organizations. The focus is on the organizational learning of the environment both internally and externally and the ability to exploit opportunities and manage threats in four moments: monitoring, analysis, decision-making and change. A structure with dynamic capabilities must

present both reactive capacities, necessary to respond to market changes, and proactive capacities that allow foreseeing and creating market changes. The combined use of dynamic capabilities, i.e. (1) ordinary defensive capabilities, (2) dynamic resilience capabilities and (3) extraordinary capabilities leads to increasing levels of maturity in the area of computer (cyber) resilience.

DiMase *et al.* (2015) propose the cyber physical systems security framework (CPSS). According to this framework, the following ten areas of concern enable an holistic analysis of the health status of the cyber-physical security: (1) Electronic and physical security, (2) Information assurance (IA) and data security, (3) Asset management and access control, (4) Life cycle and diminishing manufacturing sources and material shortages (DMSMS), (5) Anti-counterfeit and supply chain risk management (SCRM), (6) Software assurance and application security, (7) Forensics, prognostics, and recovery plans, (8) Track and trace, (9) Anti-malicious and anti-tamper, (10) Information sharing and reporting.

Boyes (2015) details the cyber security attributes that affect cyber resilience. They are divided in three main clusters: Information quality & validity system configuration (integrity, utility and authenticity); controlling access and system operations (confidentiality and possession) and continuity of operations safety personnel & assets (availability/reliability, safety and resilience).

Davis (2015) identifies factors affecting the ability of an acquirer to protect its information using a simplified supply chain model. These factors are (1) state cybersecurity requirements to suppliers using a common framework and language, (2) integrate cybersecurity into the acquirer procurement process (3) devote resource to investigate the makeup of the supply chain (i.e. which supplier organizations make up the supply chain), (4) understand how a supplier meets the acquirer's requirements when not using a common, shared, framework, and language, (5) identify acquirer information shared between the acquirer and its direct suppliers, and acquirer information shared between direct and indirect suppliers.

Sharkov (2016) shows the approach implemented in Bulgaria to design a roadmap for the cyber-security national strategy. The national strategy for Bulgarian cyber-security focus on two main aspects: the implementation of the so called "CIA triad" i.e. a triad for information security made by confidentiality, integrity and availability and the measurement of knowledge of risks and threats. Those two aspects allow the structuring of objectives and measures at three nested levels: information security, IT security and IT resilience. Threats, objectives and measures are outlined in three categories: "Known", "Unknown knowns" and "Unknown unknowns".

Tran *et al.* (2016) designed the Cyber Resilience Recovery Model (CRRM) to combat a zero-day malware attacks. The model provides insights into the strengths and weaknesses of current recovery processes and presents possible solutions for addressing changing cybersecurity threats.

Best practices as awareness trainings or the isolation of a machine infected with zero-day malware using the full containment method formatting and re-imaging a machine infected with zero day malware are indicated.

Rohmeyer *et al.* (2017) believe that complete cyber resilience tests, based on the principles of cyber war-gaming to involve all corporate domains, have the potential to increase knowledge, improve resilience and make its improvement measurable. Coordinated game events, discussion-based seminars and conferences, and special forums to support collaborative analysis of real-world events and emerging threats are also recommended.

Carayannis *et al.* (2019) outline the anatomy an ambidextrous approach to cyber security adopting a balanced scorecard, multistage approach under a 7Ps stage gate model (Patient, Persistent, Persevering, Proactive, Predictive, Preventive, and Preemptive). Such an approach "emphasizes the need to enable a complex, nonlinear, adaptive process of dynamic intangible

organizational assets, resources, and capabilities across a performance frontier” (p. 1) in order to enhance cyber resilience.
Caron (2019) highlights the potential of cyber-testing techniques in assessing the effectiveness of cyber-security controls and obtaining audit evidence.

Table 1: literature review

Author (year)	**Objective(s)**	**Approach**	**Output**
Katsumata *et al*. (2010)	To propose a proactive approach that builds protective and resilient mechanisms to provide greater assurance to the organizational information system.	Risk Management	Cyber Security Risk Management (CSRM).
Steen and Aven (2011)	To understand how risk can be assessed and treated to enhance resilience engineering.	Risk Management	The basic ideas of resilience engineering can be supported by the (A, C, U) risk perspective.
Collier *et al.* (2014)	Showing how risk-based standards can move beyond risk assessment to create systems that are more resilient to dynamic threats.	Risk Management	A risk-based cybersecurity framework
Jensen (2015)	Cyber-resilience for the efficient functioning of the maritime industry.	Risk Management	Three main tools are recommended: 1) informational campaigns 2) pressure from customers 3) "cyber premiums" on insurance policies
Bodeau and Graubart (2011)	To investigate how to evolve architectures, cyber resources, and operational processes to provide cost-effective cyber resiliency.	Resilience management	Cyber Resiliency Engineering Framework.
Burstein *et al.* *(*2012)	To diagnose an attack, to quickly switch to calculated backup actions and predict downstream events by the attackers to make the functions resilient to those attacks.	Resilience management	Strategic and tactical resiliency against threats to ubiquitous systems model (STRATUS).
Linkov *et al.* (2013)	To provide recommendations for resilience metrics.	Resilience management	Resilience Matrix Framework.
Linkov *et al.* (2013)	To provide the metrics recommended by Linkov *et al.* (2013).	Resilience management	Resilience Metrics.
Ferdinand (2015)	How to build and to maintain a state of cyber resilience in organizations.	Resilience management	The combined use of dynamic capabilities leads to increasing levels of maturity in the area of computer (cyber) resilience.
DiMase *et al.* (2015)	To provide a holistic analysis of the health status of the cyber-physical security.	Resilience management	Cyber Physical Systems Security Framework (CPSS).
Boyes (2015)	To detail the cyber security attributes that affect cyber resilience.	Resilience management	Three main clusters affecting cyber resilience are identified.
Davis (2015)	Building cyber-resilience into supply chains.	Resilience management	He identifies factors affecting the ability of an acquirer to protect its information using a simplified supply chain model.
Sharkov (2016)	To design a roadmap for the cyber-security national strategy.	Resilience management	The implementation of the so called "CIA triad" the measurement of knowledge of risks and threats: "Known", "Unknown knowns" and "Unknown unknowns".
Tran *et al.* (2016)	To combat a zero-day malware attacks.	Resilience management	Cyber Resilience Recovery Model (CRRM).
Rohmeyer *et al.* (2017)	To test Capability Effectiveness for Architectural Resiliency in Financial Systems.	Resilience management	Complete cyber resilience tests, based on the principles of cyber war-gaming to involve all corporate domains, have the

			potential to increase knowledge, improve resilience and make its improvement measurable.
Carayannis *et al.* (2019)	To outline the anatomy an ambidextrous approach to cyber security.	Resilience management	Multistage approach under a 7Ps stage gate model.
Caron (2019)	To highlight the potential of cyber-testing techniques in assessing the effectiveness of cyber-security controls and obtaining audit evidence.	Resilience management	Cyber-testing techniques are providing insight in the effectiveness of the actual implementation of cyber-security control and concise input for cyber-risk management and improvement recommendations.

Based on the theoretical background, the risk management approach resulted to be replaced by the adoption of a resilient management one in the most recent years. The second approach seems to be almost the evolution of the first one in order to enhance cyber resilience.

Research question

From the analysis of literature, some research gaps emerged. There are not yet enough studies focused on how to introduce and implement a cyber systems rather than specific practices and metrics to use. This limitation justifies our research that aims to provide a managerial approach in the implementation of a cyber resilient system. For this reason, our research question is:

How do companies effectively introduce and implement systems to enhance cyber security?

The following section describe in detail the methodology applied in order to answer this question.

METHODS

Research protocol

To answer the research question, we adopted the case study methodology to take into account the complex interrelation of variables characterizing the phenomenon of cyber security and its interrelationships with resilience. The case study is the preferred method to investigate an empirical topic by following a set of pre-specified rules and procedures; it allows a holistic and contextualized analysis, properly suited for exploratory research purposes, because it allows the identification of crucial variables while exploring a given phenomenon. In particular, this research employed a multiple case study design, because "it allows both an in-depth examination of each case and the identification of contingent variables that distinguish each case from the other" (Yin, 1984; Eisenhardt, 1989). Moreover, the essence of a case study is that it tries to illuminate a decision or a set of decisions: why they were taken, how they were implemented, and with what result (Schramm, 1971).
Even if cyber security can be considered as a (relatively) new topic, the existing theoretical background can be considered as a solid base upon which build and organize the research (Yin, 1984; Huberman and Miles, 2002). Given the above considerations, for this research there was no need of exploratory efforts with a grounded theory approach, and case studies have been designed as an inductive theory building effort, according to findings emerging from previous literature.

The approach of theory building is suitable to identify and describe key variables, linkages between them, and reasons behind these correlations.
Case studies have been designed in order to ensure a high quality of research and methodological rigor. In order to reduce possible biases, and to maximize validity and reliability, objective of the research was to diversify data collection methods and sources (Yin, 1984; Eisenhardt, 1989; Patton, 2002). Documentation included mainly primary data from semi-structured interviews, and the process of data collection for this first step of research required about 3 months.
The study involved three different cases, operating in different industries and involving different types of organizations. In this way, it has been possible to study different realities inside the same phenomenon. The chosen approach has been that of theoretical sampling (Eisenhardt, 1989; Eisenhardt and Graebner, 2007).
For each of the involved firms, we carried on a semi-structured interview with a high-level manager working in the field of IT and dealing with cyber security issues: the first part of the interview protocol has been focused on classifying firms and gathering more general data about the approach in building a cyber security system; the following sections of the interview protocol have been structured to investigate (1) *Organizational structure for cyber security, (2) Investments in cyber security (3) Risk management process for cyber security, (4) Risk analysis approach, (5) Definition of cyber resilience (6) Reactive/proactive approach against cyber attacks (7) How does the company introduce and implement cyber security systems.*
As already detailed, we selected the three cases to maximize the degree of information that could be extracted, privileging an all-encompassing view on the overall (cyber) risk management process. Given this aim, we selected three different cases that are: an international consultancy company (labelled as Case_1), a military center for cybernetic operations (Case_2) and a large Italian bank (Case_3).

Table 2: sample of cases selected

	Case_1	Case_2	Case_3
Country	U.K.	Italy	Italy
Industry	Consultancy	Military	Banking
Experience in Cyber Security	10 years	1 year	8 years
Number of employees[a]	200.000	300	89.000
Number of employees involved in Cyber Security	More than 100	More than 200	150
Cyber Security investment	1 mln €	NA	10 mln €
Key Informant	Cyber Security Manager	Military manager	Cyber Senior Analyst

a. Numbers of employees approximated

Case analysis was conducted following recommendations of Yin (1984), Eisenhardt (1989), McCutcheon and Meredith (1993), and Miles and Huberman (1994). Within-case analysis was conducted according to coding techniques (adapted from Strauss, 1987; Strauss and Corbin, 1990). Cross-case analysis was conducted according to Eisenhardt (1989), Runkel (1990), and Yin (1984), seeking matches, similarities, differences, and crossing variables, among cases, to maximize validity and generalizability of the study.

Criteria of internal validity, construct validity, external validity, and reliability were taken into account, according to Cook and Campbell (1979), Yin (1984), Eisenhardt and Graeber (2007), and are schematized in Table 2.

Table 3: validity and reliability of case studies

Characteristics	Criteria
Construct validity	• Review of transcriptions by co-authors and key informants • Explanation of data analysis procedures
Internal validity	• Research framework derived from literature
External validity	• Cross-case analysis (multiple case studies) • Rationale for cases selection: theoretical sampling
Reliability	• Adoption of an interview protocol • Development of a case study database

We performed a within-case analysis in order to test and verify the relevance of factors identified through literature analysis to understand their influence and capability of explaining cyber security's characteristics and the effects in enhancing organizational resilience.

RESULTS

The case studies enabled us to examine data at the micro level. Through these three cases we had the opportunity to analyzed data of real situations providing insights into the detailed opinion of the subjects of interest.
The different sectors analyzed allowed as to be more confident about the perceptions and recommendations common to the three interviewees.

The Multinational Consultancy Company and Cyber Security: Case 1

The first organization analyzed is a service company, a world leader in the strategic, legal and financial consulting market, where one of the cyber security managers of the Italy area was interviewed.
The interviewee was enrolled previously in a small company that dealt with facility management, holding the role of DB Administrator Junior, then in a large multinational company where he was able to approach the world of cyber security.

<u>Organizational structure for cyber security</u>

In this company, only in Italy, more than 100 employees are employed in cyber security. It is necessary, however, to make a distinction: on the one hand, there is a portion of the company that provides services to other companies and includes a hundred people who work with the interviewee; on the other hand, there is a team that deals with security for the company itself. In the second case, it is not merely an IT function but a security function, well structured, composed of a Cyber Information Security Office (CISO) that deals with the security of internal projects and with teams that are not very large, consisting in about 20 people.

Investments in cyber security

Within these three large areas, people are always the weak link in the cyber chain, and therefore there is a great need for investment in training and social engineering activities which, in the field of information security, is the study of a person's individual behavior to steal useful information. In order to strengthen the procedures and policies, on the other hand, we invest in writing rules and identifying rules to follow. And finally, the major investment in technology is the purchase of cutting-edge tools.
This company is ISO 27001 certified, which is the international standard in Europe for information security management.

Risk management process for cyber security

Among the six phases in which a risk management process is conventionally divided, the interviewee sustains that the company maintains missing a very important step: the first thing to do is the definition of the perimeter.

Risk analysis approach

Companies are large and this makes the risk analysis development difficult. Being unable to consider everything, the company considers for the risk analysis the processes that lead to a revenue, or the ones having a strong impact on the budget.
According to the interviewee, one of the most frequent errors is to look at security as a cost because it requires spending a lot of money without, however, any monetary returns.
Each year the Case_1 company invests about one million euros for the prevention of risks related to information security.

Definition of cyber resilience
An interesting observation expounded by the expert lies in the interpretation of resilience as the ability to not break down and therefore avoid economic losses, he talks about avoided damage rather than money earned.
The interviewee states that resilience does not consist in knowing how to protect oneself but in being able to deal positively with what happens. Do not get upset but know exactly what is lost and what still persists.

Reactive/proactive approach against cyber attack

Regarding the choice between a reactive and proactive approach against cyber-attacks, the interviewee claims that it is increasingly necessary to adopt a proactive approach.

The way the company introduces and implements cyber security systems

Recently, a series of initiatives gathered under the name of threat intelligence (which are going to mix with artificial intelligence) are taking place more and more in the cyber field. These are activities based on the concept of neural systems or systems that "learn". There are threat intelligence services and activities that help identifying new emerging threats day by day. These services are based on feed mechanisms that send a data stream and anticipate threats, starting from this intelligence analysis.

For some activities such as Vulnerability Assessment and Penetration Test, it is important to consult the CVSS catalogue (Common Vulnerability Scoring System) and to check if these vulnerabilities are exploitable i.e. if a hacker can pierce and enter the system via vulnerabilities. According to the interviewee, when cyber security attacks occur, it’s important that the critical infrastructure operators learn from the calamity and undertake actions in order to avoid similar situations.
Although intrusion detection systems and monitoring tools play a significant role in network security, among all the resources, the key resource is represented by human resources. The people who work within the company play the most delicate role because, having access to company data and privacy, they have the greatest responsibility. For this reason, a lot is invested in HR, allocating compulsory, online and frontal training paths with a teacher. Some of these are courses on specific topics, such as ad hoc courses on data protection. These provide information to employees about online security and social engineering, when users are connected to their workstations and access websites that could lead to potential cyber-attacks. Training programs are very often able to monitor users' progress.
The training is then tested with social engineering activities, that is, people are exposed to real receipts of false e-mails, attacks, and other simulations of attempts to extort sensitive data. These tests are done because often the possibility that, especially about online courses, people follow directions in a distracted way is underestimated. Of course, training courses will not avoid human errors but, if done correctly, they can increase awareness.
At the end of the interview, the manager was asked if he considered the company in which he works to be at the forefront, and the response was positive. In fact, they have four security operation center (SOC).

Future directions in cyber security

He also said that two paths are opening for the future in cyber: artificial intelligence and insurance.
About artificial intelligence the manager believes that we are going more and more towards a world in which man will give instructions and machines will learn. The new cyber trend will see hackers no longer attempting to puncture the system but badly instructing machines in doing it.
About insurance, the interviewee sustains that it is difficult to mitigate all the risks, so he highly recommends the development and the adoption of cyber risk insurances.

The Public Military Organization and Cyber Security: Case 2

The second interview was submitted to a cyber security manager operating in a military industry. His command deals with the protection of military networks and military information in Italy and abroad, where they have networks connected or are conducting some military actions.

Organizational structure for Cyber Security

The organization chart includes a command and staff area and an operational department, a training area and an experimental area. The mission of the Case_2 organization is cyber-defense and cyber-network-defense. Cyber defense is a static defense of the network, using monitoring systems that allow the company to observe the integrity and availability of data and networks on which they work. Cyber-network-defense, on the other hand, also includes the ability to perform Vulnerability Assessment (VA) Penetration Test (PT), i.e. the continuous verification of network efficiency and the search for vulnerabilities. The interviewer is a manager who has been working in the cyber defense sector for more than four years. He takes care of the organization, the management of human resources, the infostructure and infrastructure sector and takes care of all the educational and training activities of the personnel, which is around a few hundred people.

Investments in Cyber Security

The cyber environment has exploded over the past decades. In the last five years there has been a strong acceleration in that, above all the various european organizations, it has been ordered to adapt to those that are the dictates of the NIST and GDPR directives, regarding the protection of privacy. According to the key informant, creating a link between the academic world, the industrial world and other administrations, it is the best recipe for responding to cyber-attacks.

Risk Management Process for Cyber Security

Within a cyber risk management process, the six steps conventionally recognized in a risk management cycle are executed. It is an iterative process that, at the end of the risk mapping and after designing an action framework, dynamically restarted; creating a cycle that keeps threats under control. Each stage is assigned to a manager who, coordinating with the other sector managers, constantly monitors, possibly with the help of IT tools.
Through the analysis, the main answer to be obtained is the list of the priorities of the risks to be faced, with the related established response plans. This is because the attack surface is so wide that one cannot afford, not even as a nation, to establish countermeasures for any type of threat, as this would imply particularly high costs. Therefore, the first thing to do is to define which are the potential attacks for which it is valuable to take preventive measures and therefore, to draw a perimeter within which only the risks with greater priority assigned will fall.
The Case_2 organization has set up the main figures who must comply with what their duties and role require. Certainly, the problem arises in the case of small and medium-sized enterprises, when these do not have enough staff to dedicate to this purpose. In these cases, it is therefore necessary to proceed with an ad hoc investment program. Investing in cyber security, however, as well as being risky, often turns out to be very expensive.

Risk Analysis Approach

The best trade-off between costs, benefits and risks is for an economic approach to choices and decision-making activities. So, with protocols containing statistics, the real risks and the percentage of decisions to be taken must be addressed. The interviewee suggests that the salient point of each organization's policy should be to establish, even in agreement with the trade unions, clear rules that do not appear as simple regulatory dictates or prescriptions but rather are shared with employees. He claim that there is a need for creating a knowledge culture in which everyone must know why each policy has a reason to exist, so if the passwords must be of a certain type, if access via USB is not authorized, if it cannot be accessed during the weekend or during night hours to external systems from outside, the employee must be aware of the reason. It is mostly about establishing a continuous dialogue and a full sharing of objectives with the company.

Definition of Cyber Resilience

Resilience is a very structural and systemic concept, according to the interviewee. Even considering resilience as the ability to resume operations following a damage suffered, it must be considered that resilience can assume political, infrastructural, physical, logical and individual character. From the political point of view, resilience is the ability of the Nation to put in place policies that then allow all systems to resume functioning in unanimity. From an infrastructural point of view, system redundancy, whether infra-structural or info-structural, must exist.
We must not forget that internet was born in the ARPA (Advanced Research Projects Agency) as the ability to allow the various nodes to communicate in cases of nuclear war, to get information from EBI (European Bioinformatics Institute) ensuring the possibility of taking different paths. So, while the physical resilience is given by the redundancy of the systems, alternative A, B or C plans constitute the logical resilience since, given a failure or an inefficiency, we can still connect the various nodes though them. In this case, a great contribution could be made from the world of modeling and simulation, with those algorithms and data mining that allow us to automatically extract useful information from huge amounts of data and make them understand the correlations. Finally, resilience must be considered from an individual point of view. Resilience takes the cultural character of people and how the nation faces the issue of system security, so a large component is precisely the training of personnel, who must understand what their role is in the process of building computer resilience, in which man is certainly the weakest link. People must first be educated to protect themselves more, and this practically means being less superficial in generating passwords and not releasing, for example, their email to anyone. Resilience obviously also depends on the amount of investments that are made. Therefore, since cyber security is a "problem" of the Nation and many contributing actors are involved, procurement initiatives are necessary, thus acquiring technologies and material capacities, which are above all characterized by interoperability, that allow the various ministries to communicate. Surely all the aspects of anticipation, recovery, resistance, evolution, are necessary to build a resilient environment for cyber-attacks but we must start to build systems that are not only commercial off the shelf, that is, solutions that are given by the industrial world, and sometimes predominantly foreign, of which the main components are not known, but above all a system of national resilience must be structured that goes to look for those products that allow the Nation to have a precise control of the industrial solutions at the base.

Reactive/proactive approach against cyber attack

In order to be resilient, we cannot privilege a reactive approach or merely proactive approach, but it is above all the union of the two, a third approach that is the predictive prognostic approach, to ensure that there are capacities for resilience. The interviewee explains that he calls it “prognostic approach” because through the correlations of data it is possible to anticipate the arrival of a catastrophic event. Speaking of a reactive approach, the interviewee meant an attitude of the Nation that is ready to react in a very short time thanks to communication, to info - sharing and to industrial capabilities. However, the key informant sustains that Italy does not yet invest enough in a proactive component.

The way the company introduces and implements cyber security systems

According to the manager, the organization Case_2 focus on the artificial intelligence sector and should do it more allowing the systems to react automatically. With regards to monitoring technologies, the most advanced ones now rely on SIEM (System Incident Event Management) systems of all types, which can create the correlation between data and ranging from IBM to the

Firewall. The support of the human being is fundamental, that is, there must be a class of operators that really has awareness of the systems.
Without any doubt, the interviewee thinks the main factor is the "human" component. For this reason the activities on which Case_2 company invests the most are training, recruitment and above all development of retention that is the ability to retain people who bring value to the organization. The interviewee suggests that the education of the human system should start from schools. There is a need for education regarding the world of computer science starting from primary schools, in which basic concepts should be introduced with the primary aim of instilling curiosity, which is the basis of the growth and development of advanced studies. For training activities there is a need for cross-cultural systems, such as masters, training courses, and school activities. It would be a great benefit for the nation to have prepared people who can work in the various departments in a transparent manner and without additional costs.

Future directions in cyber security

The manager imagines that there will be another 5 or 10 years ahead in which it will be necessary to invest a lot. In the next five years he foresees that the major investment is dedicated to the retrofit of all ICT information systems; for the next ten years he also expects a lot to be invested in artificial intelligence and in secure by design. He concludes by saying that the serious mistakes that are still being made are above all that of not educating the national leadership and of entrusting the world of information technology only to a few very specialized young technicians, but who are unable to transfer the real criticality of this environment. So, the key points to which great importance must be attached will be the creation of a global and shared culture, the organization of events, conferences and awareness campaigns in order to educate policy makers and decision makers on this new reality that afflicts us.

The Multinational Banking services company and Cyber Security: Case 3

The third and last interview was submitted to a computer security consultant of one of the largest banking services companies in Italy (and in the world), which will be called Case_3. The interviewee is a senior analyst who has been dealing with computer security for about 7 years and for this reason he has had the opportunity to learn about many aspects of this area, including application security, or the security of online applications, infrastructure security, or security concerning computers and systems. It deals with defining what the mitigation actions downstream of a risk should be and what the costs are related to these actions.

Organizational structure for Cyber Security

CERT is the acronym of "Computer Emergency Response Team", it is a team designated to respond to security incidents in case of need. Usually, several levels of response are established: a team is in charge of the supervision, then real-time monitoring of all the corporate infrastructures, twenty-four hours a day, seven days a week and, when certain alarms occur that highlight problems, if the employees are not able to understand if it is indeed a real problem, a false positive or even a failed attempt, the CERT comes into play. Within this group there is a primary structure, the Security Operations Centre (SOC), from which monitoring and supervision activities are carried out, and a secondary structure that represents the real heart of CERT, corresponding to the second level of incident response, who is called to respond in cases of extreme necessity. Therefore, the CERT is specifically the body that oversees confirming the

occurrence of a real accident, and in this case indicates the countermeasures to be taken, or instead to ascertain whether it is an oversight of the system or a false positive. Incident management includes many aspects because, based on the type of accident, certain initiatives are taken rather than others.
Since the introduction of the GDPR (General Data Protection Regulation), the document already illustrated in the previous case study, there are certain areas and contexts in which there are constraints that require special attention. These are mandatory constraints and, if the threat is real and falls within a specific case, it is necessary to contact certain individuals who have been designated to receive the reports and evaluate with them whether to proceed with a publication for customers, in case the threat impacts on customers. The cyber environment is very large, and this means that the company uses a lot of resources in cyber security. It includes anti-fraud activities, intelligence and threat intelligence, prevention and protection teams, which serve to prevent and protect against threats in a preventive way. This constitutes a great gain for the company because, preventing threats rather than having to manage them when they are attacked, brings an advantage in terms of recovery costs. In this company, about 150 people doing cyber security are in different sites in Italy.

Investments in Cyber Security

It is very difficult to make an estimate on the volume of business because security does not consist only on the cyber one but includes a considerable part also of physical security. Investment is very high: about tens of millions of euros per year. Obviously, it depends on the type of organization, it will therefore be proportional to the business that the company provides and to the volume of users/customers that are managed: the greater the business volume of the company, the greater is the investment made to protect the business of the organization and its users.

Risk Management Process for Cyber Security

The cyber can be analyzed from different points of view: from the operational point of view, of the analyst, from that of processes and a whole part connected to governance.
Based on the cyber security that is analyzed, several risk management steps are performed. This way of acting is part of the risk management prevention phase.
There is a specific field related to cyber which is business continuity that gives a given company the possibility of being able to guarantee business continuity and, in a company like this in analysis, plays a fundamental role. In this regard, the company favors a proactive approach, despite it requires very high initial costs than the costs related to the reactive one.

Risk Analysis Approach

According to the interviewee, estimating costs related to a possible contingency is certainly the most effective way to react. For this reason, there are figures dedicated to prevention and protection that work on this aspect of proactivity: based on a problem that they have already experienced, they analyze what were the outcomes of accident to understand strengths and our weaknesses and go to study where we need to improve. After this analysis, they look for the right system to make an improvement in both qualitative and quantitative monitoring.

Definition of Cyber Resilience

For the key respondent of Case_3 company, resilience cannot be generalized, it must always be contextualized. In a cyber environment, adapting to change is very difficult because the change goes much faster than one can keep up with. The interviewee says that we need to know how to adapt to change based on the risks we face in this historic moment and defines resilience as the ability to stand up once fallen. In risk management context resilience is the intrinsic ability to change its functioning before, during and following a change or a disturbance, he says. This definition is well suited to the context under study as the cyber space is a field characterized by highly dynamic systems.

Reactive/proactive approach against cyber attack

The intent is to approach resilience in a dynamic way, but it is not always possible in the environment in which one finds oneself because sometimes there are constraints linked to certain business realities that could constitute obstacles for an incident response process, which should instead manage the occurrence of these issues. These difficulties often derive from another factor that must be considered, namely the fact that, within large companies, modifying processes is very difficult. What can be done, instead, is to intervene in operations by making changes that can be made even without modifying the process. This difference becomes important when it is not enough to change the approach to a problem, but it is necessary to change the way of approaching that problem.

The way the company introduces and implements cyber security systems

Within the IT world and, in general, companies that have many assets to monitor, protect and defend (meaning both physical and digital assets, both end points and servers, network devices such as routers, switches and all the pieces that allow the machines to communicate with each other), there are systems that collect the logs, or systems that maintain the computerized traces that these devices produce, thus allowing to store the information collected by the systems in the field, centralize them and aggregate them according to data models, to go then to model data based on the type of research you are going to do. These systems are called SIEM and today are indispensable for managing events of all kinds. They collect all the events that originate from the various devices and, starting from the evidences identified, they create an alarm system designed, developed, tested and then fielded. With this alarm system, trends are defined and analyzed so that the alarm is triggered whenever something is out of the baseline trend. The company Case_3 has an alarm that monitors whether an antivirus identifies specific files on the machines, and, in this case, an incident response process starts. The interviewee sustains that the weak link in modern computing and Cyber is man because he is the one who makes mistakes and introduces risks. In order to mitigate this aspect and train people, awareness campaigns are promoted. Often these campaigns are not very effective because of the availability of people, who should be ready to accept it. An effective way to enhance awareness consists in creating false attack vectors, simulations of the possible occurrence of certain conditions reported as anomalous. An example of simulation is the phishing campaign, that is a targeted attack aimed at a target group of people to steal data from them, and this means not just sensitive data but also personal data (name, surname, date of birth), from which it is possible to obtain the identity, or particular private data of interest for the hacker. A phishing simulation also allows to call users through customer service by pretending to be someone from the company staff and to have the information released. So, in practice, the phishing campaign consists in monitoring how many people fall into the trap to see if the awareness campaign has had an impact or not, because

what people aim to develop is the ability to recognize the possible attack vectors and have a complete picture of the company's perimeter of exposure.
The company in question, for the most part, carries out these activities internally but there are technologies that are still entrusted to third parties. There are very critical assets that must be managed internally and less sensitive assets that can be outsourced. Theoretically, if you have a working SIEM you can entrust the management of assets to third parties, monitor their progress remotely and make certain risk analyzes to understand what it is before contacting the manager. After the analysis, the manager is contacted for a comparison. Many appliances and software have already begun to exploit artificial intelligence, a part of computer science that allows machines to self-learn. This part of artificial intelligence is called "machine learning" and is already implemented by the company that uses this type of system. This is a very powerful technology but also very risky because, in managing a self-learning system, the biggest difficulty is being able to make it learn well and configure its learning.

Future directions in cyber security

Machine learning is certainly the future direction of this field; given that very rapidly a series of technologies are being created that can be managed by automated systems. However, the interviewee remains of the opinion that know-how and dynamism of the human beings are irreplaceable by artificial intelligence that can instead be introduced as a support, those the figure of the computer analyst has not to be eliminated. Currently, according to the interviewee's experience, the biggest mistake that can be made is cutting safety costs. Fortunately, this now happens less and less frequently because organizations have begun to understand the importance of safety, which is the pillar for the survival of any company and therefore one of the most important aspects on which to invest economic resources. A strong limitation is that security is often considered only as a cost and the real benefits are not realized. Another problem lies in the lack of sensitivity to content, both on the part of those who perform security-related activities, and on the part of those who benefit from it. It often happens that people within companies are hostile to security consultations because users may feel limited. But this is necessary to reduce their probability of making mistakes and to put them in situations of greater protection from possible attacks, preserving the reliability and integrity of the systems that the user uses. Cyber is increasingly becoming an instrument of attack and defence, a firearm that a state can focus on others and the serious thing is that the population does not yet have an awareness of the importance of data.

Findings

Table 4: Findings emerging from the multiple-case analysis

		Case_1 Consultancy	Case_2 Military	Case_3 Banking
Context	**Final Users**	External organizations	External organizations	Individuals
	Property	Private	Public	Private
	Critical structures/activities	No	Yes	Yes
	National Culture	UK and Italy	Italy	Italy
	Size (number of employees)	200.000	300.000	89.000
	Organizational Structure for Cyber Security	- Manager in charge of risk management and compliance - CISO + teams SOCS (20 people) - Risk Owner - Top Management	- Each stage is assigned to a different manager	- CERT - SOCS
	Investments amount	10 mln € per year	N.A.	10 mln € per year
Risk Management	**Cyber Risks' cost-benefit analysis method**	- Economic analysis	- Economic analysis	- Economic analysis - Qualitative analysis
	Risk Management Approach	- Proactive	- Reactive and proactive approaches	- Proactive
	Risk Management Process	- Standard steps preceded by a perimeter definition phase	- Standard steps	- Flexible and adapted to the specific area of interest
	Policy	- GDPR - ISO 27001 certification	- GDPR - NIST	- GDPR
Resilience Management	**Resilience conceptualization**	*"Ability to not break down and avoid economic losses"*	*"Resilience can assume political, infrastructural, physical, logical and individual character.* *From the political point of view, resilience is the ability to set up policies that then allow all systems to resume their unison"*	*"In Risk Management, resilience is the intrinsic ability to change its functioning before, during and following a change or disturbance"*
	Introduction and implementation of cyber resilient systems	- Artificial Intelligence - Benefit cost analysis - Penetration test - Study of historical trends - Threat Intelligence - Training - Vulnerability Assessment	- Anticipation recovery - Data mining - Penetration Test - Procurement initiatives - Recruitment and retention - SIEM system - Training - Vulnerability assessment	- Anti-fraud intelligence - Business Continuity - Clone system - Disaster Recovery - Phishing simulations - Prevention and protection - SIEM system - Study of historical trends - Threat intelligence

				activities - Training
	Future directions	- Artificial intelligence - Cyber risk insurance	- Artificial intelligence - ICT information systems - Secure by design	- Artificial Intelligence - Know-how of human beings

DISCUSSION AND CONCLUSIONS

The case analyses revealed interesting insight regarding the existence of common practices to enhance cyber resilience.

What directly emerges from the presented cases is the great importance of intangible activities. Case_1 manager mentioned “social engineering activities” as a tool for the effective human resources training, while Case_3 manager talked about “phishing simulations” used to build organizational cyber resilience.

A considerable amount of people, handling vital information and data for the organization, still ignore how easily these information and data can be stolen: Case_2 and Case_3 directly highlighted “training activities” as an important investment to build resilience from cyber attacks.

There is however the need for developing more powerful tools: all the interviewees highlighted artificial intelligence as a future direction of investment. As a matter of fact, from both Case_1 and Case_3 emerged the relevance of “threat intelligence” as a current important investment.

Nevertheless, results of the study insistently point at stressing the importance of human dimension of cyber security: all the interviewees acknowledged that major threats come from oversights and lack of information on how to effectively manage information and communication means. As a matter of fact, tools like penetration tests can provide valuable results only if paired with investments in training to prevent erroneous behaviors.

Two key characteristics of organizational resilience, namely the analysis of historical trends and speed of recovery, maintain a critical role, just like in other risk management contexts.

Case_1 (consultancy) and Case_3 (banking) both are private organizations and the key interviews acknowledged the importance of investing into social engineering activities, e.g. phishing simulations, together with the development of threat intelligence.

Case_2 and Case_3 both deals with crucial data and information (security/military and banking): this brought interviewees to recognize the importance of SIEM system, together with training activities; at the same time, they highlighted the importance of recovery as a key dimension of resilience.

Case_1 and Case_2 both share similarities in the context, because their activities and interest toward cyber security involves also third parties. They stressed the importance of perimeter definition as a key step in risk management process together with the need of investing into IT tools like penetration tests and vulnerability assessments.

Indeed, findings showed that the introduction and implementation of security systems to enhance cyber resilience doesn’t necessarily involved the development of powerful IT tools, but rather it is made through investments in knowledge and training of human resources together with the introduction of artificial intelligence.

The above consideration brings to the formulation of the following proposition:

- ❖ Investments in learning of human resources (human resources training) and of machines (artificial intelligence solutions as machine learning) are the ways organizations introduce and implement security systems that enhance cyber resilience.

This proposition provides a double contribution, for both research and practice: it represents a new insight and direction for research on managerial aspects of cyber security and expresses also a guideline for managers willing to invest in security systems in order to enhance cyber resilience.
Indeed, this study identifies an organizational factor that plays an important role in building a resilient system for cyber security: learning. There is a need for introducing the culture of cyber security and the knowledge of its main aspects rather than the unaware adoption of high-tech tools and techniques.
However, the study has some limitations. First, the set of companies studied might be too small and future research is needed with a larger sample to confirm our findings: the above proposition must to be confirmed by a quantitative study using a survey methodology.
Secondly, the theoretical framework did not consider the nature of investments in cyber security, which do not provide a direct return and can be considered as sunk costs. Future research efforts should take into account this aspect to understand the overall willingness of company to invest in cyber security, given the highlighted drawback.

ACKNOWLEDGEMENT

This work is supported by the fund “Progetto di Eccellenza” of the Department of Computer Control and Management Engineering “Antonio Ruberti” of Sapienza University of Rome that has been assigned by the Italian Ministry of Education (MIUR) for being “Department of Excellence” in advanced training programs in the field of cybersecurity.

REFERENCES

Abdullah, N. A. S., Noor, N. L. M., & Ibrahim, E. N. M. (2013). Resilient organization: Modelling the capacity for resilience. In 2013 International Conference on Research and Innovation in Information Systems (ICRIIS) (pp. 319-324). IEEE.

Borrett, M., Carter, R., & Wespi, A. (2014). How is cyber threat evolving and what do organisations need to consider?. Journal of business continuity & emergency planning, 7(2), 163-171.

Burstein, M., Goldman, R., Robertson, P., Laddaga, R., Balzer, R., Goldman, N., ... & Keller, P. (2012, September). Stratus: Strategic and tactical resiliency against threats to ubiquitous systems. In 2012 IEEE Sixth International Conference on Self-Adaptive and Self-Organizing Systems Workshops (pp. 47-54). IEEE.

Collier ZA, DiMase D, Walters S, Tehranipoor M, Lambert JH, Linkov I (2014b) Cybersecurity standards: managing risk and creating resilience. Computer 47(9):70–76

Cook, T.D., Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Randy Mc Nally, Chicago, IL.

Ferdinand, J. (2015). Building organisational cyber resilience: a strategic knowledge-based view of cyber security management. Journal of business continuity & emergency planning, 9(2), 185-195.

Eisenhardt, K.M. (1989). Building theories from case study research. Academy of Management Review, 14, 532-550.

Eisenhardt, K.M., Graebner, M.E. (2007). Theory building from cases: Opportunities and challenges. Academy of Management Journal, 50, 25-32.

Emanuel, R. N. (2017, January). Resilience and stakeholder need. In 2017 Annual Reliability and Maintainability Symposium (RAMS) (pp. 1-6). IEEE.

Hohenstein, N. O., Feisel, E., Hartmann, E., & Giunipero, L. (2015). Research on the phenomenon of supply chain resilience: a systematic review and paths for further investigation. International Journal of Physical Distribution & Logistics Management, 45(1/2), 90-117.

Huberman, M., Miles, M.B. (eds.) (2002). The Qualitative Researcher's Companion. Sage, Thousand Oaks, CA.

Kamalahmadi, M., & Parast, M. M. (2016). A review of the literature on the principles of enterprise and supply chain resilience: Major findings and directions for future research. International Journal of Production Economics, 171, 116-133.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. Risk analysis, 1(1), 11-27.

Katsumata, P., Hemenway, J., & Gavins, W. (2010, October). Cybersecurity risk management. In 2010-MILCOM 2010 MILITARY COMMUNICATIONS CONFERENCE (pp. 890-895). IEEE.

Linkov, I., Eisenberg, D. A., Bates, M. E., Chang, D., Convertino, M., Allen, J. H., ... & Seager, T. P. (2013). Measurable resilience for actionable policy. Environmental science & technology, 47 (18), 10108-10110.

Linkov, I., Eisenberg, D. A., Plourde, K., Seager, T. P., Allen, J., & Kott, A. (2013). Resilience metrics for cyber systems. Environment Systems and Decisions, 33(4), 471-476.

McCutcheon, D.M., Meredith, J.R. (1993). Conducting case study research in operations management. Journal of Operations Management, 11(3), 239-256.

Miles, M.B., Huberman, A.M. (1994). Qualitative Data Analysis: An Expanded Source. (2ed.) Sage, London, UK.

Patton, M.Q. (2002). Qualitative research and evaluation methods (3ed.) Sage, Thousand Oaks, CA.

Ribeiro, J. P., & Barbosa-Povoa, A. (2018). Supply Chain Resilience: Definitions and quantitative modelling approaches–A literature review. Computers & Industrial Engineering, 115, 109-122.

Roege, P. E., Collier, Z. A., Chevardin, V., Chouinard, P., Florin, M. V., Lambert, J. H., ... & Todorovic, B. (2017). Bridging the gap from cyber security to resilience. In Resilience and Risk (pp. 383-414). Springer, Dordrecht.

Rogers, R., Apeh, E., & Richardson, C. J. (2016, December). Resilience of the Internet of Things (IoT) from an Information Assurance (IA) perspective. In 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA) (pp. 110-115). IEEE.

Rohmeyer, P., Ben-Zvi, T., Lombardi, D., & Maltz, A. (2017, July). Capability Effectiveness Testing for Architectural Resiliency in Financial Systems. In 2017 Portland International Conference on Management of Engineering and Technology (PICMET) (pp. 1-7). IEEE.

Runkel, P. (1990). Casting nets and testing specimens: Two grand methods of psychology. Praeger, New York, NY.

Schramm W.D. (1971). How communication works. In The Process and Effects of Mass Communication. Urbana: University of Illinois Press, 4.

Sharkov, G. (2016, October). From cybersecurity to collaborative resiliency. In Proceedings of the 2016 ACM Workshop on Automated Decision Making for Active Cyber Defense (pp. 3-9). ACM.

Strauss, A.L. (1987). Qualitative Analysis for Social Scientists. Cambridge University Press, Cambridge, NY.

Strauss, A.L., Corbin, J. (1990). Basics of Qualitative Research. Grounded Theory Procedures and Techniques. Sage, Newbury Park, CA.

Tierney, K. and Bruneau, M. (2007) Conceptualizing and Measuring Resilience: A Key to Disaster Loss Reduction.TR News May-June 2007, 14-17.

Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. computers & security, 38, 97-102.

Wieland, A., & Wallenburg, C. M. (2013). The Influence of Relational Competencies on Supply Chain Resilience: A Relational View. International Journal of Physical Distribution & Logistics Management, 43(4), 300-320.

Yin R.K. (1984). Case study research: Design and methods. (2ed.) Sage, Newbury Park, CA, Applied Social Research Method Series.

DECISION SCIENCES INSTITUTE

The Critical Need for Decision Makers to Understand Cybersecurity

Stephanie Cervantes
Secure Smart Solutions
Email: stephanie@securesmartsolutions.com

Patricia A. McQuaid
Professor of Information Systems
California Polytechnic State University
San Luis Obispo, CA 93407
Email: pmcquaid@calpoly.edu

ABSTRACT

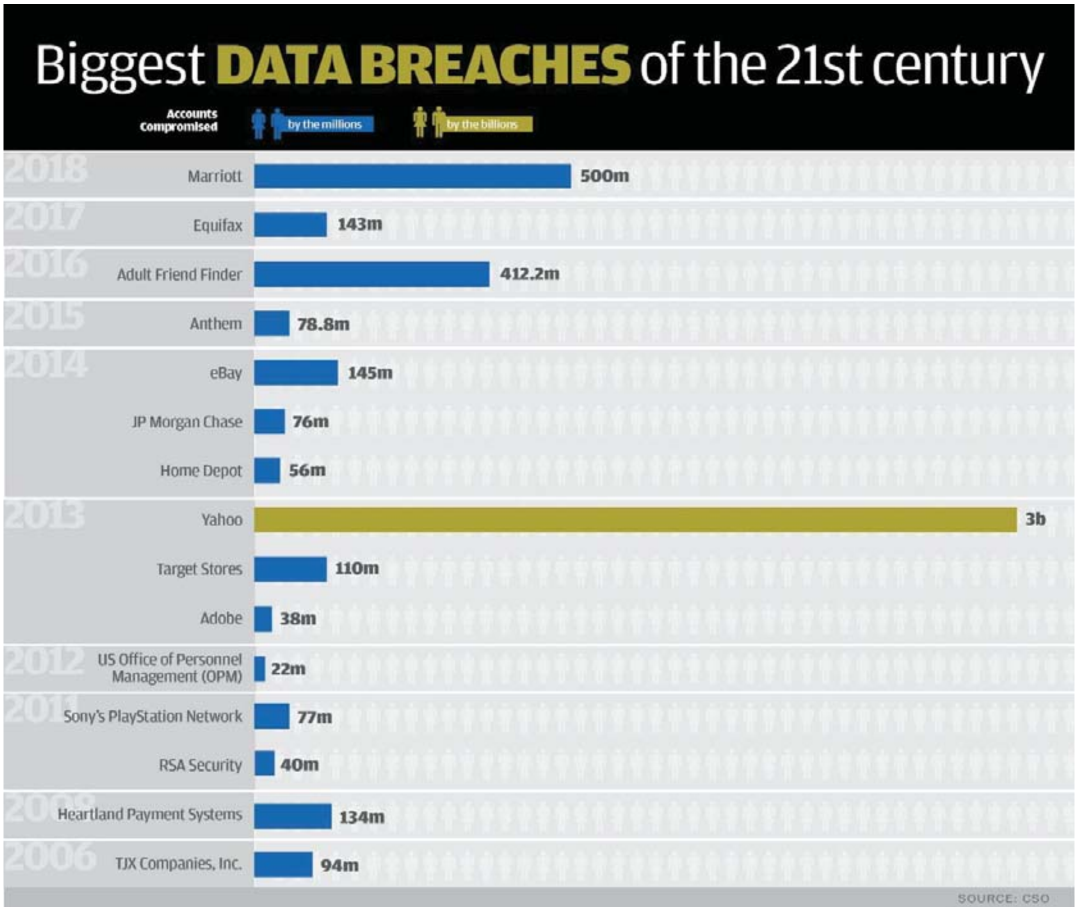
Decision Makers, aka Business Managers, control the budget without a good understanding of cybersecurity or the threats targeting their business. How can they adequately fund, staff and support a critical program without a good understanding of what the threats are and how the information they have is a desirable target for hackers and cyber criminals. This paper discusses the issues and suggests ways to address the risk and what is needed to keep up with, defend and recover from a cyber-attack.

Keywords: Cybersecurity, Decision Makers, Cyber-threats, Adversaries, Hackers, Cyber-criminals, Risk, Paradigm Change, Visibility, Detect, Contain, Remediate

INTRODUCTION

Historically, war has been waged on the ground and in the air through troops in combat. With the Internet and the growth of Internet connected devices, systems, homes, cars, etc., we have seen an increase of Nation States training cyber forces to attack other nations and individuals. It makes sense that the next world conflict or war will most likely include cyber warfare in some part, or in total. It is imperative that Decision Makers in companies or in government understand they can no longer leave everything cyber, including the understanding of cyber, to the IT or Information Security shops. It is critical that Decision Makers have a basic and beyond understanding of cybersecurity and how cyber-attacks can not only take down their network, but steal intellectual property; gain control of critical systems for ransom or nefarious purposes; be used as a launching point to attack other companies, governments or individuals, and cause serious financial and reputational loss, as well as impact the US critical infrastructure. It is not IF you will be attacked, it’s WHEN and how you quickly you detect an attack and mitigate it. Figure 1 is a compilation of some of the biggest data breaches and their reported losses, from 2006-2018 (Armerding, 2018). What they all have in common is the fact that their data was breached; what is not in common is the way they happened and the ways they could have been prevented. Much of this information is discussed in this paper.

Figure 1: Biggest Data Breaches of the 21st Century



TODAY'S THREATS

Attacks today are increasing in sophistication. We've seen them move from fairly simply scripted attacks to use of commodity malware and ransomware to custom backdoors put into networks with boot record (aka keys to the kingdom) malware that is so persistent it's difficult to remove. We've gone from the same type of malware spread throughout an organization to malware that dynamically morphs and changes as it spreads. We've gone from limited Command and Control takeovers that were typically limited to IP addresses and not Domain Name Resolution, to larger Command and Control infrastructures with more IP addresses and some use of Domain Name Resolution to taking advantage of legitimate sites for Command and Control.

We've seen disruption of services take place that has caused outages and the risk of ransomware to Internet of Things (IOT) increase. We've seen distortion of data and information happen, which has been used to spread misinformation. If you can't trust the integrity of the data because it's been compromised, or you don't know it's been compromised, what do you do?

Phishing attacks are fraudulent attempts to obtain sensitive information such as usernames, passwords and credit card details by disguising oneself as a trustworthy entity, typically in an email message. It often directs users to enter personal information at a fake website which looks very much like the legitimate site. These attacks are being used to obtain log-on information and credentials for home and office users, along with credit card information, financial information, health information, etc. (Fruhlinger, 2019). Phishing attacks are getting more sophisticated with the use of machine learning to gather information from social media, distribute fake and often times targeted messages (Goldberg, 2018).

Ransomware is still out there and being used repeatedly on those already ransomed. What many fail to realize is that once a ransom has been paid, the cyber-criminal normally has a full copy of all the information they want from your PC or network. Paying the ransom usually gets you or your company placed on a list on the Darknet of individuals and companies who have paid ransom, setting you up as a good place to be phished over and over again (Urbelis, 2017).

Cryptojacking is rising with the use of cyber-currency outside of the Darknet and criminal realms (Nadeau, 2018). Cryptojacking is a trend that involves cyber criminals hijacking home or work computers to look for, or “mine” cryptocurrency. Mining for cryptocurrency uses a lot of processing power, which is why is advantageous for the cyber-criminal to piggyback or use other systems to look for cryptocurrency. Cryptojacking runs in the background and causes serious performance issues to the infected computer or network.

Cyber-Physical attacks that use computers and their ability to interrupt or disable critical infrastructure such as cameras, traffic signals, the electrical grid, transportation, water treatment facilities, dams, and even military equipment is growing and is a new wave of warfare and terrorism. The threat to US weapon systems was brought to light in a NY Times article in October 2018 (Sanger & Broad, 2018).

THE GROWING THREATS AND WERE THEY ARE COMING FROM

The threat to IOT connected devices, medical equipment, cars (IOT connected and semi-automated), including law enforcement vehicles and medical records is growing.

Threats are coming from insiders, external partners and contractors, Internet connections, etc. That is why those in positions to make decisions, to allocate budgets and spending, need to realize how being under-protected in the cybersecurity realm can leave them wide open to attack, loss of data integrity, revenue and reputation, negligent and liable and many times out of a job. That lack of knowledge and commitment to cybersecurity can also lead to loss of life for those using a compromised or hacked device, system, weapon, vehicle, etc.

THE EXPENSE OF BEING ATTACKED

Cyber-attacks can cost involve cost of breach notification to customers; cost for credit insurance for customers; regulatory compliance (fines); public relations / crisis communications; attorney fees and litigation; cybersecurity improvements; and technical/forensic investigations, often times with law enforcement. Some of the lesser known costs can be insurance premium increases; increased cost which raises debt; operational disruption; loss of customers / strategic relationships; lost contracts resulting in lost revenue; devaluation of trade name /brand; loss of intellectual property.

UNDERSTANDING THE THREATS

The best way for those in decision making authority to understand the current cyber-threats is to become familiar with what the threats are, along with how mitigate them. The top 10 cyber-attacks are (Melnick, 2018):

1. Denial-of-Service (DoS) and distributed denial-of-service (DDoS) attacks, which include TCP Syn flood Attacks, Teardrop attack, Smurf attack, Ping of Death attack, and Botnets.
2. Man-in-the-Middle attack, which includes session hijacking, IP spoofing, and replay attacks.
3. Phishing and Spear-phishing attacks.
4. Drive-by attack
5. Password attack
6. SQL injection attack
7. Cross-site scripting (XSS) attack
8. Eavesdropping attack
9. Birthday attack
10. Malware attack

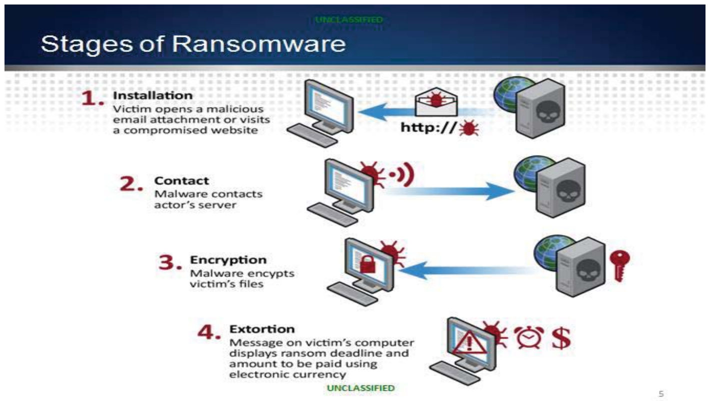
UNDERSTANDING THE TARGETS

Understanding the attacks along with who they target is important. DoS and DDoS attacks target the network and application infrastructure. Malware, zero-days and botnets target endpoint systems and servers. They try to obtain access to systems, create backdoors, and establish command and control over a large network of devices. SQL Injections and XSS attacks target databases (OWASP, 2019). Social and phishing attacks target individual users by doing a pre-attack recon eScience, building trust with fake social profiles, and then the initial infection. Password and configuration attacks target endpoint systems and servers. Their goal is initial penetration of a network or system, expansion of reach, and escalation of privileges to gain as much control possible. Smart or IOT connected device and mobile hacking targets mobile and embedded services. They are a new attack and entry point to enterprise networks. They gain access to user data through vulnerable mobile OS and applications.

Decision makers need to understand ransomware and how it works, along with the consequence of paying or not paying the ransom. Ransomware typically gets into a system or network when a user opens a malicious email, or visits an infected site, or clicks on an infected link. The malicious code is designed to crawl through the infected computer with the hope of gaining access to the network the computer is attached to. Once the ransomware has crawled the network and allowed the attacked to obtain files, confidential information, intellectual property, etc., the code executes and all or some of the files are encrypted. Then the extortion banner is displayed on the victim's computer telling them to pay the ransom in cyber currency by a certain deadline to unlock the files. Figure 2 illustrates this process.

What many do not understand is that the information has already been copied, stolen, compromised before the encryption takes place and the demand for ransom is issued. When the ransom is paid and the computer or network is unlocked, the attacker already has everything they want. Paying the ransom will unlock the data, but you can't be sure it hasn't been changed, altered or parts deleted. Additionally, paying the ransom gets the payee a place on a list for users or companies who paid ransom, setting them up to be targeted and held ransom again and again. It's better to have a solid, well tested back-up routine that you can use to recover from a ransom attack by wiping your systems and re-installing to the most recent back-up.

Figure 2: How Ransomware Works



THE ATTACKERS

Adversaries, cyber-criminals, hackers, attackers come in varying forms. There are groups with causes such as ALF, ELF, Peta, Anonymous, etc. who attack anyone they believe is a threat to, or against their cause. There are nation states such as China, Iran, North Korea, Eastern-Bloc Countries who target companies to steal intellectual property, plans (think defense contractors), weaponry or critical infrastructure systems or development, monetary targets (think banks, Fortune 100 companies, airline or just about anything you can think of. There are cyber-criminals who target individuals to steal personal information, financial credentials, and health information and credentials.

There are also the unintentional or intentional internal adversaries, your employees, their families or friends, external customers or partners that connect to your network using personal devices or using a company computer or device for personal use ignoring company policies and security.

WHAT IS A HACKER OR CYBER VILLIAN?

Hackers are a type of cyber villain who create and modify computer software and/or hardware, including computer programming, administration, and security-related items without authorization. They gain unauthorized access to computers or networks, often just for the challenge of it.

Non-Nation State Hackers or cyber-criminals typically fall into 4 categories: Black Hats, Gray Hats, White Hats and Suicide Hackers.

- Black Hats, aka Crackers, are individuals with extraordinary computing skills, who resort to malicious or destructive activities.

- White Hats are individuals professing hacking skills who use them for defensive purposes. They are also known as security analysts.
- Gray Hats are individuals who work offensively and defensively at times.
- Suicide Hackers are individuals who aim to bring down critical infrastructure for a "cause". The normally don't worry about facing jail time for their actions, believing the "cause" is greater than all.

There are also unintended cyber villains, which many call the Human Factor. Most companies are understaffed in the cybersecurity arena, which causes staff to do more with less people, causing mistakes to be made, or threats to be overlooked or missed. There is also the element that "stuff just happens". People take vacations, go to meetings, are out sick, etc. There is simply not enough time in a day for staff to do it all, which means they are not watching systems, reading logs, or correlating odd behavior on the network 24 x 7, or even 8 x 5. The world-wide shortage of cybersecurity professionals is the major reason behind this (Manpower, 2018).

ADDRESSING THE RISKS

Decision makers need to ensure the cyber risks are systematically addressed. To do this they should realize the risks. The global cost of cyber-crime in 2019 was estimated to reach $2 trillion US dollars (Juniper Research: Global Risk Reports). In 2016 the average cost of a cybersecurity breach is $4-7 million US dollars (Security Intelligence: 2016 Cost Data Breach Study) and it continues to grow in 2019. Theft of trade secrets/Intellectual Property (IPR) results in $749 billion to $2.2 trillion annually. Only 48% of breaches are caused by acts of malicious intent (IDG 2018).

Once you understand the risks and threats, you need to address them. First you must understand what you have. You cannot secure what you cannot see, or do not know about. Think about what you have in your organization and understand what it could mean to a hacker. Do you have intellectual property, confidential data, personally identifiable information, healthcare or medical information, top-secret plans or information, building and network diagrams that would provide someone to compromise you organization physically or through cyber, financial information, stock market projections or trading research or information, accounts receivables or ledger information that if changed would create financial ramifications, etc.? You should know how many ingress and egress (in and out) points there are on your network, how critical systems or information are protected and everything that resides on your network, systems and their patch levels, and who your users are. There should be a way to check the network and everything connected to it quickly with security solutions. You should be able to account for all users and ensure their access privileges are limited to need to know, and not accumulated from past positions within the organization.

DECISION MAKERS SHOULD…

It is critical to acknowledge that the cyber threat is real, all the way from the person who answers the phone to the CEO. Cybersecurity is the job of everyone who is in anyway connected.

Don't be fooled into thinking that just because you are compliant, you are secure. Compliance does not equal security. Security is a balance of control and visibility.

Penetration testing, also called pen testing or ethical hacking, “is the practice of testing a computer system, network or web application to find security vulnerabilities that an attacker could exploit. Penetration testing can be automated with software applications or performed manually. Either way, the process involves gathering information about the target before the test, identifying possible entry points, attempting to break in -- either virtually or for real -- and reporting back the findings” (TechTarget, 2018). Once a system has been updated, or a line of code changed, etc. there could be new holes opened up.

WAYS TO KEEP UP AND AHEAD OF THREATS

To stay ahead, or at least in step with attackers, you should continuously monitor your network. Often times, organizations fix high to medium vulnerabilities, leaving the low risk ones alone. Attackers are aware there is only so much time in a day, so they frequently look for low risk avenues to gain entrance. Use security solutions to do machine to machine analysis and human threat-hunters to review the data and make it contextual to your environment.

You should have solid security procedure policies, change control, full and incremental back-ups for data and critical systems that is tested regularly. You need an incident response and escalation plan that is practiced and kept up to date, along with training for all personnel.

CHANGE THE PARADIGM

Security tools by themselves are not the “magic bullet” that will solve all of your security problems. More cybersecurity professionals by themselves are not the answer either. It takes a combination of both to keep on top of cybersecurity.

Keep in mind, prevention is not enough. Tools fail or don’t see the whole picture. Individuals get re-tasked, are multi-tasking, are not reviewing logs continually, or just miss things. Relying on any of these solely can give you a false sense of security. All networks can be hacked if they are internet-facing. Visibility and quick threat detection and response is the key to preventing, minimizing and recovering from a cyber incident.

It’s time to change the paradigm! Going forward cybersecurity is all about:

- The time to detect
- The time to contain
- The time to remediate

If you don’t have the tools or the personnel to do cybersecurity, consider using a managed security service provider to fill the gaps. They normally have backgrounded staff with top level security clearances, along with the ability to monitor your network traffic round the clock. They can help the biggest security team, or the smallest fill in the gaps. The cost can be less than what you’d pay a lower-level cybersecurity analyst. A managed security service can help you identify and remediate the threats quickly, saving you time, money and reputation.

CONCLUSIONS

Regardless of the size of your company, or the industry you’re in, as a Decision Maker, it is your responsibility to ensure you have a good cybersecurity program, a secured network, and trained cybersecurity personnel. You may be the one that controls the budget, so it is imperative that you understand the problems and related risks.

REFERENCES

Armerding, T. (2018) The 18 biggest data breaches of the 21st century https://www.csoonline.com/article/2130877/the-biggest-data-breaches-of-the-21st-century.html

Fruhlinger, J. (2019) What is Phishing? How This Cyber Attack Works and How to Prevent It. https://www.csoonline.com/article/2117843/what-is-phishing-how-this-cyber-attack-works-and-how-to-prevent-it.html

Goldberg, A. (2018) How Companies Are Detecting Spear Phishing Attacks Using Machine Learning https://www.business.com/articles/machine-learning-spear-phishing/

(IDG Global State of Information Security 2018) https://www.idg.com/tools-for-marketers/2018-global-state-information-security-survey/

Juniper (2018), Cybercrime Will Cost Businesses Over $2 Trillion By 2019, https://www.juniperresearch.com/press/press-releases/cybercrime-cost-businesses-over-2trillion

Manpower Group (2018) Solving the Global Talent Shortage https://go.manpowergroup.com/talent-shortage-2018

Melnick, J (2018), Top Ten Most Common Types of Cyber Attacks (05-15-18) https://blog.netwrix.com/2018/05/15/top-10-most-common-types-of-cyber-attacks/#Denial-of-service%20(DoS)%20and%20distributed%20denial-of-service%20(DDoS)%20attacks

Nadeau (2018) What is Cryptojacking? How to Prevent, Detect, and Recover from It. https://www.csoonline.com/article/3253572/what-is-cryptojacking-how-to-prevent-detect-and-recover-from-it.html

OWASP (2018), Open Web Application Security Project, SQL Injection https://www.owasp.org/index.php/SQL_Injection

Sanger, D. & Broad, (2018), New US Weapon Systems Are Hacker's Bonanza, Investigators Find https://www.nytimes.com/2018/10/10/us/politics/hackers-pentagon-weapons-systems.html

TechTarget (2018), Pen Test (Penetration Testing) **https://searchsecurity.techtarget.com/definition/penetration-testing**

Urbelis, R. (2017) WannaCrypt Ransomware Attack Should Make Us Wanna Cry https://www.cnn.com/2017/05/14/opinions/wannacrypt-attack-should-make-us-wanna-cry-about-vulnerability-urbelis/index.html

DECISION SCIENCES INSTITUTE

A Conceptual Model of Data-Driven Decision-Making Process: A Guide for Developing an Introductory MBA Core Course in Business Analytics

Hindupur Ramakrishna
University of Redlands
Email: hindupur_ramakrishna@redlands.edu

Avijit Sarkar
University of Redlands
Email: avijit_sarkar@redlands.edu

Bindiganavale Vijayaraman
University of Akron
Email: bsv@uakron.edu

ABSTRACT

Over the last decade, many MBA programs have added Business Analytics (BA) courses into their curriculum. Unlike most other disciplinary area courses that are somewhat similar in across MBA programs, BA courses tend to vary significantly in content and structure. Thus, very little systematic guidance is available to faculty who are considering developing a BA course. This paper presents a conceptual model of the data-driven decision-making process that was the foundational guide for the methodical development of a BA course and its implementation. Some implications of this model for development of BA courses for MBA programs are also presented.

KEYWORDS: Decision-making, Business Analytics, MBA, Course Development

INTRODUCTION

One could argue that analytics owes its origin to the Hindu-Arabic numeric system (that included zero – ninth century) for data and the development of scientific method and probabilities (17th through 19th century) for modeling as presented by Buchanan and O'Connell (2006). The first well documented use of data (numeric, temporal, and geographic) for decision making is the study of the cholera epidemic by John Snow in 1854 (Tufte, 1997). With the availability of computing to add to data and modeling, business analytics has become a strategic tool for decision making and competitive advantage in many organizations since the concept was introduced in the 2000's. Davenport & Harris (2007) defined analytics as "the extensive use of data, statistical and quantitative analysis, explanatory, and predictive models, and fact-based management to drive decisions and actions." Analytics helps to discover meaningful patterns, new and novel information, insights, and knowledge, in data. Business analytics (BA) has been defined to be "a special application/subset of analytics that leverage its tools, techniques, and principles to develop solutions to ever so complex business problems" (Delen and Ram, 2018, pg. 2).

Davenport and Patil (2012) identified that a huge shortage of analytically skilled employees, such as data scientists, is a serious challenge for organizations moving into data driven decision making economy. In a report for McKinsey, Manyika et al. (2011) projected that United States alone could face shortage of 140,000 to 190,000 people with deep analytical skills and 1.5 million managers and analysts well-versed in the use of analytics to make effective decisions by 2018. Many other authors also identified huge shortages of business analysts, data engineers & data scientists.

Since BA has become so important in organizations, universities took up the challenge of developing BA knowledge and skills among their graduates by introducing degree programs in BA and courses in BA in their graduate programs and this trend has exploded over the last decade. The number of universities offering graduate programs in BA and courses related to BA in their MBA programs have increased tremendously. According to Rappa (2014), there were nearly 60 programs that were offering advanced degree in analytics in 2014.

Initially, many business schools started converting their statistics, business intelligence, and management science courses into BA courses in their MBA programs. Faculty teaching these courses were assigned to develop and teach the newly introduced BA courses. Because of the lack of clear definition of BA at that time, the BA courses were heavily dependent on the background of the faculty who were assigned to teach these courses (Gorman & Klimberg, 2014). In the beginning, faculty assigned to teach BA courses could not depend on textbooks related to BA since there were none. Some faculty used customized books by putting together articles and cases from journals, and chapters from statistics, business intelligence, and management science books. Since then, publishers have started publishing books on BA authored by some of the faculty who were teaching the BA courses. This led to a wide variety of content in BA courses depending on the background of the faculty who were developing and teaching these courses. Some of the courses had statistics orientation whereas other courses had management science, decision sciences, Management Information Systems (MIS) and business intelligence orientation depending on the courses faculty developers were previously teaching.

Thus, the content of these BA courses were not consistent and were not based on any framework or model. New textbooks that were developed by the faculty who were teaching BA courses also had different orientation namely statistics, management science, MIS, and business intelligence. Some books cover descriptive, predictive, and prescriptive analytical models, whereas other books focus on descriptive and predictive analytical models only. Only few books evolved over time to cover topics based on the data mining models such as CRISP-DM (see Wirth and Hipp, 2000), SEMMA (see SAS, 2014), and KDD (see Brachman & Anand, 1996) models which grew as industry standards for developing data mining projects that define a set of sequential steps. Some schools decided to cover all types of analytics such as descriptive, predictive, and prescriptive analytics in one course and other schools decided to offer two courses with one course covering descriptive and predictive analytics and the next advanced BA course covering prescriptive and other advanced analytic topics such as big data and text analytics.

Due to the lack of systematic guidance in developing BA courses in MBA programs, a comprehensive process model could offer a solution. This model when used would lead to more consistency in course content and it would help in developing better targeted course materials. This would also make the process of course development more explicit. Since decision making is a journey, and not just a destination, the data-driven decision-making process should focus on the entire process of defining the right business problem, identifying and collecting the right data, analyzing the data with right methods, and interpreting the results and implementing the right solution to solve the problem effectively. That is the approach taken in this paper.

This paper focuses on the development of a graduate BA course for the MBA program based on a data-driven decision-making process model developed at a U.S. university by one of the co-authors. The approach to develop and teach the first BA course in the MBA program is entirely different compared to others, whose design was either driven by the background of the course developers or by the textbook they were using. The conceptual model of data driven decision making that shaped the design of an MBA course on Business Analytics will be discussed in detail. The process that was used in developing the course, the course content along with the assignments used, and the lessons learned from implementing the model will be presented in detail in the rest of the paper. Finally, implications of using the model for developing future courses in BA will be discussed.

REVIEW OF BUSINESS ANALYTICS CURRICULUM DESIGN LITERATURE

Over the past decade, universities have been responding to an increased demand for data analysis skills put forth by employers. As a result, institutions have been building analytics education and curricula, particularly in business-focused programs. During this time, the growth in analytics and related data science masters programs has been impressive. Most of these are MS in Analytics (MSA) or MS in Business Analytics (MSBA) programs (Institute for Advanced Analytics, 2019). Due to this growth in analytics programs, academic interest in studying various aspects of analytics infusion in business education has also increased. Some prior studies have examined business analytics coursework in universities, particularly programs offered by business programs and classified them by matriculation level, subject area coverage, topic coverage, and pedagogies employed (Sircar, 2009; Phelps and Szabat, 2017). Others have viewed analytics programs through the lens of industry needs, workforce requirements, and employment opportunities (Radovilsky, Hegde, Acharya, and Uma, 2018; Mamonov, Misra, and Jain, 2015; Turel and Kapoor, 2016) and appraised the maturity of such programs relative to their potential to fulfill presumed industry needs. Detailed surveys of faculty, industry professionals, and students interested in analytics have formed the basis of successive surveys conducted at the BI Congress meetings. These surveys yielded insights about curriculum design, content coverage, and a need for interdisciplinary collaborative engagement that have been documented in the studies by Wixom et al (2011, 2014). In the remainder of this section, a handful of studies on BA curriculum research is discussed briefly.

In an early exploratory study (Sircar, 2009), business analytics was posited to lie at the intersection of management science, statistics, and information technology (IT) – all of which have been offered in traditional business curricula, both at the graduate and undergraduate levels. The study found that only 7 among top 50 business schools offered a course on business analytics at the undergraduate level, while 5 others offered related coursework in which business analytics topics were expected to be covered. Sircar (2009) acknowledged that coursework in decision sciences, statistics, and IT has traditionally always been integral to business curricula; yet, the challenge was to consolidate important topics from these three areas in a coherent package that is either a standalone course in business analytics, or a BA minor or major.

Based on a survey of information systems faculty, industry professionals, and students, Wixom et al. (2011) recommended that universities must provide a broader range of analytics skills within analytics coursework and programs. The authors specifically recommended a broad, diverse range of topics for analytics coursework that includes business domain knowledge, research methods, statistics, data management, warehousing, and mining topics, quantitative analysis and operations research (OR), and systems analysis, design, and development. Recognizing that no single department within a specific unit, for example, a business school could offer such a diverse array of topics, Wixom et al. (2011) pointed to the need to adopt a more collaborative approach of curriculum development that spans multiple departments or academic units. In a subsequent survey-based study, Wixom et al. (2014) found that an interdisciplinary approach cutting across traditional disciplinary silos involving faculty from different departments and schools was on the rise for developing, offering, and teaching analytics and related coursework. Wixom et al. (2014) further found that such an approach was increasingly commonplace in MBA programs in which the introductory analytics courses were being taught by marketing faculty. A notable contribution of this study was the nuanced recommendation that analytics curricular offerings must be distinguished based on the audience – undergraduate versus graduate, and MS versus MBA. For MBA coursework and executive education in analytics, greater emphasis on business applications of analytics at the enterprise level was recommended. Analytical skills that enable students to interpret business data and make informed decisions, along with management aspects of analytics projects were recommended to be essential for MBA analytics curricula. Viewed together, the Wixom et al. (2011, 2014) studies are instructive and provide early clues about development of a foundational curriculum model that can guide the design of a BA core course for an MBA program.

Consistent with Wixom et al. (2014), Gorman and Klimberg (2014) also noted that a greater focus on technical knowledge marks a point of distinction of MS programs in analytics compared to MBA programs, even though both seem to emphasize skills training that is applicable in the workplace. Like Wixom et al (2014), Gorman and Klimberg (2014) also delineated subject area coverage of statistics, OR, and MIS topics in MS courses in analytics and found that among surveyed programs, half of the content cover statistics topics, while the other half is largely evenly split between OR and MIS topics. Only 4 of the 32 institutions included in the Gorman and Klimberg (2014) survey offered MBA concentrations in analytics, while a vast majority offered MS in analytics degrees, limiting insights on design of MBA coursework on analytics.

Among contemporary prior research on analytics curricula, especially as it relates to the development of a conceptual model for data-driven decision-making, the recent work by Gupta, Goul, and Dinter (2015) is notable for providing the first analytics model curriculum. Based on extensive survey and review of prior literature in analytics, Gupta et al. (2015) provided a taxonomy of analytics and business intelligence topic coverage for undergraduate, MS, and MBA coursework, based on knowledge and cognitive process dimensions. Reiterating Wixom et al. (2014), Gupta et al. (2015) argued that MBA coverage of analytics should predominantly focus on strategic applications of analytical tools and technologies to solve business problems to derive competitive advantage, while MS coverage should emphasize technical dimensions such as how to build and deploy analytics applications using BI and BA tools. Gupta et al. (2015) prescribed inclusion of topics such as introductory concepts, IT frameworks and use cases for analytics, evolution of analytics in the organization, enterprise performance management, human resources for BI and analytics, business processes as they relate to analytics, enterprise application areas, predictive analytics and data mining, visualization, ethics, advanced topics, etc. in an MBA BA course.

Overall, with the exception of Gupta et al. (2015), detailed, well-researched BA curriculum development guidance especially for analytics coursework in MBA programs is virtually absent.

Much of the prior work on analytics curriculum research has focused on MS in Analytics programs, consistent with the impressive growth in such programs. Yet, despite the increasing popularity of specialized MS programs, MBA programs offer a viable option for students interested in generalist business education to be exposed to analytics. As discussed earlier, in the absence of curriculum development guidance, analytics coursework design for MBA programs usually defaults to faculty with backgrounds in quantitatively oriented disciplines such as statistics, OR, industrial engineering, mathematics, and MIS, to name a few. Naturally, the course designer's academic, scholarly, and professional backgrounds are likely to bias course design. In addition, the absence of suitable teaching resources such as textbooks and case studies exacerbate the MBA analytics course designers' problems. Faculty end up designing courses based upon available textbooks that would otherwise be more suitable for MIS, OR, or business statistics coursework, or in some cases using textbooks that have been conveniently retitled to cater to the increasing demand of analytics coursework in business schools. This is neither desirable nor optimal, especially as the demand for skilled analysts continues to remain high in industry and is in fact predicted to grow.

Acknowledging that a single MBA course is hardly sufficient to produce a skilled analytics graduate, it nonetheless lays an important foundation for an aspiring analytics professional. A curriculum design model that is built around the quartet of problem definition, evaluation of data requirements, application of analytical techniques, culminating in interpretation of results and problem solution, not only places a greater focus on business applications of analytics, as deemed appropriate in prior studies (Wixom et al, 2011, 2014; Gupta et al, 2014), it aligns analytics course design with problem-solving needs of contemporary businesses and industry at large. In addition, such a curriculum design model could guide development of appropriate teaching resources, optimal selection of course designer, and potentially pave the path for redesign of MBA coursework in MIS, quantitative methods, and strategy leading to an MBA that can produce a well-rounded analytics professional.

In addition to reviewing prior BA curriculum design literature, a limited meta-analysis of syllabi of introductory Business Analytics (BA) courses in the MBA program from 16 different business schools/universities in the US was also conducted to better understand the design of existing introductory courses on business analytics in MBA programs. All 16 business schools are accredited by AACSB. All courses were 3-hour credit courses. All courses except for one course from a university in Texas had "Business Analytics" in the title of the course. Most of these courses did not have any pre-requisites except one that listed Business Statistics as pre-requisite for the BA course. The introductory BA course is required in the MBA programs in only 5 business schools whereas the other schools use them as electives. Only three courses were mentioned to be offered online.

Except for one business school, none of the BA courses were structured based on any conceptual model as per the syllabus. One school had structured their course on the CRISP-DM, six phase model. As far as the textbook is concerned, 8 schools required text book(s), one school was found to use a course reader made available in their book store, one school provides articles available online, and two schools have optional recommended books.

A variety of analytics software tools are being used in the BA courses by the business schools in their MBA program. Some are less technical than others which requires programming skills. Six schools use Excel, followed by two schools each uses XL Miner and R. Other schools are using JMP, SAS, SAS Enterprise Miner, Weka, Power BI, Rapid Miner, Risk, and Tableau.

From the content point of view, most of the schools cover descriptive, predictive, and prescriptive analytics in their courses. Eight schools cover descriptive analytics (average – 17% of the semester with a range from 12% to 40%); all 16 schools cover predictive analytics in their BA courses (average – 48% of the semester with a range from 25% to 80%); and only 5 schools cover prescriptive analytics (average – 9% of the semester with a range from 16% to 33%). This indicates a majority of BA courses cover descriptive and predictive analytics apart from some coverage on introduction to BA and Big Data.

This limited syllabus meta-analysis confirms that the BA course in many business schools do not follow any conceptual model and are either based on the text books they are using or the faculty's perception what should be covered in the introductory BA course. This also points to a need for a model for structured course development for BA courses in MBA programs.

The next section outlines a conceptual model of data-driven decision-making process as a guide for developing an introductory MBA core course in analytics.

APPROACH OF BUSINESS ANALYTICS COURSE DEVELOPMENT

Rationale for model

When faced with the task of developing a new course, it is common for faculty to search for existing courses elsewhere or conceptual models to guide them. This is no different with a business analytics (BA) course. One of the authors was tasked with development of a new introductory BA course for the MBA program early in 2018. This section details the process used for development and implementation of the course.

From the literature review, as presented in the previous section of this paper, it is clear that a single, commonly accepted, comprehensive model to guide development of BA courses does not exist. The development of our model is an attempt to fill this gap.

Data-driven decision-making process model: development

A typical decision making model starts off with a precise definition of a problem to be solved (or a decision to be made) and ends with a solution to the problem defined. For a data driven decision-making, there are three other activities that come in the middle – data, analysis, and results. This was the basis of our preliminary model. The roots for this model can thus be traced back to Dewey's model on decision making which was further structured/detailed by Simon (Dewey, 1910; Simon, 1977). The model was further refined with concepts from a variety of models – CRISP-DM, SEMMA, and KDD models.

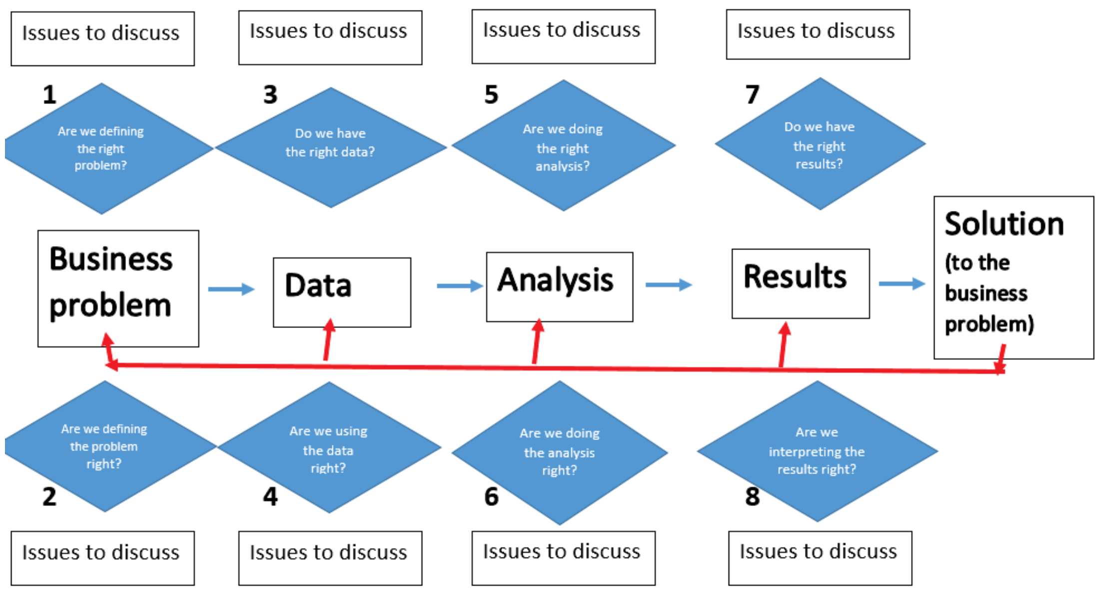
The model developed is presented in Figure 1. The model consists of the following:

- Five basic processes – defining the business problem, data, analysis, results, and solution to the stated problem
- Two sub-processes/activities for each of the first four basic processes (numbered 1 through 8 in the model) – the questions to be answered (or issues to be addressed) for each of these sub-processes are listed in Table 2.
- Connections between different processes/sub-processes.

The model thus developed maps well with Simon's decision making model. The mapping of our model with that from Simon is presented in Table 1:

Table 1: Mapping of Simon's decision making model with paper's conceptual model	
Simon's model	**Our model**
Intelligence (problem recognition/statement	Sub-processes 1 & 2 Sub-processes 3 through 6 & then 1 & 2 (Exploratory Data Analysis)
Design (generating alternative solutions)	Sub-processes 3 through 6
Choice (selecting the best solution)	Sub-processes 7 & 8 and then 1 & 2

Figure 1. Conceptual model of data driven decision-making process



The five basic processes are commonly approached in a sequential manner, in most problem solving situations. It is also Simon's argument that the three processes/phases (the intelligence, design, and choice) are followed in a sequential manner and within each phase there could still be the same three processes as each phase/process is explored. However, whenever we get stuck in any of the processes, we will most likely loopback to previous step(s) to continue the problem-solving process. For example, if we cannot get the data we need to solve the problem as stated in sub-activity 2, we may have to loop back to step 2 and redefine the problem so that appropriate data can be obtained to solve the problem. Also, in some cases, the decision making activity need not start with the first process. For example, in identifying a problem or an opportunity, we may do some exploratory data analysis (EDA) and, hence, the decision making process may start with data and proceed with the analysis processes sequentially and then loop back to problem statement. Thus, the model has the flexibility to be used in decision making contexts where exploratory or confirmatory data analysis is appropriate.

Table 2: Eight sub-processes – questions and related issues
1. Are we defining the right problem? a. How do we know we have defined the right problem? b. Who would know that?

c. Are there processes that would assure us a definition of the right problem?
2. Are we defining the problem right? a. How do we know we have defined the problem right? b. Who would know that? c. Have we clearly specified the boundaries of the problem? d. Have we understood "what is the problem" and "what is not the problem?" Have we stated them? e. Have we considered all assumptions made (whether implicit or explicit) and stated them? f. Are there processes that would assure us a definition of the problem right?
3. Do we have the right data? a. How do we know we have the right data? b. Have we defined all the variables needed? c. Have we coded the data correctly? d. Have we defined all the variables fully? e. Are there processes that would assure us a complete set of variables needed to address the problem defined? f. Are we collecting the specified data using proper methods? g. Have we used proper verification methods to assure good quality data?
4. Are we using the data right? a. Have we clearly classified the data into proper categories based on the measurement scale? (qualitative & quantitative, categorical & continuous, etc.) b. How are we handling missing values in the data set? c. How are we handling outliers? d. How are we handling incorrect data? e. Which variables will be transformed & which transformation technique is right?
5. Are we doing the right analysis? a. Do the analysis techniques match the data available? b. Would the results of the analyses address the problem defined in step 2?
6. Are we doing the analysis right? a. Have we transformed the data, if necessary, correctly? b. Have we specified the analysis fully and properly for the tool used? c. Are we using the right techniques to analyze the data? d. Are we using the intermediate results from the analysis to re-specify the model correctly for further analysis?
7. Do we have the right results? a. Have we verified the validity of the assumptions for the analysis techniques used? b. Have we asked for the right and complete outputs from the analyses?
8. Are we interpreting the results right? a. Are we interpreting the analysis results correctly? How would we assure this is happening? b. Do we need expertise from others to interpret the results?

For each of the first four basic processes, we can define two sub-activities. For example, when defining a business problem, we need to ensure that the right problem has been defined and also the problem has been defined right -- if we do not do this well, a type III error may occur, i.e., solving the wrong problem. The same structure also applies to data (having the right data versus using the data right – for example, the zip code data may be the right data, but using that data as quantitative will lead to wrong results), analysis (doing the right analysis versus doing the analysis right – doing the right analysis is a function of the data and the problem posed, but doing the analysis right is a function of the tools), and results (getting the right results versus interpreting the results right – for example, getting R^2 as a part of the regression analysis results is correct, but interpreting the value without paying attention to the "p" value would be inappropriate).

The connections between sub-activities and between some sub-activities and the first and last basic processes is also important for most problem solving situations. For example, the connection between (sub-activities) 4 and 5 is critical to understand. We may specify the data right, but fail to identify the right analysis to be done. We can see this in student work where quarter (with values assigned as 1, 2, 3, & 4) or zip codes are used in the analysis as quantitative variables.

Application of the model

Once the model is well understood and is deemed acceptable for course development, the focus will be on the following tasks:

1. Determining the depth of coverage for each aspect of the model (the processes and sub-processes) – this could differ depending on the philosophy of the designer of the course. For example, there may be courses where 10-20% the course time be spent on discussion of "data" and some other courses where the coverage may be only 5% of the course time. This also depends on the breadth and depth of students' preparation leading to this course.
2. Identifying and selecting the right course materials that map with the model – this could very well mean that a standard textbook may not be available. Using articles, from academic journals as well as from trade publications, articles that are classics (and possibly dated), and also current articles addressing relevant issues would be appropriate here.
3. Developing/selecting appropriate homework, case analyses, exams, and projects that map well with the model and with prior items 1 and 2. This could mean modifying existing materials (from textbooks) to match the model and its emphasis and also developing new materials as necessary.
4. Developing class discussion materials (like presentation materials) that map well with the model and addresses the questions raised in Table 2.
5. Choosing the right tools for analysis (like Excel, Minitab, SAS, Tableau, etc.)

The model can be used to develop any introductory BA course, in an MBA program, with different amount of focus on different aspects of analytics – descriptive, predictive, and prescriptive. We believe that the model may not change substantially in the near future, but different users could modify, add, or delete some questions raised in Table 2 to customize the model to their context and needs.

Implementation of the model

Our implementation context was a redesign of an "applied business statistics" course into an introductory analytics course in the MBA program. However, the model can be used to develop a new business analytics course in MBA programs. As the MBA program already had a "management science" course in the program, the focus of the new course was on descriptive and predictive analytics. This appears to be the case with most existing introductory business analytics courses in MBA programs as we have noted in the summary of the meta-analysis of syllabi. The designer was given complete freedom in forming the course and he exercised the freedom in making sure the content from the existing course did not have an undue influence on the new course. Though some topics overlap between the two courses, the emphasis shifted from "teaching business statistics" to "solving business problems using statistics." The new course was named "data driven decision-making" and all content (the breadth and depth) as per the model was specified before selecting the course materials (like books, articles, cases, etc.). Due to the design of the course, a traditional business statistics book or an available business analytics book was deemed to be unsuitable.

The main issue that had to be considered during the design/development of the course was the fact that many adjunct faculty would be teaching this course. Therefore, training was required to prepare them and make them comfortable teaching a radically different course compared to a traditional applied business statistics course that they were used to teaching. Also, new materials to suit the new course had to be developed – materials such as case studies, exams, projects, homework sets, etc. The materials were developed (or modified from existing materials) so that they can be mapped with the processes and sub-processes discussed in different sessions (again, with an emphasis on "solving business problems using statistical analysis" always in the forefront of discussion).

The course was offered in an eight-week format, one 4-hour session per week for eight weeks (with a stated expectation that students spend an average of 18 hours outside the classroom per week on the course). The course was offered for the first time in Nov-Dec 2018, when it was taught in both online and on-the-ground versions. The course will be taught again in May-June 2019. Some relevant parts of the syllabus developed – like course learning outcomes, reading materials, and course schedule/outline – are listed at the end of this paper.

LESSONS LEARNED AND IMPLICATIONS

Some lessons learned from the first delivery of the course – with experience of one of the authors, feedback from six adjunct faculty who taught the course and a little over 100 students are presented next.

- The model makes new course development considerably focused and straight forward.
    - It assisted the faculty involved in focused discussions on specific processes and sub-processes to be included in specific sessions. It helped faculty plan class sessions and spend appropriate amounts of time/emphasis on chosen topic areas. The model does not restrict academic freedom of faculty (in delivering the course) as long as they meet the course learning outcomes as stated.

- Students, in general, appreciate the business problem (or application) oriented focus of the course.
    - Many students specifically commented on the usefulness of the "solving business problems" focus and indicated that they have started using the concepts/tools at their place of work.
    - Students also had positive comments on the "practitioner oriented" reading materials (instead of boring academic materials).

- Do not overwhelm the students with reading materials – select/assign weekly materials that require reasonable number of hours to finish. Our students are expected to work outside the classroom an average of 18 hours per week/session. Providing "practitioner oriented" course materials, and having an emphasis on "solving business problems" makes the readings a bit more enjoyable and hence desist students from complaining about hours spent outside of course seat time per week.

- Do not overwhelm the students with too many analysis tools. We eased the students into using these tools with tutorials developed specifically for this course. Our students appreciated this extra help they received.

- Work closely with students needing extra help by providing them timely, appropriate and adequate tutorial help. As the course is fairly intense for most students, and as the course material is somewhat sequential in nature, it is important to diagnose the need for, and provide timely tutorial help to students who are having some difficulty with concepts, or tools in the course.

- A segment of students have difficulty connecting the five basic processes correctly. This was evident in the difficulty of some students with the project where they were required to pose a business problem and take it through an analytic solution. Thus, it is important to emphasize (in almost every session of the class – in our case just eight sessions) the seamless connections between the processes and sub-processes. It is best illustrated through business examples.

- As part of course assessment, a quiz as a part of the final exam was administered to all students. This consisted of six multiple choice questions – one question on problem solving process using hypothesis testing, one question on interpretation of a confidence interval estimation, and four questions of matching data with the right analysis (different types of data and different analyses). The assessment results are comparable with results from the original applied business statistics course that formed the basis for the new course. This assessment will help us with continuous improvement in the delivery of the course.
    - Anecdotal evidence volunteered by students indicate their clear happiness regarding the problem solving focus of the course and the usefulness of the course content. This is a very positive indication that the course was well received. It is very rare indeed when a student in a quantitative course voluntarily provides positive feedback!

The main implication for the course design using our model is to start with a simple design of the course (that does not overwhelm the students) and slowly add more content (i.e., readings, assignments, tools, etc.) over several deliveries of the course. In our program, this course is a pre-requisite course (and hence can be waived for students with a business degree). The diversity in the classroom is considerable with student age range of 22-55 years. This also results in considerable diversity in work experience. This will always be a challenge in course design/delivery. The implication we have discussed may be more or less severe depending on the composition of the student group.

Though the model may seem to indicate a sequential decision making process (as laid out), it is important to note there may be exceptions depending on the context. For example, in “big data” environments, the data already exists and hence the process starts with analysis (i.e., sub-process 5) and then eventually loop back to the “business problem” process. In other situations, regulatory requirements/standards, for example in healthcare environments, may constrain available data. This may again lead to starting with the “analysis” process of the model instead of the “business problem” process.

CONCLUSION

In this paper, we demonstrated how a conceptual model of the data driven decision-making process can guide in developing a graduate introductory BA course in the MBA program. This approach will help faculty in developing a BA course that is focused & consistent in course content. This well-defined data-driven decision-making process enables the faculty to break down complex problems into smaller steps that are easy for students to understand and master. This will also help students to get a clear grasp of the business problems to be solved. This can also help publishers develop BA books that are based on a standard and systematic process for BA. This should help faculty and publishers develop better course materials such as exercises, projects, cases, and discussion materials. The model presented in this paper is not rigid and provides flexibility to faculty teaching the BA course on the amount of time and coverage of topics in each phase of the data-driven decision-making process. The use of this model in refining the courses developed and in developing new courses will provide data for validation of the model/approach presented. Though the processes/sub-processes of the model developed are typically approached sequentially, it is important to note that exception to this is very much possible. An example of this could be the exploratory data analysis that could be used to identify problems. This would typically start with the “data” process and sub-processes 3 & 4, proceed with “analysis” process and sub-processes 5 & 6 and then loop back to “business problem” process.

REFERENCES

Brachman, R. J., & Anand, T. (1996). *The process of knowledge discovery in databases*. AAAI Press / The MIT Press.

Buchanan, L. & O’Connell, A. (2006). A brief history of decision making, *Harvard Business Review*, January, 32-41.

Chiang, R. H., Goes, P., & Stohr, E. A. (2012). Business intelligence and analytics education, and program development: A unique opportunity for the information systems discipline, *ACM Transactions on Management Information Systems (TMIS)*, *3*(3), 12.

Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning*, Cambridge, MA: Harvard Business Press.

Davenport, T. H. & Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century, *Harvard Business Review*, October.

Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics*, 1(1), 2-12.

Dewey, J. (1910). *How we think*. D.C. Heath & Company: New York, NY.

Gorman M. F., & Klimberg R. K. (2014). Benchmarking Academic Programs in Business Analytics, *Interfaces* 44, 329–341.

Gupta, B., Goul, M., & Dinter, B. (2015). Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School

Undergraduates, MS Graduates, and MBAs, *Communications of the Association for Information Systems*, *36*, 23.

Institute for Advanced Analytics. (2019). Graduate Degree Programs in Analytics and Data Science, Retrieved from https://analytics.ncsu.edu/?page_id=4184 on February 1.

Mamonov, S., Misra, R., & Jain, R. (2015). Business analytics in practice and in education: A competency-based perspective, *Information Systems Education Journal*, *13*(1), 4.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute, June 2011, available at https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx

Phelps, A. L., & Szabat, K. A. (2017). The current landscape of teaching analytics to business students at institutions of higher education: who is teaching what? *The American Statistician*, *71*(2), 155-161.

Radovilsky, Z., Hegde, V., Acharya, A., & Uma, U. (2018). Skills Requirements of Business Data Analytics and Data Science Jobs: A Comparative Analysis, *Journal of Supply Chain and Operations Management*, *16*(1), 82.

Rappa, M. (2014). Training Next Generation of Business Analytics Professionals, Interview with Michael Rappa, September 26, 2014, available at: https://sloanreview.mit.edu/article/training-the-next-generation-of-business-analytics-professionals/

SAS. (2014). SAS Enterprise Miner – SEMMA. SAS Institute, available online at http://www.sas.com/technologies/analytics/datamining/miner/semma.html.

Simon, H.A. (1977). *The new science of management decision*. Prentice Hall: Upper Saddle River, NJ.

Sircar, S. (2009). Business intelligence in the business curriculum, *Communications of the Association for Information Systems*, *24*(1), 17.

Tufte, E. R. (1997). *Visual and statistical thinking: Displays of evidence for making decisions*. Graphics Press, Cheshire, CT.

Turel, O., & Kapoor, B. (2016). A business analytics maturity perspective on the gap between business schools and presumed industry needs, *Communications of the Association for Information Systems*, *39*, 6.

Wirth, R. & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining, in Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining.

Wixom, B., Ariyachandra, T., Goul, M., Gray, P., Kulkarni, U. R., & Phillips-Wren, G. E. (2011). The current state of business intelligence in academia, *Communications of the Association for Information Systems*, *29*, 16.

Wixom, B., Ariyachandra, T., Douglas, D. E., Goul, M., Gupta, B., Iyer, L. S., Kulkarni, U., Mooney, J., Phillips-Wren, G. & Turetken, O. (2014). The current state of business intelligence in academia: The arrival of big data, *Communications of the Association for Information Systems*, *34*(1), 1-13.

APPENDIX

Course Learning Outcomes

Upon successful completion of this course, students will be able to:

1. Identify organizational opportunities for data-driven decision-making that create value.
2. Define each opportunity clearly and accurately, and articulate why your definition is right.
3. Specify the data needs for each opportunity identified, including its collection, and justify how that data is relevant for the context.
4. Identify and justify the right data analyses that could support decision-making in the context of the opportunity and data needs specified.
5. Perform relevant analyses using the right business analytics (BA) tools.
6. Effectively communicate the findings of the analyses performed to the decision maker(s).
7. Evaluate and articulate the value delivered by data-driven decision-making, for each opportunity identified.
8. Evaluate the ethical implications associated with BA use in organizations to assist decision making.

Reading Materials

REQUIRED TEXTS

1. Diez, David M, Barr, Christopher D, and Mine Cetinkaya-Rundel. (2017). OpenIntro Statistics (Third edition). Openintro.org. Use the following link for a free download: https://www.openintro.org/stat/textbook.php.
2. Tufte, Edward (1997). Visual and statistical thinking: Displays of evidence for making decisions. Cheshire, CT: Graphics Press. ISBN-10: 0-9613921-3-4; ISBN-13: 978-0961392130.
3. Ayers, I. (2008). Super Crunchers: Why thinking-by-numbers is the new way to be smart. New York, NY: Bantam Books. ISBN-10: 0553384732; ISBN-13: 978-0553384734
4. Davenport, Thomas H, and Jinho Kim. (2013). Keeping up with the quants. Boston, MA: HBS Publishing. ISBN-13: 978-1-4221-8725-8

REQUIRED ARTICLES

Session	Article description
1	1. Buchanan, Leigh and Andrew O'Connell (2006). A brief history of decision making. *Harvard Business Review* (January), 32-41. 2. Chottiner, Sherman, "Statistics: Toward a Kinder, Gentler Subject," *Journal of Irreproducible Results*, Vol. 35, No. 6. 3. Frick, Walter (2014). An introduction to data-driven decisions for managers who don't like math. *Harvard Business Review* (May), Accessed on 5/7/2018 from https://hbr.org/2014/05/an-introduction-to-data-driven-decisions-for-managers-who-dont-like-math 4. Hammond, John S, Keeney, Ralph L, and Howard Raiffa (2006). The hidden traps in decision making. *Harvard Business Review* (January, 118-126. 5. Liebowitz, Jay (2015). Intuition-based decision-making: The other side of analytics. *Analytics Magazine* (March/April), 38-43. 6. Lindsay, Matt (2017). The devil is in not having details, so get granular. *Analytics Magazine* (January/February), 8-12. 7. Mehrotra, Vijay (2017). Problem-solving: Keep it real with gemba. *Analytics Magazine* (May/June), 12-15. 8. Michelman, Paul (2017). When people don't trust algorithms. *MIT Sloan Management Review* (Fall), 11-13. 9. Mintzberg, Henry and Frances Westley (2001). Decision making: It's not what you think. *MIT Sloan Management Review* (Spring), 89-93. 10. Rigby, Tom (2017). Survey Sampling. *Analytics Magazine* (November/December), 44-49. 11. The Onion, "U.S. Population at 13,462," April 5, 2000, retrieved on 2/9/13 from https://politics.theonion.com/u-s-population-at-13-462-1819565581
2	12. Hymowitz, Carol, "**IN THE LEAD: Grading systems force bosses to honestly assess performance**," *The Wall Street Journal*, May 15, 2001.
3	13. Gould, Stephen Jay, "The Median Isn't the Message," *Discover*, June 1985. Retrieved on 5/7/2018 from http://www.phoenix5.org/articles/GouldMessage.html 14. Harvard Management Update (2006). Five Guidelines for Using Statistics. (May 22). Retrieved on 5/17/2018 from https://hbswk.hbs.edu/archive/five-guidelines-for-using-statistics#1 15. Berinato, Scott (2016). Visualizations That Really Work. *Harvard Business Review* (June). Retrieved on 5/7/2018 from https://hbr.org/2016/06/visualizations-that-really-work?referral=03758&cm_vc=rr_item_page.top_right 16. Davenport, Thomas D (2006). Competing on analytics. *Harvard Business Review* (January), 98-107. 17. Dhebar, Anirudh (1993). Managing the quality of quantitative analysis. *Sloan Management Review* (Winter), 69-75. 18. Duarte, Nancy (2014). The quick and dirty on data visualization. *Harvard Business Review* (April 16). Retrieved on 5/7/2018 from https://hbr.org/2014/04/the-quick-and-dirty-on-data-visualization?referral=03759&cm_vc=rr_item_page.bottom 19. Nickell, Joe Ashbrook (2002). Data mining: Welcome to Harrah's. *Business 2.0* (April), 48-54.
4	20. Paulos, John Allen, "FDA Caught Between Opposing Protesters," in *A Mathematician Reads the Newspaper*, Anchor Books, 1995. 21. Burling, Stanley, "Study Links Use of Lights in Youngsters' Rooms and Future Nearsightedness. Does Baby's Night Light Lead to Bad Eyesight?" *Philadelphia Inquirer*, May 13, 1999. 22. Denman, Chip, "Blinding Insight," *Washington Post*, May 8, 1996. 23. Gallo, Amy (2016). A refresher on statistical significance. *Harvard Business Review* (February). Retrieved on 5/7/2018 from https://hbr.org/2016/02/a-refresher-on-statistical-significance 24. Gallo, Amy (2017). A Refresher on A/B Testing. Harvard Business Review (June 28), retrieved on 5/7/2018 from https://hbr.org/2017/06/a-refresher-on-ab-testing?referral=03758&cm_vc=rr_item_page.top_right http://www.ou.edu/cls/online/lstd2323/pdfs/unit1_lamberth.pdf

	25. Kleiman, Carol, "White, male M.B.A.s found to Profit Most from Job Moves," *The Philadelphia Inquirer*, April 23, 2001. 26. Krueger, Alan B., "Better Pay for a Better College? Not Really," *NY Times*, April 27, 2000. 27. Lamberth, John, "Driving While Black," *Washington Post*, August 16, 1998, retrieved on 2/9/13.
5	28. Gallo, Amy (2015). A refresher on regression analysis. *Harvard Business Review* (November 4), Retrieved on 5/7/2018 from https://hbr.org/2015/11/a-refresher-on-regression-analysis?referral=03758&cm_vc=rr_item_page.top_right 29. Lindsay, Matt (2016). A picture is worth a thousand words. A regression is worth a few pictures. *Analytics Magazine* (July/August), 8-12. 30. Unknown (2015). Beware spurious correlations. *Harvard Business Review* (June), retrieved on 5/7/2018 from https://hbr.org/2015/06/beware-spurious-correlations?referral=03759&cm_vc=rr_item_page.bottom
6	31. Salter, Chuck (2007). She's got their number. *Fast Company* (February), 100-108.
7	32. Kauffman, Elisabeth, and Crab Orchard, "Watch for Huddling Spiders," *Time,* October 19, 1998, p. 6. 33. Samuelson, Douglas A (2011). Assessing the analysts. *Analytics Magazine* (September/October), 8-10. 34. Samuelson, Douglas A (2017). Storytelling: The write stuff. *Analytics Magazine* (May/June), 64-68.
8	35. Siegel, Eric (2013). The privacy pickle. *Analytics magazine* (November/December), 40-45.

Course Schedule (Planned)

DAY/ DATE	TOPIC	TO DO (before you attend that session)	WORK DUE
1.	Introduction to the course: Decision-making (DM), data, patterns, and solution	**Read:** **Diez et al** -- Chapter 1, **Davenport et al** – Chapter 1 **Ayers** – Introduction & Chapter 1 **Articles:** 1-11 **Tufte** – pages 5-15 **Explore:** **Think and come prepared to discuss** sampling and data collection in your organization.	
2.	Uncertainty and patterns in uncertainty	**Read:** **Diez et al** -- Sections 2.1, 2.4, 2.5, 3.1, 3.4.1, & 3.5.2) **Davenport et al** – Chapter 2 **Ayers** – Chapter 2 **Articles:** 12 **Explore:** Excel capabilities related to descriptive statistics (functions & graphs) **Think and come prepared to discuss** how uncertainty is handled in your organization.	Homework set #1
3.	Data, variables, patterns, and analytics	**Read:** **Diez et al** -- Chapter 1 **Davenport et al** – Chapter 3 **Ayers** – Chapter 3 **Articles:** 13-19 **Tufte** – pages 16-31 **Explore:** Explore descriptive analytics features of Tableau **Think and come prepared to discuss** how patterns in data are explored in your organization.	Homework set #2
4.	One variable inference	**Read:** **Diez et al** – Sections 4.1 – 4.3 **Davenport et al** – None **Ayers** – Chapter 4 **Articles:** 20-27 **Explore:** Explore features related to today's topic in Excel **Think and come prepared to discuss** how inferences using data are made in your organization.	Homework set #3
5.	Two variables inference	**Read:** **Diez et al** – Sections 5.5, 6.4, & 7.1 **Davenport et al** – Chapter 4 **Ayers** – Chapter 5 **Articles:** 28-30 **Explore:** Explore features related to today's topic in Minitab **Think and come prepared to discuss** examples of single and multiple variable(s) inferences in your organization.	Take-home mid-term exam;
6.	Multiple variables inference	**Read:** **Diez et al** – Sections 8.1, 8.2, & 8.4 **Davenport et al** – Chapter 5 **Ayers** – Chapter 6 **Articles:** 31 **Explore:** Explore features related to today's topic in Minitab **Think and come prepared to discuss** how single and multiple variable(s) inferences are made in your organization..	Homework set #4
7.	Multiple variables inference	**Read:** **Diez et al** -- None **Davenport et al** – Chapter 6 **Ayers** – Chapter 7 **Articles:** 32-34 **Explore:** Explore features related to today's topic in Minitab **Think and come prepared to discuss** analytics use in your organization.	Homework set #5
8.	Ethical aspects of data-driven DM Project presentations.	**Read:** **Diez et al** -- None **Davenport et al** – Chapter 7 **Ayers** – Chapter 8 **Articles:** 35 **Explore:** **Think and come prepared to discuss** ethical and social aspects of analytics use in your organization.	Project presentations; Final exam (in-class)

DECISION SCIENCES INSTITUTE

Redesigning an Information System Survey Course for Online MBA Students

Teuta Cata

Northern Kentucky University

Email: catat@nku.edu

ABSTRACT

Business students don't find the benefit of an information systems course in the MBA curriculum. This study uses Bloom taxonomy to redesign testing of the concepts covered in an IS survey course by shifting from Remembering to Applying level as the best testing method for adult business professionals. Learning such IS concepts in isolation from their applications in business processes was the reason for the negative perception of why an IS course should be integrated in an MBA core curriculum.

KEYWORDS: Teaching MBA Students, Teaching IS Courses, Online MBA Teaching

INTRODUCTION

It is well known in the Information Systems community that business students have difficulty on grasping the concepts of an Information Systems course for several reasons: Such courses are survey ones, which cover a broad area of IT topics without going in depth in any of them. Students are facing the challenge of remembering all those concepts without understanding the difference between them since there is no depth in the coverage of such concepts. Furthermore, those concepts are changing very fast, as this is the nature of IT, and professors need to bring to class the updates, which are not included in the textbooks. The need to stay updated in the topics is a challenge for both faculty and students that adds more pressure on preparing for class exams (Dadashzadeh, 2018).

Teaching MIS survey courses to business students should be more than just lecture-oriented classes. Faculty should consider active learning approaches where they should integrate the lectures with hands-on activities, student participation in group projects, faculty – student collaborations to explain the concepts and their implications to business practices.

The value added of IT courses to an MBA program, has been the focus of many research projects over the years. Beachboard and Aytes (2011) identified the need than an IT course in the MBA program should focus on what strength and value the IT domain add to the education of an MBA student. The authors tested the perceptions of MBA students before and after taking the MIS class as part of their curriculum. Students' perceptions before taking the class were more towards the negative side, which reflects the well-known perception of an MIS course taken by business students. The authors redesigned the IT course content to include a consulting project that allowed students to use experience-based learning as a tool to see the connection between IT and the efficiency in business operations. This course redesign made a significant change on the students' perception after the class was over. Students saw the value of IT concepts in the preparation for them to become business executives in the future.

The research on how MIS concepts should be integrated in the MBA curriculum have changed over the years, since many MBA students have some IS related background from their undergraduate studies or work related experiences. Thouin et al (2018) investigated on how an MSIS program should be designed for students who have or not IT experience, the choice of curriculum, teaching instructions, and program selections. Students who have IT background prefer a program with high level of flexibility, more integrative approaches in teaching, working in groups, having guest lectures, and courses taught by adjunct professors who work full time in the industry. They prefer a program that has a mix information on both business and technology without having a heavy focus on IT.

While research conducted by Thouin et al (2018) suggests that students learn from each other while working on group projects with their peers, may not be necessary the case for accelerated online courses. A research on accelerated online MBA courses which are primarily 5 to 8 week classes', done by Favor and Kulp (2015), show that adult learners prefer to work alone rather than with other peers in groups. Despite students' perception, that working in groups enhances their learning on the subject matter, their preference is to work alone and pace themselves. Adult learners in accelerated degree programs get frustrated by waiting on others to communicate on the projects, align performance expectations on the tasks, managing and assigning workload within groups, etc. While it is very important to have several approaches on motivating learning in an online accelerated environment, learning from each other as much as learning from the faculty, group work is not suitable for such learners.

What Teaching and Testing Methods to Use for an Online MBA?

Teaching an online-accelerated MBA class is a much different teaching experience than to a regular MIS program in a traditional setting for a full semester (16-week). Faculty who teaches to such students must adjust their teaching goals, lectures, testing materials, in order to assess correctly the knowledge obtained by such adult learners.

A quasi-experiment study for an intensive online MBA program compared two different instructional approaches related to participation on online discussions. The Socratic questions used on the experimental group to lead students to critical thinking and deeper learning improved the online learning environment for that group of students (Strang, 2011). Group comparison analysis such as Anova and Mancova identified that academic quality was much higher when the Socratic questioning conversation theory was used. It is obvious that online MBA students learn better if they challenged to speak up and participate into an online discussion initiated by faculty. The downside of such methods is that faculty need to customize and participate in the online discussions which may be a time consuming approach if they teach large sections of online MBA classes.

The principal investigator of this study has changed the peer collaboration during the semester by having online discussions where students exchange readings, videos or personal professional experience related to the IS topics covered during that week. Such online discussions are found very informative and beneficial from such adult learners in the comments they make online or at the end while they evaluate the course.

Testing of the knowledge obtained from the weekly lectures is usually done by multiple choice questions quizzes, where questions come from a test bank prepared by the publishers. Such questions are drafted as "one size fits all", without taking in the consideration how adult learners prefer to understand the material and apply it in the business setting. There is a significant disconnect on what a student wants to learn and apply in the real world from an MIS textbook and what those multiple choice question quizzes suggest what a student should remember while learning. Those questions are very isolated from a business setting, ask specific details on definitions, and hardly motivate a student to see a value in that information.

Therefore the focus of this study is to understand who are the adult learners in an accelerated online MBA program, what kind of background they have in Information Systems (undergraduate or work related ones), and what will be the best way to motivate them learn Information Systems concepts which are necessary in the education of such business professionals.

Methodology to be Used in the Study:

This experimental study will include a pre-survey; experimental treatment, and course survey at the end of the 5-week course.

1. All students enrolled in the online 5-week MBA class of Intro in Information Systems will take a short survey where they are asked to share their background which will fall in three categories: have IS background from their undergraduate studies; have IS experience from their work experience; and have no IS experience at all.
2. Students will be randomly divided in two groups in Canvas: experimental and control groups.
    a. **Experimental group:** This group of students will have a mix of MCQ and short answer questions prepared by the faculty on how the concepts of that chapter applies into business processes. The short answer questions will be out of the some of the MCQ questions pool provided to the control group and transformed to short answer ones by the faculty who teaches the class. Students will be assessed on how concepts discussed in the chapters are related to business context. Such questions will require focusing on Understanding of the concepts and Applying them in the context of Information Systems domain, which is aligned with "Apply" level of Bloom Taxonomy. The nature of short answer questions for chapter will depend on what the chapter covers. Some chapters are more technical like the database, networking technologies chapter, etc. The short answers prepared for such technical chapters will be technical scenarios and how an MBA student can make a business decision based on the technical concepts covered on the chapter. Some of the chapters covered in the course are business oriented. Students will be asked to make a decision on what will be the best software to address that business need.
    b. **Control group:** 20 MCQ provided in the test bank will be used to test the students' understanding on the chapters. Such questions are usually testing definitions or specific details of such concepts. These questions are usually very specific and isolated from the business context where such technologies are implemented. Such questions belong to "Remembering" level on Blooms taxonomy, where they basically recall, memorize and list the concepts discuss in the chapters. Control group will have 20 questions and have 20 minutes to respond to them on Canvas.
3. Course and professor evaluations will be analyzed for both groups and will be compared to see if the perception about the course really adds value to an MBA student. The course survey for both groups will have more questions which will ask if students perceive the MIS course as a useful one in the MBA core classes; if this class will help them become a more successful business leader; and what kind of IS concepts they would be interested to learn more about.

Both groups are taught the same lectures from the faculty on how to relate IS concepts to a business context.

Analysis to be Used in this Study:

T-test, descriptive analysis and maybe text mining (if enough comments are shared on course evaluation at the end of semesters) will be used to see if there is any significant difference between control and experimental groups, among students with no IS backgrounds vs the ones who have previous IS backgrounds. These tests will help to understand what IS background the current MBA students have, what kind of learners they are (based on Blooms taxonomy) and what is the best way to teach an accelerated MBA course.

Contribution of this Study:

MBA program has become one of the most well attended graduate programs. The online format gives the students the opportunity to take classes when they can, and fit those classes into their busy life. Such students need to see that an MBA degree adds value in their education, and is relevant to their current work experience. Introduction to Information Systems class has been identified as a class that doesn't align very well with their educational goals. This misalignment is not so much a mental block that an MBA student may have towards concepts in IS domain, but how such information relates into their work experiences and educational expectations. Understanding what kind of IS background an MBA student has, designing a class that is challenging and interesting, is extremely important for a faculty who designs courses which make a graduate student to see the added value of his/her graduate education.

References:

Beachboard, John; Aytes, Kregg (2011).Information Orientation, Information Technology Governance, and Information Technology Service Management: A Multi-Level Approach for Teaching the MBA Core Information Systems Course, Journal of Information Technology Education, v10, p139-160

Dadashzadeh, M. (2018). Engaging and Informing Business Students Through Group Work in the Management Information Systems Course. International Journal of Education Research. 13(1), 16-28

Judy K. Favor PhD, Amanda M. Kulp, MA, (2015). Academic Learning Teams in Accelerated Adult Programs: Online and On-Campus Students' Perceptions, *Adult Learning*, 26(4), 151-159

Kenneth David Strang, (2011) "How can discussion forum questions be effective in online MBA courses?", Campus-Wide Information Systems, Vol. 28 Issue: 2, pp.80-92, https://doi.org/10.1108/10650741111117789

Thouin, Mark F.; Hefley, William E.; Raghunathan, Srinivasan (2018), Student Attitudes toward Information Systems Graduate Program Design and Delivery, Journal of Information Systems Education, 29(1), p25-36.

DECISION SCIENCES INSTITUTE

Mastering the onboarding challenge: The case of information processing in distributed ledgers in supply chains

Dominik Roeck
University of St.Gallen, Switzerland
Email: dominik.roeck@unisg.ch

Erik Hofmann
University of St.Gallen, Switzerland
Email: erik.hofmann@unisg.ch

ABSTRACT

Distributed ledger technology (DLT) is praised to process information across organizational boundaries and thus enhance supply chain transparency (SCT). However, as first pilots show, this improvement is no sure-fire success. This case study investigates what is needed to process information with DLT to enhance SCT by studying five DLT solutions in a single industry. The cases reveal that DLT requires information processing capabilities (IPCs) in form of processes, infrastructure and IT both on an organizational and interorganizational level. Thus, in order to process information with DLT, supply chain actors need to collaborate to enable smaller organizations make use of DLT.

KEYWORDS: Supply chain management, Information processing theory, Supply chain transparency, Distributed ledger technology, Supply chain collaboration

INTRODUCTION

The importance of supply chain transparency (SCT) is increasing for both practitioners and academics. Incidents of contaminated food on a daily basis, unethical work conditions, product recalls (Saberi *et al.*, 2018) or a billion-dollar diamond fraud in India (Subramanian, 2018) are just a few examples that depict the rising importance of SCT. For example, in 2018 media reported that an Indian diamond trader was able to fraud multiple banks with forged trade documents over a period of multiple years due to a lack of SCT. The affected banks were not able to match the trade documents with the corresponding diamond. Thus, leaving the opportunity to forge documents for stones that do not exist or finance trades multiple times. Moreover, this is not the only lack of transparency in the diamond industry as proofing provenance of diamonds remains challenging. Only lately has the Czech national collection revealed that the exhibited diamonds were faked glass imitations (Malm, 2018). In addition, the problem of conflict diamonds – diamonds that are mined in war zones, smuggled and unethical-mined diamonds increase the need for more SCT in the diamond industry. Although this is only one exemplary industry, it represents typical challenges in nowadays, globally distributed supply chains. For the actors within such supply chains, the lack of transparency results in uncertainties. Due to limited transparency, the actors cannot asses the product quality (e.g. the quality of the diamond) and its origin with certainty. This exposes them to risk that affects their operations, for example in case of bad product quality and consequently their reputation, for example when trade documents are forged. This is especially a challenge for an industry such as the gemstone and diamond industry. For hundreds of years, the industry was built on trust between transaction partners. Over all these years, the industry did not underwent significant change in the behavior of trading partners. Trust was always the replacement for the lack of transparency. It appeared as if the industry had already accepted that it lacks transparency. However, the latest incidents and the customer pressure pushed gemstone and diamond actors forward in their aspiration to improve transparency sooner than later. At the same time, distributed ledger technology (DLT) has emerged with promises to address the lack of transparency and consequently the resulting uncertainties. In supply chain management (SCM), both researchers (Rao *et al.*, 2017) and practitioners (Casey and Wong, 2017) see enormous potentials due to DLTs high level of availability, immutability of stored data, the frictionless distribution of data between network members and the reduced need for intermediaries (Beck and Müller-Bloch, 2017). Based on high availability, data immutability and data distribution, researchers indicate an improvement of SCT through the usage of DLT (Saberi *et al.*, 2018). In order to reduce the uncertainty within supply chains, organizations seek to enhance SCT. However, enhancing SCT greatly depends on the processing of information in supply chain networks. This requires that DLT-enhanced SCT solutions, and the organizations that use these solutions, process information successfully. While academics and practitioners have proclaimed the potential of DLT to improve SCT, they have paid little attention to the actual processing of information of DLT to enhance SCT. In particular, research lacks the understanding of the capabilities that are required to process information in DLT-enhanced SCT solutions in order to improve SCT. This understanding is substantial for DLT adoption and consequently the potential impact that DLT will have for SCT. This includes as well the aspect how these capabilities can be built up. Moreover, in order to understand this in the context it is necessary to study, why organizations seek to process information via DLT-enhanced SCT solutions, the driver for the need to process information. This paper wishes to address this research gap as it aims to study DLT-enhanced SCT solutions and their required capabilities to process information in DLT-enhanced SCT solutions. We address this research gap by the following three research questions:

Research question 1: What are the uncertainties that drive organizations to process information via DLT-enhanced SCT solutions?
Research question 2: What capabilities are required to process information in DLT-enhanced SCT solutions?
Research question 3: How can organizations build up these required capabilities?

In order to answer these questions, we draw on information processing theory (IPT), as it presents an appropriate theoretical lens to elaborate not only the needs to process information but also the required capabilities to do so and how they can be built up. By answering these questions, the paper contributes to practice as it helps managers to understand the information processing of DLT and to identify the required capabilities. Thereby, it supports managers' decisions for or against the technology and indicate needed capabilities. In addition, it contributes to theory by positioning SCT in IPT and applying it for DLT-enhanced SCT solutions. Thereby we expand IPT to an interorganizational setting.

LITERATURE REVIEW

Transparency in physical supply chains

SCT is not only a discussed topic in practice but also in academic literature (Morgan *et al.*, 2018) and since the turn of the millennium, more and more scholars turned their attention to the phenomenon of SCT. Scholars use the terminology of SCT to summarize visibility and traceability in supply chains. Creating visibility and traceability are understood as enablers to achieve SCT (Morgan *et al.*, 2018). Supply chain visibility is defined as "the ability of supply chain partners to access information related to operations of the entire supply chain, besides the activities in which they participate" (Tsanos *et al.*, 2014, p. 436). Daugherty *et al.* (2002) define traceability as an ability to track the current or historical movement of goods within the supply chain. In addition, SCT comprises different types of SCT mechanisms that are deployed to achieve visibility and traceability. Monitoring, tracking & tracing and proofing provenance in supply chains are types of such mechanisms. They enable organizations to gather and process information in supply chains. The variety of different SCT mechanisms illustrates the multi-facetted nature of the phenomenon SCT. In line with this nature of SCT, Morgan *et al.* (2018) developed a definition of SCT based on the contributions of Carter and Easton (2011) and Carter and Rogers (2008). "SCT involves reporting to and communicating with key stakeholders to provide traceability regarding the history of the product and visibility about current activities throughout the supply chain while also incorporating stakeholder feedback for supply chain improvement" (Morgan *et al.*, 2018, p. 960).

With the emergence of sustainable supply chains, SCT moved into the spotlight. As contemporary speed of communication reveals issues in a timely manner and to a large audience, firms benefit more from reporting and engaging actively with stakeholders (Carter and Rogers, 2008). This demonstrates that enhancing transparency on the origin of products is gaining in importance (Egels-Zandén and Hansson, 2016). By providing a proof of provenance of their products as well as their components, firms seek to identify risks associated with a product's origin or the involvement of specific parties within the supply chain (Maloni and Brown, 2006). Furthermore, when looking at the diamond industry, the case of the Czech National Museum demonstrates that fake diamonds are still present in the supply chain (Malm, 2018). Moreover, the end customers are likely to be laypersons. Thus, they have little expertise to assess the originality and quality of a diamond without the help of experts. A proof of provenance enables them to combat these challenges in two ways. First, a proof of provenance discloses the geographical origin of the diamond and indicates if stones are mined in conflict areas or under unethical work conditions (Cartier *et al.*, 2018). Second, a proof of provenance provides an evidence of the originality for buyers (Saberi *et al.*, 2018). However, despite technological advancement in geochemical, isotopic or spectroscopic methods, the identification of a diamond's origin is still limited (Cartier *et al.*, 2018) and so the industry turns towards establishing a chain of custody and certificates such as the Kimberley Process Certification Scheme (Responsible Jewellery Council, 2017). Thereby, the diamond industry follows a similar approach to other industries such as food (Maloni and Brown, 2006) or textile (Egels-Zandén and Hansson, 2016) that deploy paper-based certificates and labels. Yet, the solutions in the diamond industry are vulnerable to manipulations as the paper-based certificates can be forged. At the same time, paper-based processes enable only little visibility into the chain of custody, as documents and certificates are only in possession of a single actor in the supply chain. This gives rise to financial fraud in the industry as the verification of associated trade documents remains difficult (Linde *et al.*, 2018). However, paper-based trade documents still dominate global trade, even outside of the diamond supply chain (Casey and Wong, 2017). With digitization and the emergence of DLT, the ability to have a secure and timely exchange of trade documents caught the attention of numerous supply chain actors (Blossey *et al.*, 2019). IT is then deployed to reduce the lack of transparency and thereby mitigate the risk of manipulations of trade documents (Ivanov *et al.*, 2019).

Extant literature views IT usage in general as a helpful way to make upstream processes visible to downstream buyers. This includes the visibility of information on the scope and quality of processes. Therefore, buyers use monitoring as a SCT mechanism to gain visibility into external processes. Buying firms apply IT such as RFID systems to monitor external processes and their adherence to contractual agreements such as temperature in the food industry (Abad *et al.*, 2009). Aung and Chang (2014), for instance, developed an IT system for food traceability that allows buying companies to monitor suppliers' adherence to food processing standards and quality of production processes. More recently, scholars have identified the importance of IT to enhance SCT (Zhu *et al.*, 2018). Aside from established IT systems such as Electronic Data Interchange, the emergence of DLT raises high hopes to enhance SCT. The novel technology is praised for the immutability of data within the ledger which ensures a high level of data security (Nærland *et al.*, 2017). In addition, the technology is characterized by high availability of data guaranteed by decentralization (Beck *et al.*, 2016). Moreover, frictionless information distribution within the network enhances the level of data transparency (Beck and Müller-Bloch, 2017). Data security and transparency equip the technology to be applicable in the field with a low level of trust between transaction partners (Beck *et al.*, 2016). Consequently, the call for applying DLT in supply chains enjoys growing popularity among practitioners (Casey and Wong, 2017) and academics (Rao *et al.*, 2017). The technology is seen to improve SCT as it distributes and discloses information to all participating organizations within a DLT network. The immutability and high security of the technology allows to establish a single source of truth. Thus, managers and academic scholars alike see DLT as a technology that enables major improvement for SCT (Wang *et al.*, 2018). Wang *et al.* (2019) identify that twelve out of 14 experts see DLT as option to provide improved visibility and traceability in supply chains. However, there study reveals as well, that the majority of experts see substantial challenges for DLT in issues of collaboration, information sharing, interoperability as well as data (Wang *et al.*, 2019). Despite these challenges, numerous initiatives have started to introduce DLT to enhance SCT (Blossey *et al.*, 2019). Start-ups such as Provenance.org and Everledger.io have implemented several pilots to establish a DLT-based proof of provenance, allowing supply chain actors and end customers to trace back the origin of a product. In addition, established companies within supply chains such as Maersk or Walmart have joined forces with IT specialist IBM to set up a platform to digitize and exchange trade documents. While all of these initiatives aim at providing enhanced SCT with the use of DLT, these represent the main types of functionalities that can be found in practice. Namely, they are initiatives that enable a DLT-based proof of provenance and a DLT-based trade document registry.

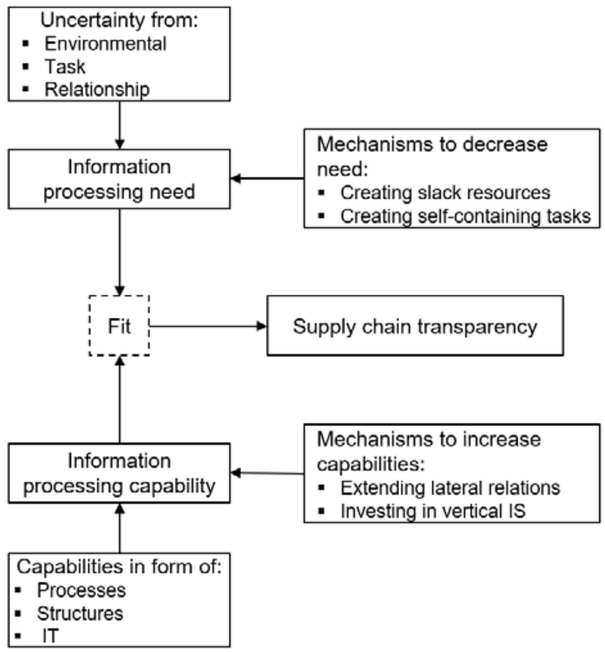
While several contributions have discussed these initiatives, the actual value of DLT for SCM and in particular for SCT and the key to overcome these challenges have not been empirically examined as of yet. Early contributions built on a conceptual approach without empirical evidence due to the novelty of the technology and so, two important aspects remain understudied. First, the adoption of the technology lacks empirical insights to guide practitioners in the technology adoption and associated decisions. Saberi *et al.* (2018) were one of the firsts to identify barriers for DLT adoption in sustainable supply chains, even though the study lacks empirical grounding. The second aspect is the value of DLT for SCM and the related phenomena such as SCT. While extant literature revealed substantial potentials, however, only a limited number of studies build their argumentation on proof of concepts (e.g. Beck and Müller-Bloch, 2017) or even pilot projects (e.g. Baruffaldi and Sternberg, 2018). Thus, more empirically grounded research is needed to guide practitioners for approaching DLT to enhance SCT. With the focus on DLT-enhanced SCT solutions, the most common application in nowadays supply chains (Wang *et al.*, 2019), the need is even larger as SCT in general is still understudied in supply chain management research (Wieland *et al.*, 2016). Nowadays, IT is gaining in importance as it supports SCT mechanisms to process information. For SCT, this processing of information is crucial, as it does not only entail the processing of information internally but also across organizations' boundaries. When analyzing the potential and challenges of novel ITs such as DLT, it is necessary to study the processing of information of these technologies. Therefore, we draw on IPT to study the information processing of DLT-enhanced SCT solutions.

Linking information processing theory to supply chain transparency

Galbraith (1974) introduced IPT, which is comprised of three key elements. These are the information processing need (IPN), the information processing capability (IPC) and the fit between both (Galbraith, 1974). As organizations are exposed to uncertainties that stem from their environment, their tasks and relationships, they have the need to process information in order to cope with uncertainty (Tushman and Nadler, 1978). Environment uncertainty are market and industry specific characteristics that arise from external factors that a company can't control (Premkumar *et al.*, 2005). Task uncertainty is related to the specific task an organization or department carries out while relationship uncertainty reflects the uncertainty that arises from the collaboration with external organizations or other internal departments that often have different goals and interests (Tushman and Nadler, 1978). In order to fulfill the need to process information, IPC in the form of processes, structures or IT are required (Bensaou and Venkatraman, 1995). Following IPT, organizations have to create a fit between IPN and IPC to cope with uncertainty and thus achieve improved performance (Tushman and Nadler, 1978). According to Galbraith (1974), organizations have two options to create the fit between IPN and IPC. They can deploy mechanisms to decrease the IPN such as creating slack or they deploy mechanisms that increase the IPC. Organizations select the mechanisms that are most effective in their setting to achieve the fit (Galbraith, 1974). While early scholars used IPT to study the design of organizations, Premkumar *et al.* (2005) expanded the scope of IPT to interorganizational studies. With their taxonomy, the authors demonstrated an approach to structure IPN and IPC in an interorganizational setting. Yet their focus is still the IPN and IPC of an individual organization. Moreover, Foerstl *et al.* (2018) study uncertainties in sustainable supply management and reveal that the fit between IPN and IPC depends upon the sustainability-related uncertainties that trigger the need. In their study, they use the buying firm as the unit of analysis, thus focusing on an individual organization again. While the mentioned studies observe supply chains, they all focus on a single organization within the supply chain. They disregard the interaction of supply chain partners and its impact on information processing within supply chains. However, SCT depends on the interaction of multiple supply chain partners (Carter and Rogers, 2008). Thus, SCT calls for information processing beyond the boundary of organizations within the supply chain (Zhu *et al.*, 2018).

Despite literatures' focus on individual organizations, IPT is well suited to be applied as a theoretical lens for SCT. Following IPT, a lack of transparency can cause uncertainty (Williams *et al.*, 2013). Thus, the IPN constitutes a need to process information in order to reduce the lack of transparency that causes the uncertainty. If mechanisms to reduce the IPN are costly such as creating slack (e.g. inventory) to deal with demand uncertainty, organizations decide to achieve the fit between IPN and IPC by deploying mechanisms to increase their IPC. Thus, they seek to enhance SCT by improving their IPC. SCT mechanisms are forms of extending lateral relations and investing in vertical information systems (IS). Figure 1 illustrates this positioning of SCT in IPT.

Figure 1: Positioning supply chain transparency in information processing theory



By this, SCT is a state that is reached by enhanced information processing which leads to improved supply chain performance. With their study, Williams *et al.* (2013) complement this notion. The authors apply IPT to model SCT by demand, supply and market visibility and operationalize performance by responsiveness. Their study reveals that internal integration is moderating the effect of SCT on responsiveness. This study demonstrates that SCT is not the sole cause for improved performance, which is taken into account in this paper. Drawing on IPT, Srinivasan and Swink (2015) explore that both demand and supply visibility are required for the development of analytics capability in an organization. Therefore, they put analytics capability in the focus. SCT is seen as a prerequisite for analytics capability. However, the authors confirm the positive effect of increased IPC for operational performance. Zhu *et al.* (2018) view analytics capabilities as enablers for SCT. The authors reveal that analytics capabilities have a positive effect on SCT. In our study, we combine the insights of Srinivasan and Swink (2015) and Zhu *et al.* (2018) and view SCT as an output of IPC and state that information is needed to fuel IPC, thus requiring visibility.

METHODOLOGY

Study design

The aim of this study is to contribute in three ways. First to help understanding DLT-enhanced SCT solutions. Second, to shed light into the drivers to deploy these solutions. Third, exploring the required capabilities to process information in these DLT-enhanced SCT solutions. Therefore, we chose to apply an inductive multiple case study design according to Yin (2017). We selected an inductive case study approach to generate mid-range theory from empirical data in the sense of Ketokivi and Choi (2014). As our research questions require an in-depth understanding of the studied phenomenon, case study research is well-suited to address these questions in detail (Ketokivi and Choi, 2014). Furthermore, the novelty of DLT use in general and specifically in SCM (Nærland *et al.*, 2017) and therewith the small number of observable units, call for a qualitative approach such as case study research (Eisenhardt, 1989). Alongside the study design, we drew on the seminal work of Galbraith (1974) and Premkumar *et al.* (2005) on IPT as our theoretical lens. The unit of analysis is the capability needed to process information in DLT-enhanced SCT solutions. Due to the variety of different DLT applications to enhance SCT (Blossey *et al.*, 2019), we decided to study the two major types that we have illustrated in sub-section “Transparency in physical supply chains”. Following this, our case study design is a multiple case design as illustrated in Table 1. The units of observation were five DLT pilots in the diamond industry that presented early pioneer implementations of the technology (Cartier *et al.*, 2018). We conducted fifteen semi-structured interviews of these five different DLT pilots over the timespan of 15 months. This period allowed us to analyze the drivers for DLT use, the needed IPCs and how firms can build these IPCs. The interviews lasted between 60 minutes to four hours. The interviews were conducted in English and German with interviewees in different positions such as CEOs, managing directors, heads of development and project managers. Furthermore, we conducted interviews with multiple organizations in the pilot to gather data from multiple angles. This included interviews with the technology providers, which developed the DLT solution or that developed identifying mechanisms. In addition, we conducted interviews with the user of these solutions as well as the initiators (in some cases they are at the same time the users). The interviews were conducted on site, Skype or via phone. Subsequently, we transcribed the interviews and sent the interviewees a copy for review to ensure correctness.

For our investigation we developed a conceptual framework as illustrated in Figure 1, drawing on IPT by Premkumar *et al.* (2005). The focus of our investigation was to uncover two aspects. First, we covered all four elements of IPT including IPN, IPC, the fit between both and the resulting improvement of SCT with our questions (see appendix) to gain an in depth understanding. Second, we paid special attention to the uncertainties and the IPC in the studied DLT pilots. These aspects were taken into account during case selection and data collection.

Case selection and sampling

For our study design, we sought to study multiple cases of DLT applications that aimed at providing more SCT by processing information with the use of DLT. We chose to select cases from only one industry in order to reach homogeneity of the objects under study. In particular, this is due the reason that different industries face different types of uncertainties (Zhu *et al.*, 2018). Following IPT, these would trigger different types of IPNs (Foerstl *et al.*, 2018) and would make it difficult to analyze the cases on a cross-case analysis. The diamond industry represented one of the first industries to implement multiple DLT pilots at this early stage (Cartier, Ali and Krzemnicki, 2018). After numerous fraudulent incidents (e.g. Malm, 2018; Subramanian, 2018) the processing of information along the supply chain (e.g. product’s origin) was gaining more importance in the diamond industry. Consequently, we defined the improvement of SCT in the diamond industry by the usage of DLT as the first selection criteria. As a second selection criteria, the DLT applications needed to represent a valid unit of observation according to Wüst and Gervais (2017) in order to avoid observing pointless uses of DLT. Furthermore, we decided to study only pilot implementations of DLT in the diamond industry to be able to observe perceived improvements of SCT, which constituted our third selection criteria. Following Eisenhardt (1989) and Seawright and Gerring (2008), our case selection followed a two-step sampling approach. In the first step, we selected cases within the diamond industry as well as with the same goal, namely to enhance SCT by the usage of DLT. In addition, all cases were initiated from experienced actors within the diamond supply chain that possess an in-depth understanding of the industry but were also characterized by an openness to adopt DLT as a novel technology in the field. By this, we established a homogenous base for our case study. In the second step, we selected cases from the initiating companies differing in mining and trade volume (big = 10 carat, medium = 5 to 10 carat or small = less than 5 carat) and in value steps within the diamond supply chain. Selecting cases from different volumes and value steps allows to include different perspectives form different market players. While companies with larger volume have a dominant market position in the diamond industry, smaller companies face different challenges. In addition, different value steps enabled to combine perspectives of mining companies, cutter and polisher as well as auditors. Our case study selection allows to uncover multiple factettes from different angles. This ensured a fuller picture of the entire supply chain and avoided making observations from a single perspective. Thereby, we were able to fulfill both objectives emphasized by Seawright and Gerring (2008). “(1) a representative sample and (2) useful variation on the dimensions of theoretical interest” (Seawright and Gerring, 2008, p. 296). Table 1 gives an overview of the five cases under study.

Table 1: Case overview

ID	**Type of functionalities**		**Interviewed organization**	**Interview partners**	**No. of interviews**
	DLT-based proof of provenance	**DLT-based trade document registry**			
1	x		Gemmology lab, technology provider	CEO; Managing Director; Head of R&D; Gemmologist Head of Sustainability	5
2	x	x	Mining company, technology provider	CEO; Business developer; Director of Communication	3
3	x		Mining company, technology provider	CEO; Communication Manager, Software developer	3
4		x	Industry interest group	Project Manager	2
5	x		Cutter, technology provider	CEO; Vice President & Head of Sustainability	2
Sum	4	2			12

Data collection

For the purpose of data collection, we designed a semi-structured interview protocol following Yin (2017). In the protocol, we defined our research questions (section “Introduction”) and theoretical framework (Figure 1), the data collection procedures, the questions we addressed in the interviews and a guide for the study report. Before collecting data, we applied the three defined selection criteria and followed our sampling approach. Therefore, we conducted desk research and combed press releases, blogs, webpages and industry magazines. Afterwards, we contacted the selected case organizations by sending out a brief description of our research project to the organizations. When the organizations agreed to be participants in our research project, we arranged an appointment for a first onsite interview or a Skype/telephone interview. We made sure to interview people that were not only involved in the pilots but also had different tasks within the pilot project to be able to both gather in-depth data and receive responses from multiple perspectives. Before conducting the interviews, we did a comprehensive research on the diamond industry, and on all pilots and involved organizations. Thereby, we collected data from websites, newspapers, conferences, and industry reports that we used both for preparation in beforehand and for data triangulation subsequently to the interviews. Afterwards, we conducted the interviews. When allowed, we recorded them electronically and made notes. When we were not able to make verbatim records, we took comprehensive notes. The records were transcribed and sent out to the interviewees for review.

Data analysis

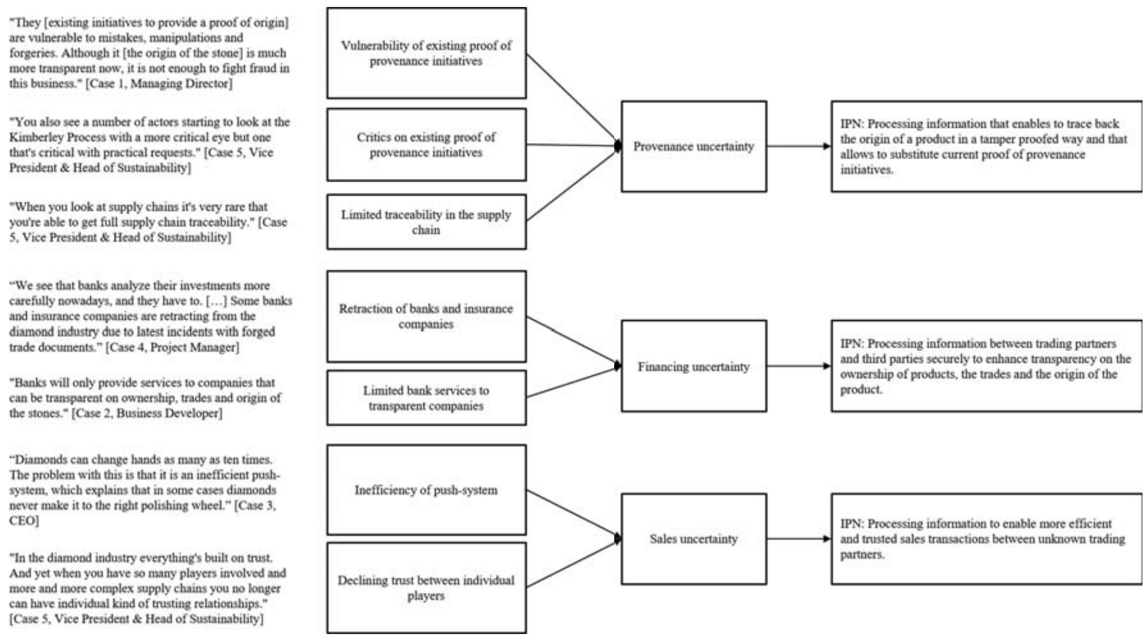
As reviewed above, the extant literature illustrates only few insights into the processing of information for gaining SCT with DLT. Thus, we applied an inductive approach (Eisenhardt, 1989). In line with Mantere and Ketokivi (2013) this approach enabled our inductive reasoning on a three-step journey. First, it helped us to understand the phenomenon of DLT-enhanced SCT in the diamond industry. Second, it allowed us to explain why and how DLT solutions were applied in the diamond supply chain to enhance SCT. Third, it enabled us to develop propositions to guide practitioners when seeking to apply DLT solutions to enhance SCT.
In line with grounded theory following Corbin and Strauss (1990), our data analysis included four steps. First, we conducted in vivo coding on all our transcripts and notes. This supported a deeper understanding and allowed us to identify emerging topics in our data. Herein we followed the approach of Gioia *et al.* (2013) that uses practitioners wording for the first codes with "informant-centric terms" (Gioia *et al.*, 2013, p. 18). As a second step, we created open codes based on our initial codes. With the support of the collected data and the data from additional sources, we were able to contextualize the emerging findings (Gioia *et al.*, 2013). Third, we consolidated the open codes into focused codes by inductive reasoning (Mantere and Ketokivi, 2013). In the following section, we present interview quotes and the corresponding focus codes. Finally, we identified the links of our focused codes to our theoretical framework and reduced the data in our focused codes (FC) that were not related to our framework.

RESULTS

The uncertainties that require more supply chain transparency

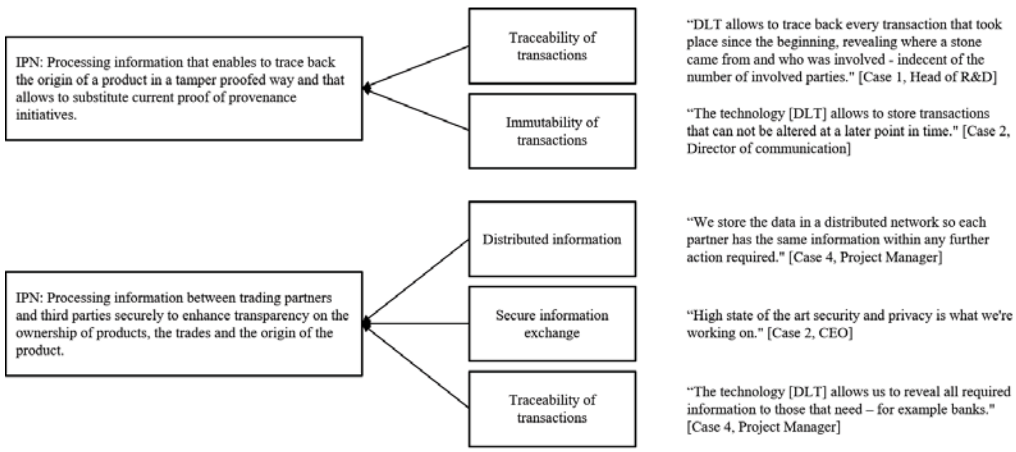
By revisiting the challenges of the diamond and gemstone supply chain from the introduction section, we go on to present the industry briefly and thereby derive the inherent uncertainties that bring to light the IPNs. Along a diamond's journey to the end customer it passes multiple actors from mining productions, to cutting and polishing on to jewelry manufacturing and retail sales (Linde *et al.*, 2018). Especially the gemstone industry is characterized by a large number of small mining facilities in rural and isolated areas. These are the four stages before a diamond or gemstone is sold to the end customer. However, the number of involved parties is much higher. Stones are typically traded multiple times before they make it to the next stage. In addition, these trades require substantial financing. Thus, banks and insurance companies are involved as well. Furthermore, from mine to ring, the stones travel thousands of kilometers, crossing a number of borders and with that require the involvement of customs authorities and additional third parties. With this increasing complexity, the supply chain lacks in transparency that result in incidents such as illustrated in the introduction. This lack of transparency causes uncertainties that characterize the diamond supply chain. Derived from our data, we identified three types of uncertainties: Provenance, financing and sales uncertainties. They create the need to process information between actors within the diamond supply chain. Consequently, these uncertainties result in IPNs, which we illustrate in Figure 2. While on the left side, we present selected statements form interviewees, we condense these to first level codes following the approach of Gioia *et al.* (2013). Based on these codes, we have identified the three types of uncertainties that drive organizations to process information.

Figure 2: Uncertainties that cause IPNs



Facing these IPNs, the initiating companies were attracted to the emerging technology DLT and its promise. As illustrated in Figure 3 the novel technology entails several characteristics that can be beneficial to process information within the diamond industry. These are illustrated in middle and are linked to the corresponding IPNs (on the left side). In summary, the technology enables to build DLT-enhanced SCT solutions that enable improved traceability while establishing a high level of security and immutability, which leads to increased trust, in an industry that is built on trust. Consequently, these initiating companies have developed DLT-enhanced SCT solutions that include two types of functionalities: DLT-enhanced proof of provenance and DLT-based trade document registry. Table 1 presents an allocation of these functionalities to the individual cases while sub-section "Transparency in physical supply chains" describes these functionalities.

Figure 3: IPNs and the corresponding DLT characteristics



Based on our analysis we made the following observation that addresses research question 1:

Observation 1: While organizations face uncertainties stemming from a lack of transparency, they initiate DLT-enhanced SCT solutions to increase transparency and rebuild trust within the supply chain.

The required IPCs in DLT-enhanced SCT solutions

In order to adopt and integrate these DLT-enhanced SCT solutions with their respective functionalities and thus process information successfully to cope with the uncertainties, organizations need to establish specific IPCs in form of processes, infrastructure and IT. These IPCs depend on the specific functionalities. In the following, we present the required IPCs by each functionality and link them to the IPNs.

<u>DLT-based proof of provenance</u>

In order for this functionality to enhance gapless product traceability and provide a proof of provenance, all involved actors have to provide information on their actions related to a specific diamond. Therefore, this functionality requires a wide range of IPCs illustrated in Table 2 from all involved actors. In order to match the IPN to trace back a products origin in a tamper proofed way, all organizations have to establish processes that they view as trusted for data entry and establish joint standards. Only by this, a gapless and trusted traceability can be achieved. Both capabilities are on an interorganizational level, as they require the joint agreement and adherence to these processes. Internally, the actors have to ensure adequate processes that enable this adherence and required infrastructure as well as IT that is necessary for processing the required information. A DLT-based proof of provenance requires to identify a diamond or gemstone in order to enter data to the corresponding diamond. Although there are multiple solutions in place, especially smaller actors may not have the required infrastructure and IT and thus have limited experience in processing the required information. Especially novel and more advanced procedures such as tagging and identifying diamonds and gemstones with nanoparticles that include DNA, are new to most of the actors in the gemstone and diamond industry. Hence, it requires to advanced technical capabilities. However, especially in the gemstone industry that includes a great number of small mining facilities, such advanced procedures will only be applied in a couple of years. Thus, it requires alternative procedures that enable to identify the stones more easily but still securely. Therefore, the involved actors have to establish common processes and standards to enable the integration of this solution for every player in the supply chain, no matter how advanced their technological capability is.

Table 2: IPCs for a DLT-based proof of provenance

IPCs	Cluster	Range	Cases
Capability to establish a process for data entry in DLT-solution that the entire supply chain accepts as trusted	Process	Interorganizational	1, 2, 5
Capability to comply (technologically) with required data standards and formats	Process	Interorganizational	1, 2, 3, 5
Capability to change established work process and incorporate data entry process for DLT-solution	Process	Organizational	1, 2, 3, 5
Expertise of employees to enable correct and timely data entry	Process	Organizational	1, 2, 3, 5
Technological expertise to handle instruments to read out identifiers of diamonds (e.g. nanoparticles)	Infrastructure	Organizational	1, 3, 5
Access to stable and secure internet for secure data entry	IT	Organizational	1, 2, 3

<u>DLT-based trade document registry</u>

In line with the previous functionality, a DLT-based trade document registry requires an interorganizational agreement on data entry and data standards in order to enable a solution that can be used by all parties. Within an organization, substantial change management capabilities have to be developed. As one interviewee stated:

> *“Well the diamond industry is still a 500 plus year old industry that has been based primarily on direct trading behaviors. But even as more people came to the industry and it didn't necessarily become more sophisticated.“* Vice President & Head of Sustainability

A DLT-based trade document registry addresses the IPN to process information securely, mostly in form of documents, between trading partners and banks. This processing of information across organizational boundaries is substantially changed, requiring all parties to buy in and comply with newly established processes. The functionality also requires the capability to enter and transfer the required information securely. A capability that all supply chain actors have to build up on their own but that is crucial for the entire supply chain.

Table 3: IPCs for a DLT-based trade document registry

IPCs	Cluster	Range	Cases
Capability to establish secure data entry process that is verified by all involved parties	Process	Interorganizational	2, 4
Capability to comply (technologically) with required data standards and formats	Process	Interorganizational	2, 4

Change management to enable the transfer from paper-based processes to DLT-based	Process	Organizational	2, 4
Expertise to enable secure data entry in DLT platform	IT	Organizational	2, 4

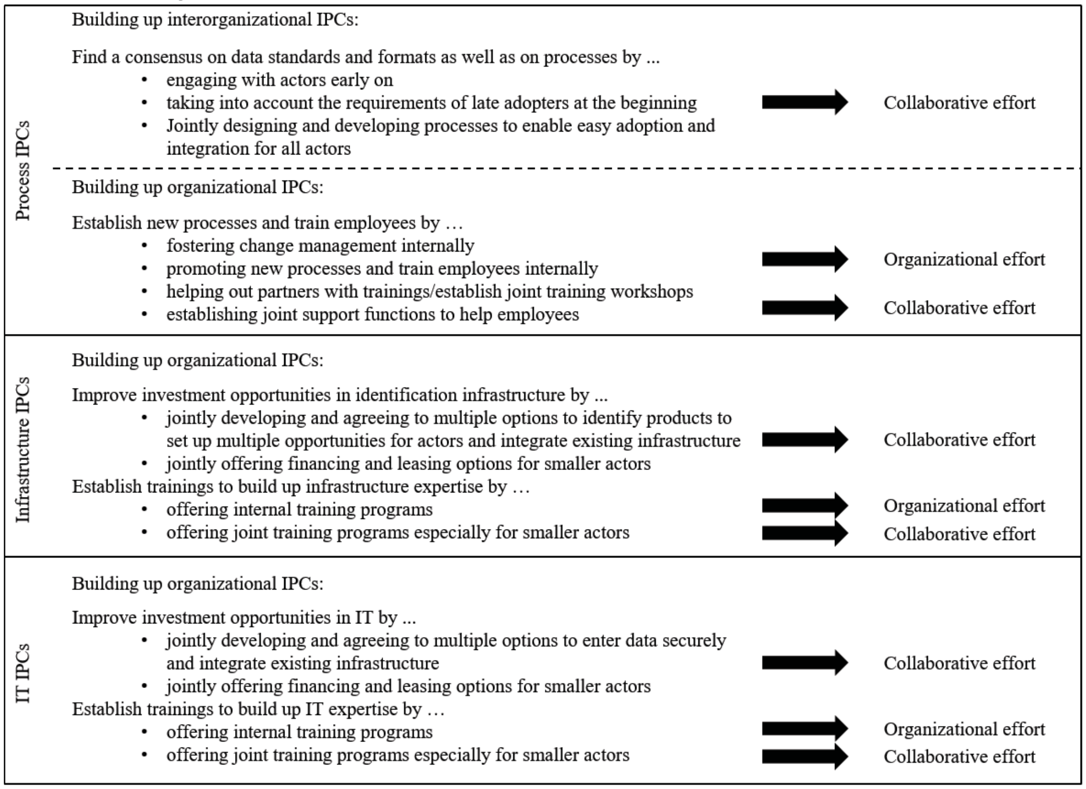
Observation 2: Adopting and integrating DLT-enhanced SCT solutions are joint initiatives that require interorganizational and organizational IPCs in form of processes, infrastructure and IT.

Existing literature (e.g. Zhu *et al.*, 2018) has focused on organizational IPCs, leaving interorganizational IPCs in the shadow. Drawing on the latest contributions of DLT in SCM, it becomes more obvious that DLTs are interorganizational solutions to process information between multiple organizations (Wang *et al.*, 2019). Hence, interorganizational IPCs are required.

Building up the required IPCs

In addition to identifying these IPCs, we have also interrogated the interviewees how they addressed these IPCs or how they will address them in the future. By this, we engaged in a vital dialogue that did also include reflections of interviewees and outlooks. In Figure 4 we summarized the insights from our case studies that reveals how organizations can build up the required IPCs to use and integrate DLT-enhanced SCT solutions. Our analysis reveals that interorganizational IPCs require a collaborative effort in order to build up these capabilities. E.g. this is the case for jointly establishing and agreeing to data standards and formats for data entry in the DLT-enhanced SCT solution. This also included to establish joined trainings and workshops that enable to share the experience and establish joint standards e.g. for the identification of stones and the uploading of data in the blockchain. This is especially crucial for smaller supply chain actors as in the gemstone industry. Bigger actors, associations and technical solution provider helped smaller actors with on-site trainings. Organizational IPCs, representing these IPCs that have to be built up by each party individually. They require the effort of each organization. This includes measurements to build up process, infrastructure and IT IPCs. However, our interview data disclosed that individual organizations should not be left alone with this task. This is especially crucial when the supply chain entails smaller actors that may not have the expertise or financial power to adopt and integrate these DLT-enhanced SCT solutions due to the required IPCs. Therefore, a collaborative effort can help them to build up the needed IPCs in form of processes, infrastructure and IT. Joint training programs, financial help for investments in infrastructure and IT as well as support functions are collaborative measurements that can help smaller actors to build up IPCs and therefore reduce the corresponding challenges. Especially in the gemstone industry, smaller players need the support of other supply chain partners to be able to adopt and integrate such DLT solutions. In some cases, the initiating firms have set up training programs online and onsite to build up the needed capabilities for tagging stones after mining.

Figure 4: Measures to build up process, infrastructure and IT IPCs



When analyzing how organizations can build up the required IPCs, we identified two additional observations that addresses research question 3.

Observation 3: DLT-enhanced SCT solutions require organizational effort but also collaborative effort to build up the required IPCs.

Observation 4: As DLT-enhanced SCT solutions require the adoption and integration of various actors within the supply chains, collaborative efforts can be adequate to help smaller organizations to build up the required IPCs.

Following IPT, DLT-enhanced SCT solutions are mechanism to increase information processing and constitute forms of extending lateral relations and investments in vertical IT.

DISCUSSION AND CONCLUSIONS

With our study, we contributed to practice in three ways. First, we have identified the uncertainties that drive organizations to process information via DLT-enhanced SCT solutions, a discussed phenomenon in the entire field of SCM. We did this by applying a case study approach in the diamond industry, an industry that has proven to be an early adopter of DLT in the field of SCM. Second, we have identified the IPCs that are required to process information via DLT-enhanced SCT solution and revealed that these IPCs are on an organizational but also

on an interorganizational level. Third, we have illustrated how organizations can build up the these IPCs. We have identified that both organizational as well as collaborative effort is needed, to adopt and integrate DLT-enhanced SCT solutions, as they require the adoption and integration of a variety of different supply chain actors. As our study was limited to the diamond and gemstone industry, these findings should lead to further research in other industries. This study reveals important learning for other industries that are characterized by similar structures with heterogenous actors along the supply chain with, strong variation of sales volumes, diverse technical and IT capabilities as well as infrastructure such as the food industry. As enhancing SCT involves processing information over multiple organizational boundaries, the findings are especially in the heterogenous setting important. They emphasize that DLT is not a intrafirm project but rather a interorganizational project, involving diverse actors that may need the help of other supply chain actors to master the onboarding challenges.
With our paper, we also contribute to theory by allocating the phenomenon of SCT in IPT. We identified DLT-enhanced SCT solutions as a form of investment in vertical IT to increase IPC and thus cope with uncertainty. While IPT was applied in the first contributions of SCT (e.g. Williams *et al.*, 2013), we extended IPT literature as we demonstrated that IPCs are to be found on an organizational and on an interorganizational level. Although several scholars have applied IPT in interorganizational settings (e.g. Zhu *et al.*, 2018), our contribution is the first to showcase interorganizational IPCs. Especially the collaboration to enable the adoption and integration of smaller supply chain actors illustrates how vital interorganizational IPCs are for DLT solutions. Extant literature has both dealt with adoption of interorganizational IS and IPT, yet literature has focused on the organization to study adoption and information processing. Our contribution emphasizes the need for collaboration to adopt, integrate and especially use DLT solutions. A finding that neither adoption literature nor IPT literature has taken into account up to now. Future research should address this gap and study the connection of collaboration and technology adoption as well as information processing.

REFERENCE LIST

Abad, E.; Palacio, F.; Nuin, M.; Zárate, A.G.d.; Juarros, A. and Gómez, J.M. *et al.* (2009), “RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain”, *Journal of Food Engineering*, Vol. 93 No. 4, pp. 394–399.

Aung, M.M. and Chang, Y.S. (2014), “Traceability in a food supply chain: Safety and quality perspectives”, *Food Control*, Vol. 39, pp. 172–184.

Baruffaldi, G. and Sternberg, H. (2018), “Chains in chains: Logic and challenges of blockchains in supply chains”, *Proceedings of the 51st Hawaii International Conference on System Sciences*, Waikoloa, Hawaii, pp. 3936–3943.

Beck, R. and Müller-Bloch, C. (2017), “Blockchain as radical innovation: A framework for engaging with distributed ledgers”, *Proceedings of the 50th Hawaii International Conference on System Sciences*, Waikoloa, Hawaii, pp. 5390–5399.

Beck, R.; Stenum Czepluch, J.; Lollike, N. and Malone, S. (2016), “Blockchain: The gateway to trust-free cryptographic transactions”, *24th European Conference on Information Systems (ECIS)*, pp. 1–14.

Bensaou, M. and Venkatraman, N. (1995), “Configurations of interorganizational relationships: A comparison between U.S. and Japanese automakers”, *Management Science*, Vol. 41 No. 9, pp. 1471–1492.

Blossey, G.; Eisenhardt, J. and Hahn, G. (2019), “Blockchain Technology in Supply Chain Management: An Application Perspective”, *Proceedings of the 52nd Hawaii International Conference on System Sciences, 52nd Hawaii International Conference on System Sciences*, Maui, January 8-11.

Carter, C.R. and Easton, L.P. (2011), “Sustainable supply chain management: Evolution and future directions”, *International Journal of Physical Distribution & Logistics Management*, Vol. 41 No. 1, pp. 46–62.

Carter, C.R. and Rogers, D.S. (2008), “A framework of sustainable supply chain management: Moving toward new theory”, *International Journal of Physical Distribution & Logistics Management*, Vol. 38 No. 5, pp. 360–387.

Cartier, L.E.; Ali, S.H. and Krzemnicki, M.S. (2018), “Blockchain, chain of custody and trace elements: An overview of tracking and traceability opportunities in the gem industry”, *Journal of Gemmology*, Vol. 36 No. 3, pp. 212–227.

Casey, M.J. and Wong, P. (2017), “Global supply chains are about to get better, thanks to blockchain”, *Harvard Business Review*. Available at: https://hbr.org/2017/03/global-supply-chains-are-about-to-get-better-thanks-to-blockchain (Accessed: 14 September 2017).

Corbin, J.M. and Strauss, A.L. (1990), “Grounded theory research: Procedures, canons, and evaluative criteria”, *Qualitative Sociology*, Vol. 13 No. 1, pp. 3–21.

Daugherty, P.J.; Myers, M.B. and Richey, R.J.G. (2002), “Information support for reverse logistics: The ifluence of relationship commitment”, *Journal of Business Logistics*, Vol. 23 No. 1, pp. 85–106.

Egels-Zandén, N. and Hansson, N. (2016), “Supply chain transparency as a consumer or corporate tool: The case of Nudie Jeans Co”, *Journal of Consumer Policy*, Vol. 39 No. 4, pp. 377–395.

Eisenhardt, K.M. (1989), “Building theories from case study research”, *Academy of Management Review*, Vol. 14 No. 4, pp. 532–550.

Foerstl, K.; Meinlschmidt, J. and Busse, C. (2018), “It’s a match! Choosing information processing mechanisms to address sustainability-related uncertainty in sustainable supply management”, *Journal of Purchasing and Supply Management*, Vol. 24 No. 3, pp. 204–217.

Galbraith, J.R. (1974), “Organization Design: An Information Processing View”, *Interfaces*, Vol. 4 No. 3, pp. 28–36.

Gioia, D.A.; Corley, K.G. and Hamilton, A.L. (2013), “Seeking qualitative rigor in inductive research: Notes on the Gioia methodology”, *Organizational Research Methods*, Vol. 16 No. 1, pp. 15–31.

Ivanov, D.; Dolgui, A. and Sokolov, B. (2019), “The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics”, *International Journal of Production Research*, Vol. 57 No. 3, pp. 829–846.

Ketokivi, M. and Choi, T. (2014), “Renaissance of case research as a scientific method”, *Journal of Operations Management*, Vol. 32 No. 5, pp. 232–240.

Linde, O.; Geyler, O. and Epstein, A. (2018), *The global diamond industry 2018*. Available at: Bain & Company and AWDC (Accessed: 25 February 2019).

Malm, S. (2018), *Horror at Prague National Museum as curators discover their ‘priceless’ diamonds are cut glass and rubies are synthetic*. Available at: https://www.dailymail.co.uk/news/

article-5477647/Horror-Prague-National-Museum-finds-diamonds-fake.html (Accessed: 4 February 2019).

Maloni, M.J. and Brown, M.E. (2006), "Corporate social responsibility in the supply chain: An application in the food industry", *Journal of Business Ethics*, Vol. 68 No. 1, pp. 35–52.

Mantere, S. and Ketokivi, M. (2013), "Reasoning in organization science", *Academy of Management Review*, Vol. 38 No. 1, pp. 70–89.

Morgan, T.R.; Richey, R.J.G. and Ellinger, A.E. (2018), "Supplier transparency: Scale development and validation", *The International Journal of Logistics Management*, Vol. 4 No. 1, pp. 959–984.

Nærland, K.; Müller-Bloch, C.; Beck, R. and Palmund, S. (2017), "Blockchain to rule the waves: Nascent design principles for reducing risk and uncertainty in decentralized environments", *Proceedings of the International Conference on Information Systems (ICIS) 2017*.

Premkumar, G.P.; Ramamurthy, K. and Saunders, C.S. (2005), "Information processing view of organizations: An exploratory examination of fit in the context of interorganizational relationships", *Journal of Management Information Systems*, Vol. 22 No. 1, pp. 257–294.

Rao, S. *et al.* (2017), *Call for papers - Blockchain: Applications and strategies for supply chain research and practice* (Accessed: 19 April 2018).

Responsible Jewellery Council (2017), *2017 RJC annual progress report*. Available at: https://www.responsiblejewellery.com/annual-progress-report/ (Accessed: 28 February 2019).

Saberi, S.; Kouhizadeh, M.; Sarkis, J. and Shen, L. (2018), "Blockchain technology and its relationships to sustainable supply chain management", *International Journal of Production Research*, pp. 1–19.

Seawright, J. and Gerring, J. (2008), "Case selection techniques in case study research", *Political Research Quarterly*, Vol. 61 No. 2, pp. 294–308.

Srinivasan, R. and Swink, M. (2015), "Leveraging supply chain integration through planning comprehensiveness: An organizational information processing theory perspective", *Decision Sciences*, Vol. 46 No. 5, pp. 823–861.

Subramanian, S. (2018), *Did this jeweler to the stars commit the biggest bank fraud in India's history?* Available at: https://www.bloomberg.com/news/features/2018-06-28/did-this-jeweler-to-the-stars-commit-the-biggest-bank-fraud-in-india-s-history (Accessed: 4 February 2019).

Tsanos, C.S.; Zografos, K.G. and Harrison, A. (2014), "Developing a conceptual model for examining the supply chain relationships between behavioural antecedents of collaboration, integration and performance", *The International Journal of Logistics Management*, Vol. 25 No. 3, pp. 418–462.

Tushman, M.L. and Nadler, D.A. (1978), "Information processing as an integrating concept in organizational design", *Academy of Management Review*, Vol. 3 No. 3, pp. 613–624.

Wang, Y.; Han, J.H. and Beynon-Davies, P. (2018), "Understanding blockchain technology for future supply chains: A systematic literature review and research agenda", *Supply Chain Management: An International Journal*.

Wang, Y.; Singgih, M.; Wang, J. and Rit, M. (2019), "Making sense of blockchain technology: How will it transform supply chains?", *International Journal of Production Economics*, Vol. 211, pp. 221–236.

Wieland, A.; Handfield, R.B. and Durach, C.F. (2016), “Mapping the landscape of future research themes in supply chain management”, *Journal of Business Logistics*, Vol. 37 No. 3, pp. 205–212.

Williams, B.D.; Roh, J.; Tokar, T. and Swink, M. (2013), “Leveraging supply chain visibility for responsiveness: The moderating role of internal integration”, *Journal of Operations Management*, Vol. 31 No. 7-8, pp. 543–554.

Wüst, K. and Gervais, A. (2017), “Do you need a blockchain?”, *IACR Cryptology ePrint Archive*, Vol. 2017, pp. 375–381.

Yin, R.K. (2017), *Case study research and applications: Design and methods*. 6th edn. Los Angeles: SAGE.

Zhu, S.; Song, J.; Hazen, B.T.; Lee, K. and Cegielski, C. (2018), “How supply chain analytics enables operational supply chain transparency”, *International Journal of Physical Distribution & Logistics Management*, Vol. 48 No. 1, pp. 47–68.

DECISION SCIENCES INSTITUTE

Where to from here? Conceptualizing operating strategies for platforms and business ecosystems in smart manufacturing

Sabine Baumann
Jade University of Applied Sciences
Email: sabine.baumann@jade-hs.de

ABSTRACT

Smart manufacturing requires increasingly digitized production settings in order to achieve full flexibility of production environments to fulfil vastly individualized customer requests. Although future value-creation architectures will increasingly be designed and implemented through standardized technological platforms, the operating strategies for the necessary platforms and multi-industry business ecosystems are not yet well understood. This paper aims to mitigate these limitations by presenting the emerging research stream on platforms and business ecosystems. It illuminates the novelty and the usefulness of ecosystem-based approaches for production and operations management, to instigate a cumulative, interdisciplinary body of knowledge.

KEYWORDS: Industry 4.0, platforms, manufacturing industries, ecosystems, value creation networks

INTRODUCTION

Smart manufacturing or Industry 4.0 (I4.0) as it is known in Germany aims at fulfilling extremely individualized customer requests and moves away from predetermined supply chains towards highly flexible production environments. This flexibility challenges the role of the focal firm as coordinator and central decision maker in a production and supply network. Besides implementing the best real-time combination of actors for each individual order (Holweg & Helo, 2014, Stein et al., 2014) focal firms face managerial problems. The actors that contribute to the focal offer's value proposition increasingly operate via platforms and a growing number of complementary offers come from new partners outside the traditional industry settings (Gawer & Cusumano, 2014, Kapoor, 2018). These so-called business ecosystems bring their own set of rules and often ensuing loss of direct control for the focal firm over its production and supply network. In addition, the operating strategies for platforms and multi-industry business ecosystems are not well understood.

This paper aims to mitigate these limitations by presenting the emerging research stream on platforms and business ecosystems. It illuminates the novelty and the usefulness of ecosystem-based approaches for production and operations management in order to instigate a cumulative, interdisciplinary body of knowledge. The author investigates the key concepts in research on platforms and business ecosystems in strategic management and in operations and production management literature and their relationship with established perspectives of value chains and of production and supply networks. Grounded in this analysis the author derives elements and constructs for a cumulative concept for operations and production management on platforms and business ecosystems for I4.0 settings.

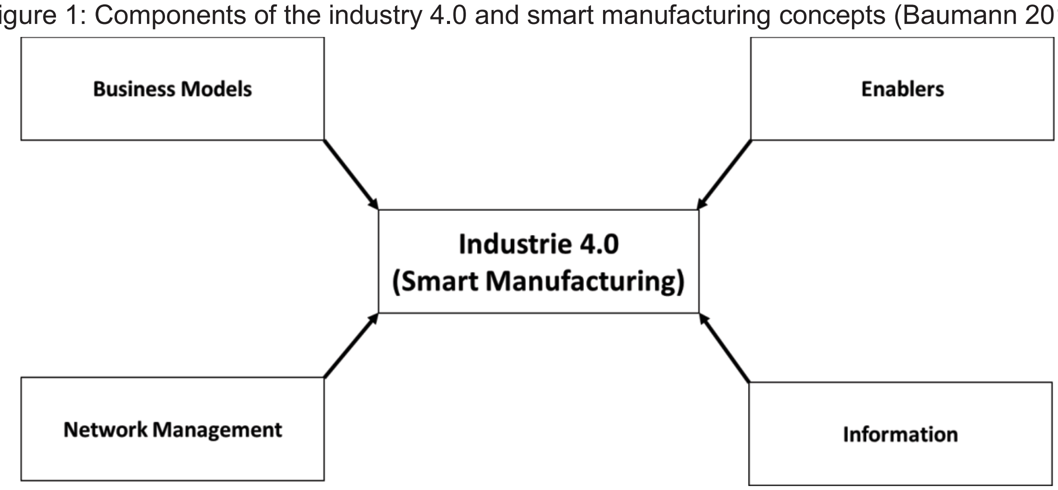
The paper is organized as follows: the first section provides a brief description of the I4.0 concept. The following method section presents information on the design and implementation of the literature review of operations and production management and strategic management research. The main part of the paper depicts the key concepts of the emerging research stream on platforms and business ecosystems and how they relate to established perspectives of value chains and of production and supply networks. Finally, the elements and constructs for a cumulative concept for operations and production management on platforms and business ecosystems are being derived for I4.0 settings. The paper closes with a conclusion and an outlook on future research.

LITERATURE REVIEW

The Concept of Smart manufacturing/Industry 4.0

The German government coined the term Industrie 4.0 (industry 4.0) in their High-Tech Strategy Action Plan 2020 that addresses the increasing global competition on product quality and production costs faced by the German manufacturing industry. Nevertheless, even before the launch German manufacturing companies had started to exploit emerging technologies such as digitalization, the internet of things (IoT), internet of services (IoS) or cyber-physical systems (CPS) in search of new digital business models.
Although the concept Industry 4.0 is now widely discussed, neither academia nor practice has arrived at a generally accepted definition of what the concept comprises in terms of the actual objectives, the relevant technologies, or the applicability for different industrial sectors. Following Baumann (2018) this paper defines industry 4.0 as a concept in which companies, machines, devices and computers cooperate through digital technologies in horizontal and vertical networks to manufacture highly-individualized products. By connecting plants, equipment, machines, products and workpieces through information and communication technologies (ICT) into cyber-physical systems these networks continuously share information in real-time" (p. 2). The term smart manufacturing will be used interchangeably with I4.0 in this paper as it expresses the same ideas.
The I4.0 concept can be captured into four components: business models, enablers, information, and network management (see figure 1), which are briefly explained (see Baumann 2018 and the literature mentioned there for a more detailed description).

Figure 1: Components of the industry 4.0 and smart manufacturing concepts (Baumann 2018)



Business Models
I4.0 business models either address cost reductions or capturing new business opportunities such as innovative products and services. The majority of new digital business models is based on data analytics, for example predictive maintenance. Extended models move to service offers where the customer pays for equipment use, but no longer buys the equipment. From a cost reduction perspective I4.0 increases production flexibility and adaptiveness, speeds up reaction times or improves overall effectiveness of manufacturing equipment.

Network Management
The best real-time combination of actors regarding objectives such as timing, cost, quality levels, etc. for each individual order can only be achieved through extended vertical, horizontal, and lateral collaborations. Actors face a trade-off between installing efficient production and supply networks for particular products while being flexible enough to be part of alternative networks for required product differentiations. Furthermore, the role of the focal firm as principal coordinator becomes increasingly decentralized in I4.0 settings (see also the discussion of business ecosystems below).

Enablers
I4.0 grounds in a variety of enablers, such as *Cyber-physical systems (CPS), Internet of Things (IoT), Internet of Services (IoS),* and *Smart factory. CPS* connect computation with physical processes, so in manufacturing CPS refers to monitoring and controlling the processes via computer networks (through the use of multiple sensors, actuators, control processing units, and communication devices) and synchronising information related to the shop floor. Manufacturing execution systems (MES) link enterprise resource planning (ERP), production planning and control (PPC), and the actual shop floor to integrate all network participants. Thus, MES allow the shop-floor to become a marketplace where the allocation of capacity is self-coordinated with multi-agent systems. In *IoT* settings (physical) things connect to the internet with the use of RFID tags, sensors, actuators, or small computers. These “smart things” interact with each other and cooperate with their neighbouring ‘smart’ components in order to reach common goals. *IoS* makes services available through web technologies or software as a service (SaaS) and allows companies and consumers to access to the resources and service functions of another party. The idea of the *Smart factory* proposes a decentralised production network comprised of human beings, machines and resources. Digital connectivity enables an automated and self-optimised production and delivery free of human interventions or the customary central steering entity. Advanced data analytics derived from the operations of these connected systems provides decision support to both machines and humans.

Information
I4.0 grounds in extensive data collection, (big) data analytics and the implementation of efficient decision support and data-based self-coordination systems. The processes within the information processing chain in I4.0 systems also need to be evaluated from the perspective of value added by these processes, i.e. a detailed analysis of their accompanying value architectures in order to compare value propositions of competing I4.0 solutions (Wunck & Baumann, 2017).

DESIGN/METHODOLOGY/APPROACH

A content-analysis-based literature review is being completed, in which an inductive and multi-level coding procedure is applied on >120 articles on platforms and business ecosystems, spanning the literature in production and operations management and strategic management. The following section shares the key tenets in research on platforms and business ecosystems in strategic management and in operations and production management literature and how they relate to established perspectives of value chains and of production and supply networks.

FINDINGS

Cooperation and Value Creation Scenarios

The cooperation of companies, machines, and computers through digital technologies is central to the provision of individualised products and services in I4.0 settings. These collaborations are realized by integrating technical and organisational processes into order-based value creation architectures that include all actors that contribute to the focal offer's user value proposition. Central elements of these value creation architectures are platforms and business ecosystems which are introduced in the following section.

Platforms

The concept of platforms is being discussed differently in economic (e.g., Rochet & Tirole, 2003 or Cennamo & Santalo, 2013) and engineering (e.g, Meyer et al., 1997 or Jiao, Simpson & Siddique, 2007) research. Economics investigates transactions on platform markets between different customer groups as well as the influence of network effects on competition. In contrast, engineering research regards platforms as technological designs that help firms generate modular product innovations (Brax et al., 2017; Gawer, 2014; Sköld & Karlsson, 2012). Engineers interpret platforms as technological architectures that have been specifically developed (including their interfaces) so that their components can be utilised systematically across different products of a product family. Design and usage of these platforms allow for economies of scope in product development and production. Consequently, implemented platforms bear permanent structures while changes occur for newly developed or redeveloped modules, in order to achieve economies of scope through platform-based product development (Gawer, 2014; Jiao, Simpson & Siddique, 2007; Sköld & Karlsson, 2012).
In economic models platform competition is driven through the adoption of a platform through multiple user groups who themselves are being compelled through network effects. In the relevant literature network effects are distinguished into direct and indirect effects. Direct (or horizontal) network effects emerge through changes in the number of users of a platform (or a technology or a particular product): the higher the number of users, the higher the value for the users. Indirect (or vertical) network effects arise when the benefit of the platform (or technology or product) is not determined by the usage, but is determined by applications or complementary products that are based on or leveraged through the platform. Consequently, the influence on value is not determined by a direct relationship, but through a vertical network of interdependent complementary goods. It follows that it is not single products or components of the network that generate the value, but instead value is created through an inherent interdependency of the demand of two or more user groups. This mutual interdependency of two or more market sides in turn produces a self-supporting feedback loop (Hagiu & Wright, 2015; McIntyre & Srinivasan, 2017).
From the perspective of economic theory, value creation of platform relies on connecting user groups that would otherwise not be linked and could not transact with each other. The price settings of the platform operator coordinate these user groups. Developing the optimal price menu therefore is one of the central problems of platform management. Prices need to be set such that they motivate users from the relevant groups to participate and transact on the platform. Specifically, attracting an increasing number of users from multiple user groups

creates benefits for both the platform participants (increased value) and the platform operator (increased profits) because of the indirect network effects. This means that network effects are an essential attribute of platforms as they reflect the exogenous demand dependencies of different user groups and consequently shape platform competition. Thus, the economic perspective addresses the question under which conditions platforms could become dominant and or even drive out all other platforms in a „Winner-takes-all" rivalry (Eisenmann, Parker & van Alstyne, 2006; Rochet & Tirole, 2003).

<u>Business ecosystems</u>

Digitization not only provides data to develop new data-driven business models, but also requires collaborations with companies that have not been part of the traditional supply chain of an industry. Through these collaborations so-called Business Ecosystems emerge, i.e. company networks that link incumbents from different industries and on horizontal, vertical, and/or vertical level. Simply put a business ecosystem comprises the set of partners that contribute to the value promise of the focal firm (focal offer). The focal offer can consist of a product and/or service that has been created with or without a platform-based architecture. The biological term ecosystem conveys the mutual interdependency of the partners that together operate as a combined system. In addition the analogy suggests that the partners in a business ecosystem may well come from different industries and that the system persistently changes over time: partners compete with each other, new partners join and others leave the system, connections between ecosystem partners change (Kapoor, 2018).

Business ecosystem participants mutually influence each other because of simultaneously existing interdependencies and complementarities. Interdependencies between partners arise through the combination of their offers within the system architecture in the production of the final product and/or service. Opposed to that complementarities emerge because the combination of the partners' offers provides or increases value for the user/consumer. Economics investigates for how companies can survive competition in business ecosystems given these interdependencies and complementarities (e.g., Adner & Kapoor, 2010); Baldwin & Woodard, 2011); Hannah & Eisenhardt, 2017; Jacobides, Cennamo & Gawer, 2018 or Pierce, 2009).

The business ecosystems perspective differs from long-established concepts such as the value chain and the supply chain regarding focus and also the subject of investigation. Porter's value chain centers on the focal firm and investigates how the focal firm can derive competitive advantage in a specific market through organising its value creation activities to achieve cost leadership or differentiation (Porter, 1996). Supply chain research studies the efficient design and coordination of object flows (goods, information, values) along all stages of a supply chain through which a product is manufactured and then distributed to the final customers (MacCarthy et al., 2016). In particular, it investigates coordination problems in buyer-supplier-relationships along the entire supply chain (Pullman & Dillard, 2010). As the supply chain typically involves different companies, supply chain research consequently extends its emphasis beyond the focal firm. Compared to business ecosystems supply chain investigations do not consider complementarities of demand and the structure of interdependencies (Kapoor, 2018; Kapoor & Lee, 2013).

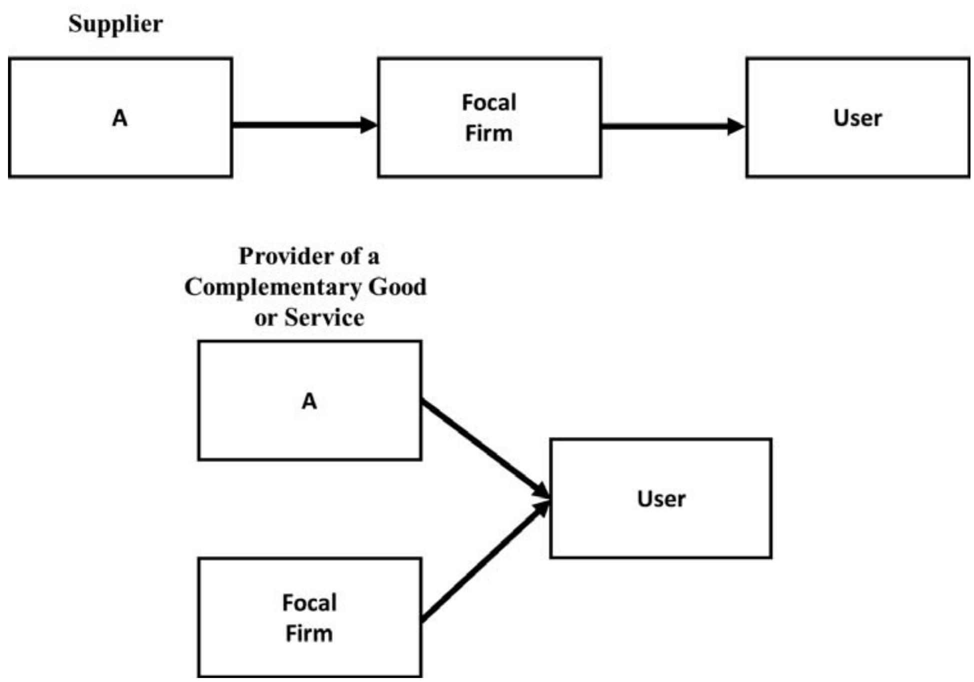
Management of value creation architectures in industry 4.0

The management of value creation architectures in I4.0 comes with a number of challenges. Particularly relevant are the integration of technological platforms, coordinating partners in business ecosystems with complementary offers and dealing with bottlenecks.

Bottlenecks
Bottlenecks can occur in any system that consists of more than one component. They limit the performance of the entire system (dominance of the minimum sector). Accordingly, partners in business ecosystems can constrain the performance and growth of the focal firm’s offer due to their output, capacities or incurred costs. Similarly, interactions between components can be determined through the technological architecture such that improvements in the performance of one partner reduce or balance weaknesses of another partner and vice versa. Identifying and dealing with bottlenecks that result from technological innovations or business models therefore is one of the central problems of business ecosystem management. Besides developing value creation architectures with reduced bottleneck effects business ecosystem management concerns resource allocations between and mutual support of business ecosystem partners in order to resolve bottleneck situations (Adner & Kapoor, 2016; Kapoor, 2018).

Partners with complementary products and services
Providers of complementary products and services (complements) are central to business ecosystems, because they increase the value of the focal firm’s offer (e.g., smartphone apps or infrastructure for charging electric vehicles). The function of the complements differs depending on how they contribute to the focal offer. A strict complementarity between two goods exists, if the goods have no usage value without each other (e.g. razor and blade). A super-modular complementarity describes a relationship where an increase in complements (availability, performance) increases the value of the focal offer (e.g., smartphone and number of apps, or electric vehicles and charging stations). Furthermore, complements can be generic or specialised regarding the focal offer (Jacobides, Cennamo & Gawer, 2018).

Figure 2: Roles in business ecosystems: supplier vs. provider of complementary goods (based on Kapoor, 2018)



The interdependency between providers of complementary products and services at the market level differs from that between a focal firm and its suppliers in the context of a supply chain. The latter concerns a sequential supply-side interdependency determined by the focal firm who decides how the supplied product or service is integrated into the focal offer. In the case of a complementary product or service offer (at the market level) the users of the focal offer decide on the integration of complementary offers. Figure 2 visualises the difference.

Compared to managing traditional supplier relationships the focal firm faces much higher operational challenges in business ecosystems. Relationships with suppliers are typically steered via bilateral contract-based formal mechanisms that regulate amounts, timing, and prices of supplies. Opposed to that the relationship with providers of complements concerns establishing multi-lateral alignment of offers with all relevant partners in the business ecosystem. This alignment relationship is about ensuring successful joint value creation (e.g. standards of cooperation, roles in business models) and the consecutive value appropriation across the partners, including rules to resolve conflict. In addition, the function of the complementary offer (strict, super-modular) poses significant alignment challenges, in order for the focal firm's value promise to be actually realised (Kapoor, 2018).

Platforms

Many value creation architectures organise the offer of complementary products and services via a central platform architecture. It is important to note that the management of product-based and platform-based business ecosystems differs considerably. Platform-based ecosystems are being orchestrated by the platform owner, who designs the platform architecture and determines the rules of participation. As such the platform owner also decides on the alignment structures with providers of complementary goods and services. In contrast, the rules in product-based business ecosystems are being determined jointly by the product company and the providers of complements.

In platform-based ecosystems the owner also defines the interface of the platform architecture through which providers of complements can access the platform as well as any changes over time (e.g., enhanced platform generation). Since every platform has its own specific interfaces providers of complements have to invest for their participation. They also have to decide in how many platform they want to participate (single- vs. multi-homing). Multi-homing provides them with a higher number of market opportunities, but increases costs through the necessary product adaptations for each platform. From the platform operator's perspective increased multi-homing reduces the relative value of their platform compared to competitors (Eisenmann, Parker & van Alstyne, 2006; Kapoor, 2018).

Product-based business ecosystems contain a single-sided market transaction between the focal product firm and the user. Platform-based ecosystems, however, relate to two- or multi-sided markets, on which the platform company interacts with complement providers and users that influence each other's platform benefits through mutual network effects (Rochet & Tirole, 2006). This multi-sided market interaction is a critical aspect of the necessary alignment structure, because it has to be designed such that it reinforces the platform value proposition through adequate price setting and subsidies (Rochet & Tirole, 2006, Rochet & Tirole, 2003; Eisenmann, Parker & van Alstyne, 2011).

In the context of I4.0 companies are expected to be able to develop and provide highly individualised products and service as participants of a value creation architecture. Depending on the order requirements the value creation architecture is assembled from a set of possible actors (Baumann, 2018). That means the architectures typically differ for different orders. Ecosystem participants therefore have to take a number of strategic decisions. As focal firms they have to determine, for example, the structural design of the value creation architecture, the composition of the business ecosystem and which parts of the value creation they want to cover themselves (Kapoor, 2018; Pierce, 2009). Providers of complementary goods and services have different roles depending on whether they participate in product- or platform-based ecosystems. Their opportunities are limited by required technologies and necessary funds as well as their market position. Small and medium-sized enterprises (SME) in particular have to ensure that they are part of competitive value creation architectures, while competing successfully for orders in the business ecosystem. They face a trade-off between differentiation (unique selling proposition) and efficiency (e.g., cost minimisation, resource allocation). An increased degree of differentiation may allow for a higher share of value appropriation, but at the same time limits potential inclusion in a large number of value creation architectures. A higher standardisation improves connectivity, but with the danger of being easily replaceable (Baumann, Eulenstein & Wunck, 2017).

CONCLUSION

The purpose of this paper was to shed light on the evolution and current state of platform and business ecosystems concepts in the literature and to synthesize the strategic management and the operations and production management perspectives into a multifaceted view for I4.0 settings. The concepts were derived through a content-analysis-based literature review on platforms and business ecosystems, spanning literature in production and operations management and strategic management. It was found that the number of studies in operations and production literature on business ecosystems is still low, while strategic management research is more advanced in their discussions. However, the latter often lacks the application perspective which is offered by the dynamic I4.0 settings. Exploring the ecosystem perspective in different research fields therefore facilitates to explain firms' manufacturing and supply strategies and to evaluate outcomes in I4.0 settings through the lens of complementarities and interdependencies. By identifying and defining key elements and constructs the paper proposes a first draft for a platform and business ecosystem framework for extending future production and operations management research.

Collaborations of companies, machines, and computers through digital technologies are central elements of I4.0. This cooperation is implelented through so-called business ecosystems, i.e. sets of actors that contribute to the focal offer's user value proposition. The ecosystem perspective explicitly links supply-side and demand-side of focal offers and also considers contributions from multiple industries. It relates to traditional views of managing production and supply networks though investigating interdependencies between actors, because their offers are connected within a system-level architecture. In addition, it simultaneously considers complementarities between actors whose respective offers help to create or enhance the user value proposition. Consequently, interdependencies embody structural relationships between offers through their value contributions and the way changes in one offer affect the value contribution of other offers, whereas complementarities represent an economic value creation relationship between offers endorsed by the user.

Nevertheless, business ecosystems incur a number of operational challenges. For I4.0 settings the integration of technological platforms, coordinating partners in business ecosystems with complementary offers and dealing with bottlenecks are of particular relevance. These subjects must be tackled on the level of the value creation activities, the perspectives of the different actors and the structure of the business ecosystem. Therefore additional research both in the production and operations management and the strategic management fields is required on three levels: the focal offer (innovation, technology), the focal firm and the business ecosystem.

REFERENCES

Adner, R., & Kapoor, R. (2016). Innovation ecosystems and the pace of substitution. Re-examining technology S-curves, *Strategic Management Journal*, *37*(4), 625–648.

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems. How the structure of technological interdependence affects firm performance in new technology generations, *Strategic Management Journal, 31*(3), 306-333.

Baldwin, C.Y., & Woodard, J.C. (2011). The Architecture of Platforms: A Unified View, in Gawer, Annabelle (Ed.), *Platforms, markets and innovation*, Cheltenham, 19–44.

Baumann, S. (2018). Value Creation Networks for Industry 4.0: best Practices in German Manufacturing Industries, in *Proceedings of the 49th DSI Conference* "Role of Data and Decision Analytics in Global Operations", Chicago (USA), 1647-1659.

Baumann, S.; Eulenstein, O., & Wunck, C. (2017). Digitized Value Creation Architectures between Standardization and Flexibility – A Study on Plastics Manufacturing (Digitalisierte Wertschöpfungsarchitekturen im Spannungsfeld von Standardisierung und Flexibilisierung – Eine Untersuchung am Beispiel der Kunststoffbranche), in Winkler, H, Berger, U., Mieke, C., & Schenk, M. (Eds): *Flexibility of Manufacturing in Industry 4.0 (Flexibilisierung der Fabrik im Kontext von Industrie 4.0)*, Berlin, 2017, 185-206.

Brax, S.A., Bask, A., Hsuan, J., & Voss. C. (2017). Service modularity and architecture – an overview and research agenda, *International Journal of Operations & Production Management*, *37*(6), 686-702.

Cennamo, C., & Santalo, J. (2013). Platform competition. Strategic trade-offs in platform markets, *Strategic Management Journal*, *34*(11, 1331–1350.

Eisenmann, T., Parker, G., & van Alstyne, M.W. (2011). Platform envelopment, in *Strategic Management Journal*, *32*(12), 1270–1285.

Eisenmann, T., Parker, G., & van Alstyne, M.W. (2006). Strategies for two-sided markets, *Harvard Business Review*, *84*(10), 92–101.

Gawer, A. (2014). Bridging differing perspectives on technological platforms. Toward an integrative framework, *Research Policy*, *43*(7), 1239–1249.

Gawer, A., & Cusumano, M.A. (2014). Industry Platforms and Ecosystem Innovation, *Journal of Product Innovation Management, 31*(3), 417–433.

Hagiu, A., & Wright, J. (2015). Multi-sided platforms, *International Journal of Industrial Organization*, *43*, 162–174.

Hannah, D.P., & Eisenhardt, K.M. (2018). How firms navigate cooperation and competition in nascent ecosystems, *Strategic Management Journal*, *39*(12), 3163–3192.

Holweg, M., & Helo, P. (2014). Defining value chain architectures: Linking strategic value creation to operational supply chain design, *International Journal of Production Economics*, *147*, 230–238.

Jacobides, M.G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems, *Strategic Management Journal, 39*(8), 2255–2276.

Jiao, J., Simpson, T.W., & Siddique, Z. (2007). Product family design and platform-based product development. A state-of-the-art review, *Journal of Intelligent Manufacturing, 18*(1), 5–29.

Kapoor, R. (2018). Ecosystems. Broadening the locus of value creation, *Journal of Organization Design, 7*(12), 1–16.

Kapoor, R., & Lee, J.M. (2013). Coordinating and competing in ecosystems. How organizational forms shape new technology investments, in *Strategic Management* Journal, *34*(3), 274–296.

MacCarthy, B. L., Blome, C., Olhager, J., Srai, J. S., & Zhao, X. (2016). Supply chain evolution – theory, concepts and science, *International Journal of Operations & Production Management*, *36*(12), 1696–1718.

McIntyre, D.P., & Srinivasan, A. (2017). Networks, platforms, and strategy. Emerging views and next steps, in *Strategic Management Journal*, *38*(1), 141–160.

Lehnerd, Alvin P., & Meyer, Marc H. (2014). *The power of product platforms*, New York.

Pashaei, S., & Olhager, J. (2017), The impact of product architecture on global operations network design, *Journal of Manufacturing Technology Management*, *28*(3), 353–370.

Pierce, L. (2009). Big losses in ecosystem niches. How core firm decisions drive complementary product shakeouts, in, *Strategic Management Journal*, *30*(3), 323–347.

Porter, M. E. (1996). What is strategy? *Harvard Business Review*, *74*(6), 61–78.

Pullman, M.E., & Dillard, J. (2010). Values based supply chain management and emergent organizational structures, *International Journal of Operations & Production Management*, *30*(7), 744–771.

Rochet, J.C., & Tirole, J. (2006). Two-sided markets. A progress report, *The RAND Journal of Economics*, *37*(3), 645-667.

Rochet, J.C., & Tirole, J. (2003). Platform Competition in Two-Sided Markets*, Journal of the European Economic Association*, *1*(4), 990–1029.

Sköld, M., & Karlsson, C. (2012). Product platform replacements, challenges to managers, *International Journal of Operations & Production Management*, *32*(6), 746-766.

Stein, A., Heddier, M., Knackstedt, R., & Becker, J. (2014). Configuring the Supply Chain Operations Reference Model, *International Journal of Engineering, Science and Technology (IJEST)*, *6*(3), 17–29.

Wunck, C., & Baumann, S. (2017). Towards a Process Reference Model for the Information Value Chain in IoT Applications, in *Proceedings of the 2017 IEEE European Technology and Engineering Management Summit (E-TEMS) Digital Transformation for Advanced Manufacturing – Managing Technological Challenges*, Munich (Germany), 1-6. IEEE Xplore.

DECISION SCIENCES INSTITUTE

Management of Online Server Congestion using Optimal Demand Throttling

Sandun Perera
College of Business, University of Nevada, Reno, NV 89557, sperera@unr.edu
Varun Gupta
Black School of Business, Penn State Erie, Erie, PA 16563, vxg15@psu.edu
Winston Buckley
Mathematical Sciences, Bentley University, Waltham, MA 02452, wbuckley@bentley.edu

ABSTRACT

We model a network congestion management problem of an online-service-provider (OSP) that uses bandwidth throttling as a tool, and show that it is optimal for the OSP to throttle the demand whenever it reaches a threshold level and downgrade the service-speed by a fixed factor while the throttling is employed. Our numerical simulations suggest that it is always optimal for OSPs to induce negative (demand) growth rates during throttling to reduce unfavorable future demand. Moreover, our comparative statics analysis explains how OSPs should handle user demand with higher growth rates and volatility, and service networks that face higher demand fluctuations.

KEYWORDS: Server/Network congestion management, Bandwidth throttling, Stochastic optimal control.

1. INTRODUCTION

Demand for bandwidth exceeds the capacity of networks [Within the context of this paper, 'server' and 'network' can be treated as the same. Thus, we use these two words interchangeably throughout the paper.] during peak demand periods/points, sometimes leading to severe network congestion which generally triggers user backlash towards the online-service-providers (OSPs). For example, in June 2016, a large number of users (several hundred thousand) who were trying to watch the famous HBO drama *Game of Thrones* online, triggered network congestion resulting in more than 15,000 users being denied network access (Hecht 2016). Thereafter, HBO issued an apology and promised to avoid a repeat of the event. In another incident in 2017, due to a server crash at the Foxtel streaming service (an OSP in Australia), the users were again left without access to the same HBO drama series (The Guardian 2017). Foxtel issued the following statement:

"We are devastated that due to unprecedented demand, we are experiencing problems with our online services this evening. Obviously, this comes at a time when so many of our users are wanting to watch the first episode of Season 7 of Game of Thrones".

Users were not happy with Foxtel's network management and criticized the OSP's inability to prepare for the high demand. It is evident that Foxtel had no clear strategy to avoid service interruptions in this situation.

Network congestion usually occurs when a large number of users access the network at the same time or when some users consume a very large amount of network capacity, during busy periods such as peak usage times or planned network maintenance (AT&T.com 2018).

Bandwidth throttling is a commonly used intervention by most OSPs that involves an intentional reduction in bandwidth (speed) whenever network/server congestion occurs. Throttling helps OSPs to preserve the network resources by regulating the traffic that goes through their network, minimizing potential bandwidth congestion and server crashes. Such an intervention is necessary due to a limit on network resources, such as bandwidth capacity in terms of speed, number of servers, and volume of data transfer, etc.

Although the practice of bandwidth throttling has been considered somewhat controversial and unethical [A recent article by Stevens (2018) in the New York Times, reported that Verizon throttled the data connection for firefighters in California, when the firefighters were in the midst of a massive wildfire operation. The chief of the Santa Clara County Fire Department, Anthony Bowden, indicated that Verizon's bandwidth throttling interfered with the ability of the firefighters to function effectively in their operations. In the interest of public safety, the technology department at the Fire Department requested Verizon to remove throttling, however, Verizon recommended the department to switch to a different (more expensive) internet plan that would limit throttling. At a later time, Verizon announced a new plan for the first responders which has no restrictions and have priority access, at a discounted price (VOA.com 2018).] (Krämer et al. 2012, Kourandi et al. 2015, Easley et al. 2018), it is however widely used by most OSPs as a part of their congestion management policy (O'Donoghue 2014). In their defense, OSPs argue that throttling helps to maintain the service quality for all their users, instead of allowing a few users with disproportionately high usage, to negatively impact the experience of other users in a shared network. OSPs such as *Verizon* for cellular and broadband internet, *Spectrum* for broadband internet, *Xfinity* for digital cable TV and broadband internet etc., are known to employ bandwidth throttling policies and these are usually stated in their terms of service. For example, Verizon.com (2018) states on its website:

"On certain plans, we may prioritize your 4G LTE data behind other traffic. If the cell site you are connected to begins experiencing high demand during the duration of your session, your 4G LTE data speeds may be slower than the other traffic's. Once the demand on the site lessens, or if you connect to a different site not experiencing high demand, your speed will return to normal. Any such network management practices will be disclosed in the descriptions of impacted plans.".

Moreover, Spectrum.net (2018) assures its users of the benefits of its network congestion management policies as follows.

"We monitor our network at all times to ensure the highest quality service. During periods of congestion, we employ certain network management principles ('congestion management') to maintain quality of service for all users, rather than allowing a few users with disproportionately high usage to impact the experience of others. While times of congestion are rare, it's important to us that all of our users experience exceptional service at all times."

While server congestion management via bandwidth throttling has been widely used over the last decade in the industry, this topic has been largely neglected in the academic literature. To the best our knowledge, this is the first analytical study that aims at achieving an optimal throttling mechanism for server congestion management when user demand is stochastic.

We model a server congestion management problem as a stochastic (impulse) control problem. The demand of the OSP is modeled as a continuous stochastic—geometric Brownian—process. The OSP can momentarily use demand throttling (i.e., speed reduction) as a tool to regulate demand. This demand throttling procedure involves costs in both the fixed and proportional forms. Moreover, the dynamics of the demand during throttling is assumed to be different in order to capture the reactions from the users as they experience inferior delivery speeds during throttling; this change in demand dynamics is modeled using another stochastic process of which the growth rate is governed by the intensity of throttling. The cost of maintaining demand is assumed to be a quadratic function of the level of demand (i.e., convex and increasing). The objective of the OSP is to minimize the total expected discounted cost of maintaining, and throttling demand. Within this setting, we show that the optimal strategy is to throttle the demand whenever it reaches a threshold level and downgrade the speed by a fixed factor every time throttling is employed. In other words, we show that the optimal strategy is to follow a simple threshold level policy.

A dynamic programming approach is employed to solve our stochastic control problem. We first derive the Quasi-Variational Inequality (QVI) associated with this optimization problem. We then prove, using a verification theorem, that a solution of the QVI coincides with the optimal (minimal) cost of the original problem. While we can only numerically validate the existence and uniqueness of the proposed optimal throttling strategy, we also provide the conditions under which our proposed solution solves the QVI associated with the original problem.

Our numerical study reveals that OSPs should be more patient while applying throttling on user demand with higher growth rates and volatility, and the intensity of throttling for such demands should be higher. Moreover, we find that service networks that face higher demand fluctuations (or volatility) during throttling should be subjected to throttling with lower intensities but applied more frequently. Another finding of our work is that it is always optimal for OSPs to induce negative growth rates while throttling is in place. This result is interesting as OSPs exploit users' willingness to exit the server due to inferior service experience during throttling. Thus, OSPs may use throttling not only to regulate the current demand but also to reduce unfavorable future demand. Furthermore, our numerical study emphasizes the importance of analyzing the trade-off between internal and external non-monetary cost factors (such as the one-time cost of damaging the OSP's public reputation and the cost of damaging the goodwill of current users etc.) and provides guidelines for making decisions on throttling intensities under different scenarios.

The rest of the paper is organized as follows: A section with related literature is provided next. Our demand throttling model and its solution using the QVI approach are presented in Sections 3 and 4, respectively. Section 5 is devoted to a comprehensive numerical study with comparative statics analysis; it also expounds on the implications of the numerical results. We discuss how congestion management with limited capacity is contained within our setting as well as other important implications of our work in Section 6. Section 7 concludes our work with directions for future research and a short summary. Additional simulation results and proofs are provided in appendices.

2. RELATED LITERATURE

We first discuss the stream of literature on network congestion management and then briefly review the related literature on impulse control theory at the end of this section.

The impact of large-scale demand surges resulting in network congestion on web servers and online networks has been extensively studied in computer science, systems science, and other related engineering streams. A few examples of such studies include Stavrou et al. (2004), Deshpande et al. (2007), Kim et al. (2015), and Tada et al. (2017). These studies mainly focus on different engineering mechanism designs to make the OSP resilient and/or prone to network congestion. Our paper focuses on an optimal demand throttling mechanism that reduces the risks associated with network congestion and manages demand during congestion while keeping non-monetary costs associated with throttling at a minimal level. Thus, our contribution lies more on the (operational) policy making domain. This is a very important but largely neglected area of research at the interface between information systems (IS) and operations management (OM).

One important feature of our model is that OSPs treat all users equally when throttling is employed. In other words, our throttling policy is non-discriminatory, that is, it obeys the principle of *network (net) neutrality*. In contrast, a vast majority of the extant literature on network (congestion) management policies focuses on non-neutral policies, where OSPs may provide preferential treatment of certain web-traffic on their networks and may charge a premium for that. Guo et al. (2013) explored the net neutrality issue from both supply (content provider) and demand (customer) sides. One of the key results in their study is that OSPs generally prefer non-neutral network management policies. However, such non-neutral network management policies may not always lead to decrease in customer welfare. From a modeling perspective, they model two different classes of customers based on their usage—heavy and light—who incur a cost (disutility) due to congestion. Thus, they capture the difference in demand levels based on congestion. They assume that customer arrivals are Poisson, whereas, we model aggregate customer demand as a geometric Brownian process. In another study, Guo and Easley (2016) examine the impact on content innovation in the context of net neutrality when customers may experience network congestion. An important result in their study is that a higher level of content generation by providers can also lead to an increase in network congestion. This an important implication for OSPs (e.g., Foxtel) who are looking to mitigate network congestion when there are new offerings by their content providers (e.g., HBO).

Feuillet et al. (2014) study the complex interaction between different classes of stochastic demand. They define classes as *stable* and *unstable*, where these classes depict different user types based on their network usage—low (stable) usage and heavy (unstable) usage. Further, they assume that these classes obtain bandwidth proportional to their incoming demand from the moment of the initial surge until the demand reaches its equilibrium. They use the asymptotic behavior of the unstable class to derive important qualitative insights on how bandwidth allocation affects the overall network performance. In another study, Cho et al. (2016) develop an analytical model using game-theory to analyze the effects of subsidizing internet access for users. They find that the *equilibrium* network management choice of subsidizing the users depends on market conditions, such as the revenue rates of the content providers and the cost to consumers. While these studies, as well as ours, consider congestion management problems in a broader sense, the focus of our study is different.

In a more recent paper, Jordan (2017) studied *zero-rating* program and associated (bandwidth) throttling policies for open internet networks. The zero-rating program has been offered by several OSPs in the United States where some specific network traffic is not included in a user's usage-based pricing policy. These programs are often accompanied by bandwidth throttling (T-Mobile Binge On 2018), where once the user reaches a certain threshold, their network speed is reduced by the OSP as per their contract. Jordan (2017) discussed in detail the no-throttling rule, introduced by Federal Communications Commission (FCC), that provides protection to users against OSP's discriminatory practices that inhibit the delivery of particular content, applications, or services to a user. However, it is important to note that the existing throttling strategies are criticized due to their discriminatory nature. On the other hand, our study focuses on optimal non-discriminatory throttling strategies.

Easley et al. (2018) discuss and categorize key issues and corresponding trade-offs in the context of the debates on net neutrality. Bandwidth throttling is one of the key issues discussed in their paper which is concerned with a *gatekeeper* (OSP) at the infrastructure level that controls and regulates the demand flow in the network. This control can happen at various levels, such as the origin of data, at the software level, hardware level, last-mile to the user, etc. They present several network management tools such as *packet discrimination* including OSP-driven bandwidth throttling, OSP-driven prioritizing of certain network protocols, content-provider driven packet discrimination, content-providers paying the OSPs to lift the data cap for their consumers (zero-rating programs), etc. Furthermore, their study discusses ways of preserving net-neutrality using data caps without discrimination. While we also do not discriminate against users in our setting, we use the aggregate-level bandwidth throttling as opposed to the individual-level data caps.

There is a second stream of literature that investigates how incentivizing users to change their behavior in consuming online services help to mitigate network congestion. Incentive schemes are used as demand management tools employing pricing to shape customer demand and help OSPs to tackle congestion. Dyaberi et al. (2012) conduct experiments with different incentive (pricing) schemes for cellular internet services. They found that users changed their network usage due to economic incentives and disincentives; specifically, users reduce their consumption more due to incentives as compared to disincentives. More recently, De et al. (2017) study quality of service (QoS) based pricing schemes to induce content sharing by end-users in peer-to-peer (P2P) networks. They develop a game-theory based model to study the changes in QoS measures based on the network size as OSPs evolve. For instance, their study indicates that for small networks, higher upload capacity can induce content sharing, however, the opposite is true when the network is large. Zheng et al. (2018) study the change in user behavior due to different pricing policies (unlimited, prepaid, and usage-based) used by different internet service providers (ISPs). They developed a deterministic as well as a stochastic model to capture this user behavior which maximizes a user's net utility from using the ISP's service. They derived analytical and numerical results for an ISP that maximizes its profit based on this specific user behavior without capturing bandwidth throttling. Although we do not consider any explicit incentive scheme in our setting, one could view throttling as a form of disincentive from a service user's standpoint.

We conclude our review on network congestion management highlighting several other notable papers. Anthony and Chung (2013) propose a speed-selection mechanism that allows users to choose different speeds for their data, that helps to limit throttling without increasing the costs for the OSP or without using pricing to manage demand. They show that by adopting this mechanism, the throttling factor can be lowered which reduces the aggregate user dissatisfaction. Li et al. (2017) develop a tool 'lib erate' that can help users to expose the network management policies used by OSPs and then help them to efficiently avoid such measures; they also discuss different throttling policies used by OSPs including AT&T and T-Mobile. In a P2P network setting, Johar et al. (2011) study how customers share content when they experience congestion under three broad categories—total delay, jitter and proportion of packets lost—and argue that successfully managing network congestion can improve end-user performance and may decrease costs for OSPs. In another study, Krämer et al. (2012) explore the allocation problem of existing network capacity through different QoS tiers offered by OSPs for prioritizing the network traffic. Their results indicate that QoS tiering can be more efficient in the short-run as well as long-run because of incentives due to increased demand leading to increased expenditure on infrastructure. Notably, Krämer et al. (2012) assume that the average user demand for service is constant and does not depend on network congestion or OSP's sensitivity to the congestion, whereas, we capture the difference in average user demand before and during the congestion (throttling) period.

Overall, from the extant literature and observed practices of network congestion by OSPs, there is no doubt that bandwidth throttling is a critically important congestion management tool for OSPs. However, research on optimal bandwidth throttling mechanisms is limited or nonexistent. As such, this paper aims to fill this important gap in the literature. To the best our knowledge, this is the first analytical study that aims at achieving an optimal non-discriminatory throttling mechanism for bandwidth when user demand is stochastic. We next briefly review the related literature on stochastic (impulse) control models.

We formulate our congestion management problem as a stochastic impulse control model and use the dynamic programming (DP) approach to solve the problem; this classical DP approach is commonly referred as the Quasi-Variational Inequality (QVI) approach in the literature (Bensoussan and Lions 1984). This celebrated method has been used extensively in the literature (Constantinides and Richard 1978, Sulem 1986, Presman and Sethi 2006, Cadenillas et al. 2010, Yamazaki 2017, etc.). In particular, Harrison et al. (1993), Korn (1997), Ormeci et al. (2008), Feng and Muthuraman (2010), Dai and Yao (2013a,b), Mitchell et al. (2014), Wu and Chao (2014), Muthuraman et al. (2015) and He et al. (2017) use the QVI approach when the underlying stochastic process follows a Brownian process, whereas Cadenillas and Zapatero (1999, 2000), Lumley and Zervos (2001), Ohnishi and Tsujimura (2006), Bensoussan et al. (2012), and Bensoussan and Chevalier-Roignant (2019) employ the QVI approach when the underlying process specifically follows a geometric Brownian process, as in our setting. We also refer our readers to Ivanov et al. (2018) and Dolgui et al. (2019) for state-of-the-art applications of control theory in the context of Industry 4.0 and digitalization.

While all these seminal papers use the QVI approach as a common tool, their model formulations and underlying applications are different; moreover, each article has its own merit in terms of use of the technique and managerial insights. In a similar spirit, we apply the QVI approach to solve a problem at the interface of IS and OM for the first time. Even among the models that use a geometric Brownian process as the underlying process, our model has several distinct features. First, rather than intervening only at the stopping times, we allow intervention (or throttling) over a random period of time which we call the *throttling period*. Second, the dynamics of the underlying process is allowed to be different during this throttling period and the corresponding dynamics depend on the throttling factor associated with each throttling period. Finally, the effective-demand (process) is only a fraction of the underlying process during the throttling period. These features clearly distinguish our model from the extant models. The model that is somewhat closest to ours is that of Bensoussan et al. (2012). Specifically, these authors also use a different stochastic process (over a period called *market reaction period*) to capture market reactions to central bank interventions in the foreign exchange market. However, they apply the impulsive type of (instantaneous) interventions while the throttling is applied over a period of time in our setting. In addition, the dynamics of the underlying process during the throttling period is governed by the intensity of the intervention (i.e., throttling factor) in our setting whereas the corresponding dynamics during the market reaction period of Bensoussan et al. (2012) is entirely governed by exogenous parameters. Consequently, our model is novel, and our work contributes to the extant literature on both the IS-OM interface and stochastic control theory.

3. DEMAND THROTTLING MODEL

Our model focuses on an online service provider (OSP) that tries to keep its demand for bandwidth below a certain threshold level to avoid server/network congestion. We assume that the demand (aggregate bandwidth usage) follows a geometric Brownian process. Formally, let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with an information set (filtration) $\{\mathcal{F}_t\}_{t\geq 0}$ generated by a 1-dimensional Brownian process $B(t)$. Then, the demand process $X(t)$ at time t, when there is no throttling (control), is the solution of the following stochastic differential equation (SDE):

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t), \tag{1}$$

where μ is the growth rate, $\sigma > 0$ is the volatility parameter of demand, and $X(0) = x > 0$.

OSP manages demand by throttling it for a random period of time. Throughout the throttling period, OSP slows down its service (or bandwidth speed) so that each user experiences a slower speed. Mathematically, service speed will be multiplied by a *fixed* fraction ρ, where $0 \leq \rho < 1$, in this period. For example, if $\rho = 70\%$, then each user receives only 70% of the speed that he/she received before throttling is employed. At the end of this period, the speed will be reinstated. Moreover, the intensity of throttling is represented by $\gamma := (1 - \rho)$ and γ will be referred to as *throttling factor (intensity)* hereafter.

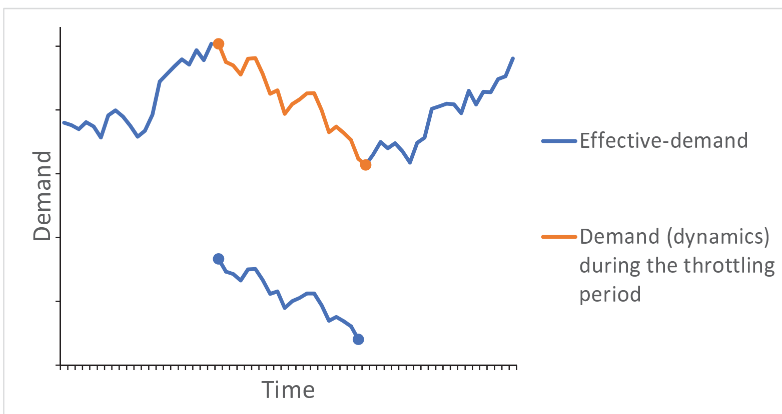
We now turn our attention to the aggregate demand and its dynamics during the throttling period. Due to the lack of speed in this period, multiple users (especially, the users who are more sensitive to the speed) may decide to stop using the service network temporarily; in extreme cases, some users may switch to different service providers permanently because of their inferior service experience [When Verizon throttled the data connection for firefighters in California, the Fire Department relied on other agencies for their internet service and firefighters used their personal internet devices, and finally, the department switched to a more expensive data plan to avoid throttling (Stevens 2018). It is also reported that, as a result of throttling, users switch to other network carriers including using Wi-Fi instead of cellphone internet, or using a mobile hotspot of another user, or in extreme cases users abandon the service intermittently (Hildenbrand 2017).]. Thus, the dynamics of the demand process should switch during the throttling period. Moreover, recognizing that the rate of exiting the service should be proportional to throttling factor γ, we assume that the demand follows the process $\tilde{X}$ during the throttling period and its dynamics is governed by the following SDE:

$$d\tilde{X}(t) = (\mu - \tilde{\mu}\gamma)\tilde{X}(t)dt + \tilde{\sigma}\tilde{X}(t)dB(t), \tag{2}$$

where $\tilde{\mu} > 0$ and $\tilde{\sigma} > 0$ are constants. Here, $\tilde{\sigma}$ represents the volatility of demand during the throttling period, whereas $\tilde{\mu}_g := (\mu - \tilde{\mu}\gamma)$ is the growth rate of demand. Our selection of the growth rate $\tilde{\mu}_g$, with the assumption $\mu < \tilde{\mu}$, allows us to have both the positive and negative growth rates for the demand during the throttling period. We consider two cases for discussion—$\mu \leq 0$ and $\mu > 0$. If $\mu \leq 0$, then it is clearly seen that $\tilde{\mu}_g \leq \mu < 0$. Thus, the growth rate of the demand during the throttling period is negative and lower than that of pre-throttling demand; moreover, the magnitude of the negative demand is proportional to γ. This is clearly expected at any level of throttling when $\mu \leq 0$. However, if $\mu > 0$, then the growth rate $\tilde{\mu}_g$ during the throttling period could either be positive or negative depending on the intensity of throttling. Specifically, $\tilde{\mu}_g > 0$ if $\mu/\tilde{\mu} > \gamma$, and $\tilde{\mu}_g < 0$ if $\mu/\tilde{\mu} < \gamma$. Thus, the growth rate changes its direction only if the intensity of throttling exceeds the threshold of $\mu/\tilde{\mu}$.

While the dynamics of the demand is governed by $\tilde{X}$ during the throttling period, the effective (aggregate) demand during this period is $\rho\tilde{X}$ and this will be the (total) bandwidth usage during the throttling period. Figure 1 illustrates the demand and effective-demand around a throttling period. The figure highlights two important features of our model. First, the demand dynamics during the throttling period is different as users experience slower speeds and the corresponding dynamics depend on the throttling factor (see also (2)). Second, the effective-demand (or true usage) during the throttling period is only a fraction of the demand during that period. These features somewhat distinguish our model from the extant impulse control models (*cf.* Cadenillas and Zapatero 1999, Lumley and Zervos 2001, Ohnishi and Tsujimura 2006, Bensoussan et al. 2012, Bensoussan and Chevalier-Roignant 2019). Finally, note that the demand dynamics reverts to the original dynamics given in (1) at the end of the throttling period. However, the demand process evolves from the level of demand at the end of the throttling period under the reinstated speed; in order to compute this starting level of demand, one should multiply the effective-demand at the end of the throttling period by $(1-\gamma)^{-1}$ (or, equivalently by ρ^{-1}). We have mathematically defined these in the sequel (see the system (3)-(5)).

Figure 1: Effective-demand during the throttling period



Despite the fact that demand throttling is employed over a period of time, we refer to the starting instance of the throttling period as the *throttling time* throughout this article. OSP decides on the times $\{\tau_n\}_n$ at which to initiate demand throttling and also on the throttling intensities (factors) $\{\gamma_n\}_n$. Formally, we can define OSP's throttling policy as an impulse control.

Definition 1. *OSP's throttling policy is a double sequence* $\nu = (\tau_1, \tau_2, \ldots, \tau_i, \ldots; \gamma_1, \gamma_2, \ldots, \gamma_i, \ldots)$, *where* $\tau_1 < \tau_2 < \cdots$ *are* $\mathcal{F}_t$*-stopping times (the throttling times) and* $\gamma_1, \gamma_2, \ldots$ *are the corresponding throttling factors (intensities) employed over the throttling periods initiated at these times such that each* γ_i *is* $\mathcal{F}_{\tau_i}$*-measurable.*

As we noted above, demand throttling is applied over the entire throttling period although the throttling is initiated as an impulse at each throttling time. This feature of our throttling (impulse-control) policy is novel and distinguishes our policy from the classical frameworks in Bensoussan and Lions (1984), Øksendal and Sulem (2007), etc.

Now, let $X_x(t)$ (resp., $\tilde{X}_x(t)$) denote the demand process $X(t)$ (resp., $\tilde{X}(t)$) with $X(0) = x$ (resp., $\tilde{X}(0) = x$). Moreover, we denote the duration of the i-th throttling period by T_i; and assume that T_i's are independent and identically distributed, and has the distribution function F_T. (Later in the paper, we further assume that T_i's are also independent of $B(t)$, and hence independent of the demand processes.) Then, if we apply a throttling policy $\nu = (\tau_1, \tau_2, ..., \tau_i, ...; \gamma_1, \gamma_2, ..., \gamma_i, ...)$ to the demand process $X_x(t)$, the corresponding throttled (or controlled) demand process $X_x^{(\nu)}(t)$ can be defined by

$$X_x^{(\nu)}(t) = X_x(t)\ ;\ 0 \le t < \tau_1, \tag{3}$$

$$X_x^{(\nu)}(t) = (1-\gamma_i)\tilde{X}_{X_x^{(\nu)}(\tau_i -)}(t) \text{ for } \tau_i \le t \le \tau_i + T_i\ ;\ i = 1, 2, ..., \tag{4}$$

$$X_x^{(\nu)}(t) = X_{(1-\gamma_i)^{-1}X_x^{(\nu)}(\tau_i + T_i)}(t) \text{ for } \tau_i + T_i < t < \tau_{i+1}\ ;\ i = 1, 2, ... \tag{5}$$

Equation (3) above means that the throttled demand process simply follows the original demand process (1) until the first throttling time τ_1. OSP initiates each throttling period and throttles demand according to equation (4). At each τ_i, as soon as throttling is applied, the dynamics of the demand process switches to $\tilde{X}$; while this new process $\tilde{X}$ evolves from the level of demand just before throttling (i.e., $X_x^{(\nu)}(\tau_i -)$), OSP's effective demand of this period is derived as in (4). In particular, equation (4) assures that each user receives only $(1-\gamma_i)$ percentage of the original speed throughout the throttling period. At the end of the throttling period (i.e., $t = \tau_i + T_i$), OSP reinstates the original speed by multiplying the current speed by $(1-\gamma_i)^{-1}$. This results in a change in dynamics as well as a change in the level of demand as indicated in (5). Specifically, the post throttling period demand evolves from the initial level of $(1-\gamma_i)^{-1}X_x^{(\nu)}(\tau_i + T_i)$ according to the original demand dynamics in (1) until the next throttling time.

To capture the qualitative factors associated with this problem, we have incorporated a running cost and a throttling cost in our model. The running cost is assumed to be a quadratic function of the level of demand x. We denote this cost by $f(\cdot)$ with $f(x) := \beta x^2$ for some $\beta > 0$. This cost is motivated by the potential risks associated with server congestion; these risks include, but not limited to, server crashes, queueing delay, packet loss and the blocking of new connections. Note that, as the level of demand increases, so does the risk associated with congestion. Moreover, the marginal risk of having congestion should be increasing as the level of demand increases. Consequently, the running cost associated with the server congestion should be convex. Notwithstanding, we assume the specific form of $f(x)$ above to simplify the analysis.

Next, we assume that the cost of throttling is $K + c\,\zeta$, where K and c are positive constants and $\zeta > 0$ is the immediate change (drop) in the level of effective-demand at the launch of throttling; for example, if the effective-demand instantly drops down from level y to ρy when throttling is initiated, then $\zeta := (1-\rho)y = \gamma y$. Here, ζ is the size of the immediate impact of throttling. The costs associated with throttling is driven by qualitative features of our problem. Note that frequent demand throttling is not encouraged in practice as throttling immediately frustrate users (irrespective of the intensity of throttling) and that significantly hurts the public reputation of the OSP each time throttling is initiated. One way to incorporate these infrequent throttling in our model is to consider a fixed cost associated with throttling; this part of the throttling cost does not depend on the intensity of throttling. On the other hand, demand throttling with higher intensities should frustrates users more compared to relatively lower intensities; thus, the throttling cost should also include a cost that is proportional to the intensity of throttling.

The objective of the risk-neutral OSP is to minimize the total expected discounted future costs of managing congestion via demand throttling. For a given discount rate $r > 0$, an initial demand level $x > 0$ and a throttling policy ν, we can define this (cost) objective as below.

$$C^{(\nu)}(x) := \mathbb{E}\left[\int_0^\infty e^{-rt} f(X_x^{(\nu)}(t))dt + \sum_{i=1}^{\infty} e^{-r\tau_i}\left\{K + c\gamma_i X_x^{(\nu)}(\tau_i -)\right\}\right]. \tag{6}$$

For a given x, we are interested in the minimizer(s) of $C^{(\nu)}(x)$ over a set of admissible class of throttling policies which we denote by $\mathcal{V}$ and define as follows:

Definition 2. *The set $\mathcal{V}$ of throttling policies is called admissible, if for all $x > 0$, $\nu \in \mathcal{V}$,*
(i) a unique solution $X_x^{(\nu)}(t)$, $t \geq 0$, of the system (3)–(5) exists,
(ii) $\lim_{i\to\infty} \tau_i = \infty$ a.s.,
(iii) $E\left[\int_0^\infty e^{-rt} f(X_x^{(\nu)}(t))dt\right] < \infty$,
(iv) $E\left[\sum_{i=1}^\infty e^{-r\tau_i}\left\{K + c\gamma_i X_x^{(\nu)}(\tau_i -)\right\}\right] < \infty$,
(v) $\gamma_i \in (0, 1]$, for $i = 1, 2, ...$, and
(vi) $\tau_{i+1} - \tau_i \geq T_i$, for $i = 1, 2, ...$

While conditions (i)-(iv) above are quite common in the literature (*cf.* Chapter 6 of Øksendal and Sulem (2007)), it will be clear in the sequel that some of these conditions can be replaced by simple sufficient conditions; for example, condition (iii) is implied by $(r - \sigma^2 - 2\mu) > 0$. Moreover, condition (v) is necessary for a meaningful analysis in our setting. Condition (vi), however, is somewhat uncommon and has been imposed only in impulse control models with random reaction periods (*cf.* Bensoussan et al. 2012, Perera et al. 2018). This condition implies that OSP waits until the current throttling period is over before implementing the next throttling period. We believe that this assumption is reasonable for our application. For example, if there is a chance of exceeding a certain threshold again while the throttling is in place, then the optimal decision under the current policy would have been to throttle by a larger factor; furthermore, it makes perfect sense for an OSP to wait and study the impact of the current throttling period before implementing a new throttling. This restriction is also supported by the fact that fixed costs associated with throttling is large enough so that frequent throttling is not economical. We recognize that, under a more general class of admissible policies, it is theoretically feasible to have scenarios where the next throttling time lands within the current throttling period. However, considering our application and for mathematical simplicity, we assume that $\tau_{i+1} - \tau_i \geq T_i$ for all i.

Finally, let Φ denote the value function (optimal minimal cost) of our congestion management problem, i.e., for all $x > 0$,

$$\Phi(x) = \inf\left\{C^{(\nu)}(x); \nu \in \mathcal{V}\right\} = C^{(\nu^*)}(x),$$

where ν^* is the minimizer of $C^{(\nu)}(x)$ over $\mathcal{V}$. The analysis of the value function Φ is not trivial. Therefore, in the next section, we convert this dynamic optimization problem into a set of point-wise optimization problems using the dynamic programming approach. This classical approach is referred to as the QVI approach in the literature (Bensoussan and Lions 1984).

4. QUASI-VARIATIONAL INEQUALITY AND ITS SOLUTION

The derivation of the QVI associated with our demand throttling model is standard. Thus, we will be brief and only present the essential steps in the derivation. In our setting, at any given state x, the OSP can either choose to throttle (optimally) or stay put. If the OSP chooses either of these alternatives immediately and thereafter follow the optimal (feasible [Note that OSP has to wait until the immediate throttling period is over before initiating another throttling period. Thus, if OSP has chosen to throttle immediately, then any policy that has a throttling time within the immediate throttling period will not be feasible.]) throttling policy, then the cost of the alternative policy should be no less than that of the 'optimal' policy which is $\Phi(x)$. Let us consider each of these scenarios separately. First, consider the case where the OSP starts throttling immediately with intensity γ and then follows the optimal policy starting with the state $\tilde{X}_x(T)$, which is the state at the end of the immediate throttling period. The cost of this policy is

$$K + c\gamma x + \int_0^\infty \mathbb{E}\left[\int_0^T e^{-rs} f\left((1-\gamma)\tilde{X}_x(s)\right) ds + e^{-rT}\Phi\left(\tilde{X}_x(T)\right)\Big| T = t\right] dF_T(t).$$

The first two terms of the above cost expression represent the throttling cost which consists of a fixed cost K and a proportional cost $c\gamma x$. The first part of the final term is the running cost during the throttling period which is governed by the effective-demand level $(1-\gamma)\tilde{X}_x$. The second part of the final term is the cost of following the optimal policy, starting with the state $\tilde{X}_x(T)$, at the end of the throttling period. If the OSP chooses this immediate throttling decision optimally over $\gamma \in (0, 1]$, then the corresponding cost can be written as

$$\inf\left\{K + c\gamma x + \int_0^\infty \mathbb{E}\left[\int_0^T e^{-rs} f\left((1-\gamma)\tilde{X}_x(s)\right) ds + e^{-rT}\Phi\left(\tilde{X}_x(T)\right)\Big| T = t\right] dF_T(t)\ ;\ \gamma \in (0,1]\right\}. \tag{7}$$

The cost expression in (7) is referred to as the *minimum cost operator* in the literature (Øksendal and Sulem 2007) and we denote it by $\mathcal{M}\Phi(x)$. As we explained above, the cost of this alternative policy should be no less than that of the optimal policy. Thus, we have $\Phi(x) \le \mathcal{M}\Phi(x)$.

Our minimum cost operator is distinct from the standard literature (Bensoussan and Lions 1984, Øksendal and Sulem 2007). In particular, we have included a throttling period where no additional control is allowed. Moreover, the dynamics of the demand process is governed by a different process during this period. With the change in demand dynamics during the throttling period, our model is akin to the impulse control model with random reaction periods (ICRRP) proposed by Bensoussan et al. (2012). However, our model has two distinct features. First, it uses the effective-demand as the underlying stochastic process throughout the throttling period; in other words, the controller applies the impulse throughout the throttling period. Second, the dynamics of the $\tilde{X}$ process depends on the value of γ which is the intensity of throttling.

Next, we consider the other alternative policy where the OSP stays put (without throttling) for a short duration of length $\varepsilon > 0$ and then follows the optimal throttling policy starting with the state $X_x(\varepsilon)$. It is easily seen that the cost of this alternative policy is

$$\mathbb{E}\left[\int_0^\varepsilon e^{-rt} f(X_x(t))dt + e^{-r\varepsilon}\Phi(X_x(\varepsilon))\right].$$

From our intuitive argument at the beginning of this section, or from the principle of optimality (Bellman 1957), we know that the cost above should be no less than the optimal cost $\Phi(x)$, for any $x > 0$. Therefore, for any $\varepsilon > 0$, we have

$$\Phi(x) \leq \mathbb{E}\left[\int_0^{\varepsilon} e^{-rt} f(X_x(t))dt + e^{-r\varepsilon}\Phi(X_x(\varepsilon))\right].$$

Dividing both sides of the above expression by ε and rearranging the terms, we obtain

$$\frac{\mathbb{E}\left[e^{-r\varepsilon}\Phi(X_x(\varepsilon))\right] - \Phi(x)}{\varepsilon} + \frac{\mathbb{E}\left[\int_0^{\varepsilon} e^{-rt} f(X_x(t))dt\right]}{\varepsilon} \geq 0.$$

Now, letting $\varepsilon \to 0$, we derive $A\Phi(x) + f(x) \geq 0$, where

$$\begin{aligned} A\Phi(x) &:= \lim_{\varepsilon \to 0} \frac{\mathbb{E}\left[e^{-r\varepsilon}\Phi(X_x(\varepsilon))\right] - \Phi(x)}{\varepsilon} \\ &= \mu x \frac{\partial \Phi(x)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 \Phi(x)}{\partial x^2} - r\Phi(x). \end{aligned}$$

Here, the differential operator A is defined on the set of twice continuously differentiable functions.

Finally, note that at any initial level of demand $x > 0$, the OSP can either choose to throttle (optimally) or stay put. Therefore, one of the weak inequalities from $\Phi(x) \leq \mathcal{M}\Phi(x)$ and $A\Phi(x) + f(x) \geq 0$ should be strict whereas the other is an equality. Thus, we should also have the following complementary slackness condition:

$$[A\Phi(x) + f(x)]\,[\Phi(x) - \mathcal{M}\Phi(x)] = 0\,.$$

We summarize the QVI associated with the value function of our problem in the proposition below.

Proposition 1 (QVI)**.** *For a given $x > 0$, the value function $\Phi(x)$ of our congestion management problem, which is assumed to be continuously differentiable [We only need $\Phi(\cdot)$ to be continuously differentiable in order to approximate it by a sequence of twice continuously differentiable functions using a well-known approximation result (cf. Appendix D of Øksendal 2003).], satisfies the following QVI:*

$$\begin{aligned} \Phi(x) &\leq \mathcal{M}\Phi(x), \\ A\Phi(x) + f(x) &\geq 0, \\ [A\Phi(x) + f(x)]\,[\Phi(x) - \mathcal{M}\Phi(x)] &= 0\,. \end{aligned}$$

A solution of the QVI above divides the interval $[0, \infty)$ into two disjoint regions—a continuation (non-throttling) region, $\mathbb{C} = \{x : \Phi(x) < \mathcal{M}\Phi(x)$ and $A\Phi(x) + f(x) = 0\}$, and a throttling region, $\mathbb{I} = \{x : \Phi(x) = \mathcal{M}\Phi(x)$ and $A\Phi(x) + f(x) > 0\}$. While OSP does not need to take any action when the demand is strictly inside the continuation region, throttling should be initiated as soon as the level of demand reaches the boundary of the continuation region provided there is no ongoing throttling.

The rest of this section will be dedicated to the solution of the QVI. However, it should be noted that one needs to verify that a solution of the QVI indeed coincides with the original value function Φ. This requires a standard verification argument which we have presented in Appendix A for the sake of completeness.

Before we turn our attention to the solution of the problem, we derive an important identity. It is well-known that the solution of (1) is given by

$$X_x(t) = x\ \exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)\right\}.$$

Then, using the exponential moments of the process, it follows that, for $q \in \mathbb{R}$,

$$\mathbb{E}\left[(X_x(t))^q\right] = x^q \exp\left\{\left[\left(\mu - \frac{1}{2}\sigma^2\right) q + \frac{1}{2}\sigma^2 q^2\right] t\right\}. \tag{8}$$

A similar identity holds for $\tilde{X}$ process, i.e.,

$$\mathbb{E}\left[(\tilde{X}_x(t))^q\right] = x^q \exp\left\{\left[\left((\mu - \tilde{\mu}\gamma) - \frac{1}{2}\tilde{\sigma}^2\right) q + \frac{1}{2}\tilde{\sigma}^2 q^2\right] t\right\}. \tag{9}$$

Next, we conjecture that the solution of the QVI associated with our congestion management problem is a policy governed by a positive threshold U and a control factor Υ, where $0 < \Upsilon \leq 1$; the optimal strategy is to initiate a throttling period with intensity Υ whenever the level of demand reaches the threshold U outside a throttling period. The continuation (non-throttling) and throttling regions are $(0, U)$ and $[U, \infty)$, respectively. One could view the proposed policy as a modified version of the classical control-band policy where the optimal policy is to bring down the state process to level $u := (1 - \Upsilon)U$ whenever it reaches a threshold level $U > 0$. However, our policy is distinct in that the demand throttling is applied over a period of time as oppose to an instantaneous (impulsive) control.

We now follow standard techniques to find the optimal cost and policy parameters. Note that, using the techniques of Boyce and Prima (1997), we can solve $A\Phi(x) + f(x) = 0$ for $x \in (0, U)$, and show that

$$\Phi(x) = ax^{\alpha_1} + bx^{\alpha_2} + \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) x^2,$$

where $\alpha_1 := \frac{\frac{1}{2}\sigma^2 - \mu + \sqrt{\left(\mu - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2 r}}{\sigma^2}$, $\alpha_2 := \frac{\frac{1}{2}\sigma^2 - \mu - \sqrt{\left(\mu - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2 r}}{\sigma^2}$, and a and b are constants to be determined. Observe that $\Phi(x)$ above should be no greater than the cost of no throttling policy for any $x \in (0, U)$. From (6), we can compute the cost of no throttling policy (denoted by $C^{(0)}(x)$) as follows:

$$\begin{aligned} C^{(0)}(x) &= \mathbb{E}\left[\int_0^\infty e^{-rt} f(X_x^{(\nu)}(t))dt\right] \\ &= \int_0^\infty e^{-rt}\beta\, \mathbb{E}\left[(X_x^{(\nu)}(t))^2\right] dt \\ &= \int_0^\infty \beta\, e^{-rt} x^2 e^{(2\mu+\sigma^2)t} dt \\ &= \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) x^2, \end{aligned}$$

where the third equality is due to (8). Moreover, from (iii) of Definition 2, we can deduce that $(r - \sigma^2 - 2\mu) > 0$. Therefore, we have

$$0 \leq ax^{\alpha_1} + bx^{\alpha_2} + \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) x^2 \leq \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) x^2, \tag{10}$$

for any $x \in (0, U)$. Note from their definitions that $\alpha_1 > 0$ and $\alpha_2 < 0$. Now, letting $x \to 0$ in the inequality on the right-hand side of (10), we deduce that $b \leq 0$; however, taking the limit $x \to 0$ in the inequality on the left-hand side of (10), we conclude that $b = 0$. Then, it is easily seen that $a \leq 0$. Consequently, for $x \in (0, U)$,

$$\Phi(x) = ax^{\alpha_1} + \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) x^2. \tag{11}$$

We now turn our attention to the throttling region $[U, \infty)$, where $\Phi(x) = \mathcal{M}\Phi(x)$. In order to compute $\mathcal{M}\Phi(x)$ explicitly and for analytical tractability, we assume that T follows an exponential distribution with rate $\lambda > 0$ (or mean λ^{-1}), i.e., $F_T(t) = 1 - e^{\lambda t}$ for $t > 0$. This assumption is reasonable in our setting. In particular, observe that $Prob(T > t_0) < 0.0025 \approx 0$ for $t_0 > 6/\lambda$. Thus, the probability of throttling for a longer period is very low under the exponential distribution. Since no additional throttling is allowed during a throttling period, short-lived throttling periods are desirable for our setting.

Now, define

$$R(x,\gamma) := \int_0^\infty \mathbb{E}\left[\int_0^T e^{-rs} f((1-\gamma)\tilde{X}_x(s))ds + e^{-rT}\Phi(\tilde{X}_x(T)) \middle| T = t\right] dF_T(t).$$

Then, using (9) and (11), it can be shown (see Appendix B) that

$$\begin{aligned} R(x,\gamma) = {} & \frac{\beta(1-\gamma)^2 x^2}{[r + \lambda - 2(\mu - \tilde{\mu}\gamma) - \tilde{\sigma}^2]} + \frac{a\lambda x^{\alpha_1}}{[r + \lambda - (\mu - \tilde{\mu}\gamma)\alpha_1 + \frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1 - 1)]} \\ & + \frac{\beta\lambda x^2}{(r - \sigma^2 - 2\mu)[r + \lambda - 2(\mu - \tilde{\mu}\gamma) - \tilde{\sigma}^2]}. \end{aligned}$$

Consequently, we have

$$\Phi(x) = \begin{cases} ax^{\alpha_1} + \left(\frac{\beta}{r-\sigma^2-2\mu}\right) x^2, & \text{if } x \in (0, U), \\ K + c\Upsilon x + R(U, \Upsilon), & \text{if } x \in [U, \infty). \end{cases} \tag{12}$$

Using standard 'principle of smooth fit' at the boundary U of the continuation region, we derive the following two equations

$$aU^{\alpha_1} + \left(\frac{\beta}{r - \sigma^2 - 2\mu}\right) U^2 = K + c\Upsilon U + R(U, \Upsilon), \text{ and} \tag{13}$$

$$a\alpha_1 U^{\alpha_1 - 1} + \left(\frac{2\beta}{r - \sigma^2 - 2\mu}\right) U = c\Upsilon. \tag{14}$$

Moreover, we know that the function $G(\gamma) := K + c\gamma U + R(U, \gamma)$ should be minimized at $\gamma = \Upsilon$. Then, using the first order condition $G'(\gamma) = 0$, we derive

$$\frac{\partial}{\partial\gamma} R(U, \gamma) \mid_{\gamma=\Upsilon} = -cU. \tag{15}$$

We need to solve the non-linear system (13)-(15) to compute the values of U, Υ and a. The parameters U and Υ uniquely determine our optimal policy. While it is customary to prove that there exists a unique solution for the non-linear system (13)-(15), we now briefly explain why it is not feasible to prove this result analytically in our case.

The existence and uniqueness of a solution for a non-linear system such as ours have been studied in the extant literature whenever it is analytically feasible to do so (Constantinides and Richard 1978, Øksendal 1999, Ohnishi and Tsujimura 2006, Perera and Buckley 2017, and the references therein); otherwise, the authors rely on numerical results (Cadenillas and Zapatero 1999). In particular, Ohnishi and Tsujimura (2006) employ techniques from Øksendal (1999) to prove the existence and uniqueness of a solution for a system when the underlying stochastic process follows a geometric Brownian process, as in our case. One common theme among all of these proofs is the 'separability' of the policy parameters in the system [For example, in Ohnishi and Tsujimura (2006), for a given $A_1 < 0$, parameters b and β are separable in equation (3.26) of the system (3.26)-(3.28).]. However, in our system (13)-(15), for a given $a < 0$, parameters U and Υ are not separable in both (13) and (15); this is because the growth rate during the throttling period depends on Υ. Consequently, it is not feasible to analytically prove the existence and uniqueness of a solution for our system. Therefore, we rely on extensive numerical simulations to address the issues of existence and uniqueness. Fortunately, as will be seen in Section 5, our computer program produces consistent and meaningful solutions for our system (13)-(15) under all feasible scenarios even though we occasionally had to impose obvious constraints such as the positivity of Υ.

Before we present our numerical results, we provide the conditions under which a solution Φ of the system (12)-(15) solves the QVI defined in Proposition 1. The proof of this result is presented in Appendix C.

Proposition 2. *Let Φ be a continuously differentiable solution of* (12)-(15) *such that*

$$U > \frac{-(\mu - r)c\Upsilon + \left\{(\mu - r)^2 c^2 \Upsilon^2 + 4\beta r[K + R(U, \Upsilon)]\right\}^{\frac{1}{2}}}{2\beta} \tag{16}$$

and

$$\Phi(x) < K + \inf_{0<\gamma\leq 1} \{cx\gamma + R(x, \gamma)\}, \tag{17}$$

for $x \in (0, U)$. Then, Φ is a solution to the QVI defined in Proposition 1.

5. NUMERICAL SIMULATIONS AND THEIR IMPLICATIONS

We use Newton's method to solve the non-linear system (13)-(15) in MATLAB [`https://www.mathworks.com/help/optim/ug/fsolve.html`]. Considering the admissibility constraints, convergence, existence, and uniqueness of the optimal solutions, we use the following parameters in our baseline model: $\mu = 0.01$, $\sigma = 0.10$, $\tilde{\sigma} = 0.15$, $\tilde{\mu} = 0.02$, $\beta = 1.0$, $r = 0.06$, $c = 2.0$, $K = 20$ and $\lambda = 0.5$. Using our MATLAB program, we derive the numerical approximations (rounded to three decimal places) for U, Υ, and a as:

$$U = 3.178, \ \Upsilon = 0.735, \ a = -6.944.$$

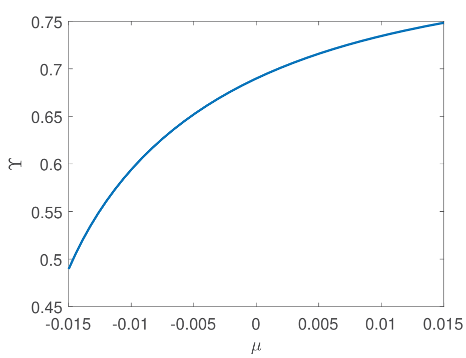
The inbuilt function *fsolve* in MATLAB also provides the range of errors for the system (13)-(15) and it is easy to verify that the conditions in the system of equations are satisfied within the default tolerance limit of 1×10^{-3} for the function. We obtain the square root of the sum of the square of errors in the three equations, viz. (13)-(15), as $2.15 \times e^{-19} \ll 1 \times 10^{-3}$.

The above numerical solution suggests that the OSP should start throttling the demand when it reaches the level $U = 3.178$. Moreover, intensity $\Upsilon = 0.735$ implies that the speed of delivery should be reduced by 73.5% during the throttling period.

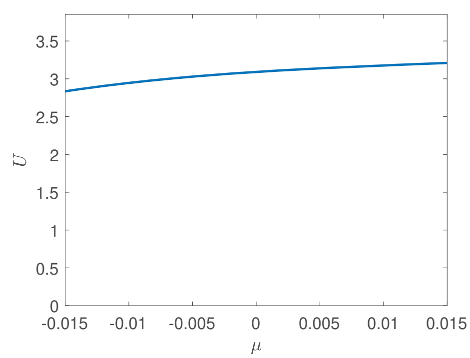
One of the main advantages of our efficient MATLAB program is that we can use it to solve multiple instances of the problem, and thereby, perform a comparative statics analysis to derive some interesting operational and managerial implications. In our comparative statics analysis, we use the parameters from the baseline case described above and change one parameter at a time to study the impact of the change in that parameter. The rest of this section is devoted to this comparative statics analysis.

5.1 The Impact of Change in the Growth Rate and Volatility of Unthrottled Demand

Figure 2(a) indicates that the throttling factor increases as μ increases. This means that, in order to alleviate higher growth rates, OSPs should employ severe throttling. It should also be noted from Figure 2(b) that U is slightly increasing in μ implying that the OSP will be somewhat patient before taking such severe actions against higher growth rates that can lead to network congestion.



(a) Υ versus μ



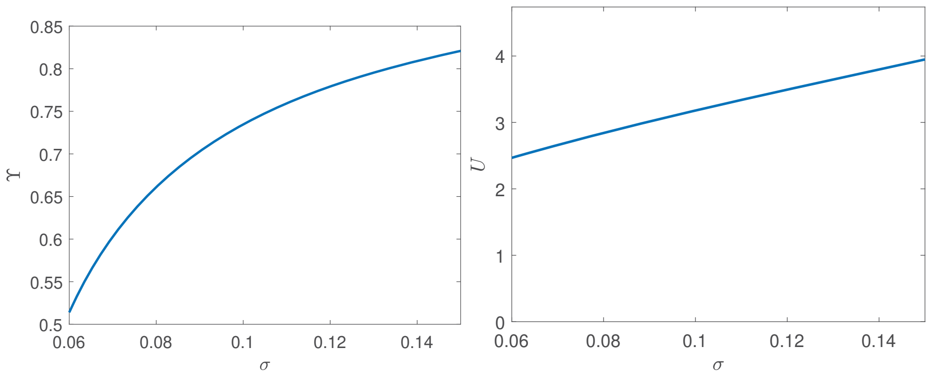
(b) U versus μ

Figure 2: Optimal throttling factor and threshold level as μ increases.

It is evident from Figure 3 that the impact of an increase in volatility, σ is similar to an increase in μ. In both of these cases, we also notice that the throttling factor is concave and increasing in the respective parameters. Thus, the marginal impact of throttling increases with its intensity.

5.2 The Impact of Change in the Dynamics of Demand during Throttling Period

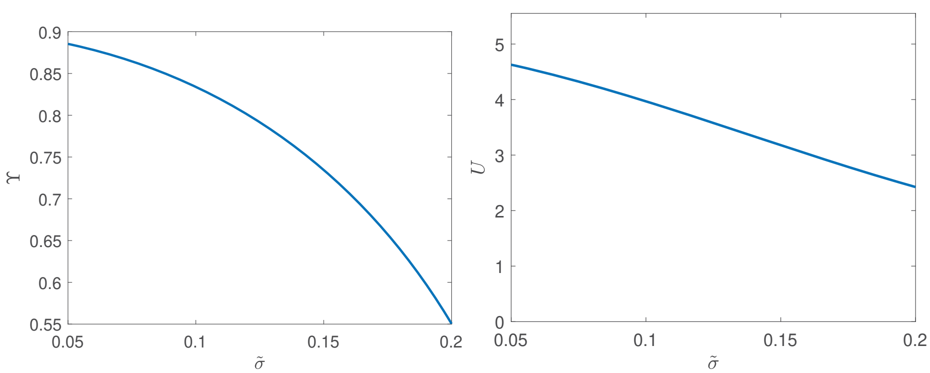
During the throttling period, the demand is driven by two parameters $\tilde{\mu}$ and $\tilde{\sigma}$. Figure 4 indicates that both Υ and U are decreasing in $\tilde{\sigma}$. From the impulse control literature (*cf.* Cadenillas and Zapatero 1999, Cadenillas et al. 2010), we know that, the OSP should wait longer to throttle



(a) Υ versus σ (b) U versus σ

Figure 3: Optimal throttling factor and threshold level as σ increases.

(intervene) as the volatility increases; moreover, the throttling factor is expected to be increasing in volatility. This is also what we observed in Section 5.1. In comparison with that traditional intuition, our observation sounds somewhat counter-intuitive and it is worth pursuing further discussion.

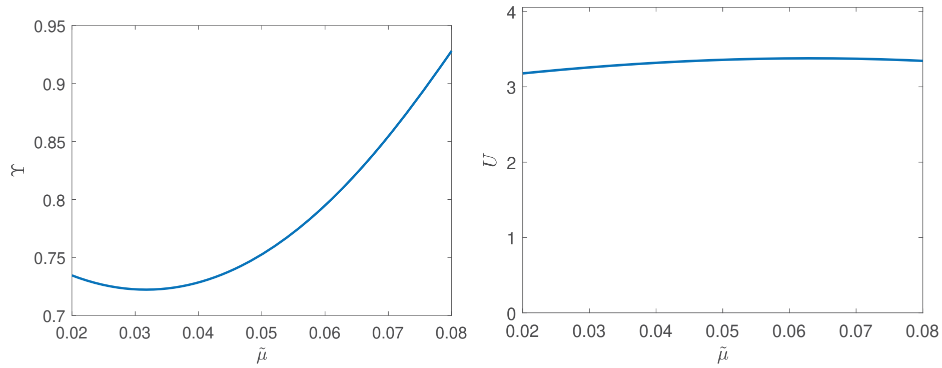


(a) Υ versus $\tilde{\sigma}$ (b) U versus $\tilde{\sigma}$

Figure 4: Optimal throttling factor and threshold level as $\tilde{\sigma}$ increases.

In our setting, $\tilde{\sigma}$ represents the volatility of the throttled demand. Thus, a higher value of $\tilde{\sigma}$ implies that users are more sensitive to throttling. If the OSP knows that its users are very sensitive to throttling, then it would lessen the strength of throttling; this indirectly forces the OSP to throttle more frequently implying a lower value of U. Finally, this observation is consistent with our model in which the sensitivity of users to throttling depends on the strength of throttling as opposed to the frequency of throttling.

We now turn our attention to $\tilde{\mu}$. While the value of U is fairly stable as a function of $\tilde{\mu}$ (see Figure 5(b)), the throttling factor is initially decreasing and then increasing as a function of $\tilde{\mu}$ (see 5(a)). However, it should be noted that the growth rate of the demand process (2) during the throttling period is $\mu - \tilde{\mu}\Upsilon$. From Figure 6, it is easily seen that the optimal throttling factor is selected such that this growth rate is negative and decreasing in $\tilde{\mu}$. Thus, optimal throttling not only helps the OSP to manage demand during the throttling period, it also helps the OSP to force out more reactive users (with a larger $\tilde{\mu}$) from its service.



(a) Υ versus $\tilde{\mu}$ (b) U versus $\tilde{\mu}$

Figure 5: Optimal throttling factor and threshold level as $\tilde{\mu}$ increases.

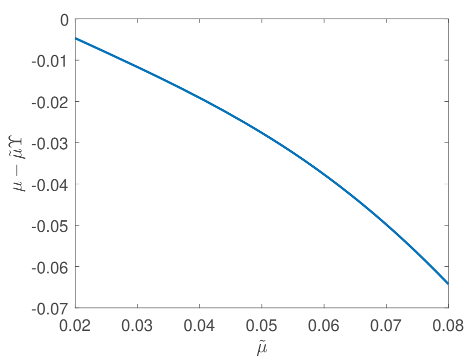
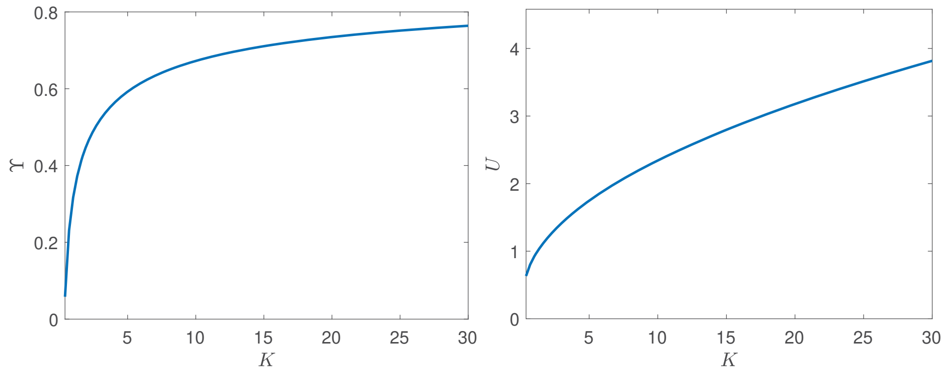


Figure 6: The growth rate of the demand during the throttling period as a function of $\tilde{\mu}$.

5.3 The Impact of Change in Cost Parameters and Discount Rate

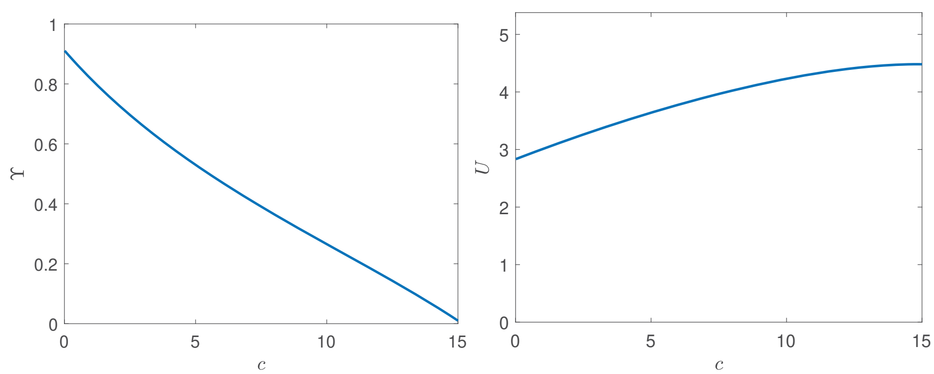
These results are parallel with the extant literature (*cf.* Bensoussan et al. 2012, Cadenillas et al. 2010, Cadenillas and Zapatero 1999, 2000) while their interpretations are different. First, we note from Figure 7 that both Υ and U are increasing in K. This means that it is optimal for the OSP to wait longer if the fixed cost is higher and then employ a stronger throttling. This fixed cost represents one-time costs such as cost of damaging the reputation of the OSP. Thus, it makes perfect sense for the OSP to wait and then apply stronger throttling if this fixed cost is higher. Conversely, the proportional cost parameter c, in general, represents the direct cost associated with the impact of throttling intensity on the current users negative experience due to throttling. Thus, for a larger value of c, the OSP should employ throttling with less intensity and be patient before throttling again. This phenomenon is clearly illustrated in Figure 8.

The cost of maintaining the demand should be higher for a larger value of β at a given level of demand. Thus, for a larger value of β, we expect a smaller U and a higher Υ. This is exactly what we observe in Figure 9. As we explained in our model, the cost of maintaining accommodates for the potential risks associated with the server congestion. Therefore, the higher



(a) Υ versus K (b) U versus K

Figure 7: Optimal throttling factor and threshold level as K increases.



(a) Υ versus c (b) U versus c

Figure 8: Optimal throttling factor and threshold level as c increases.

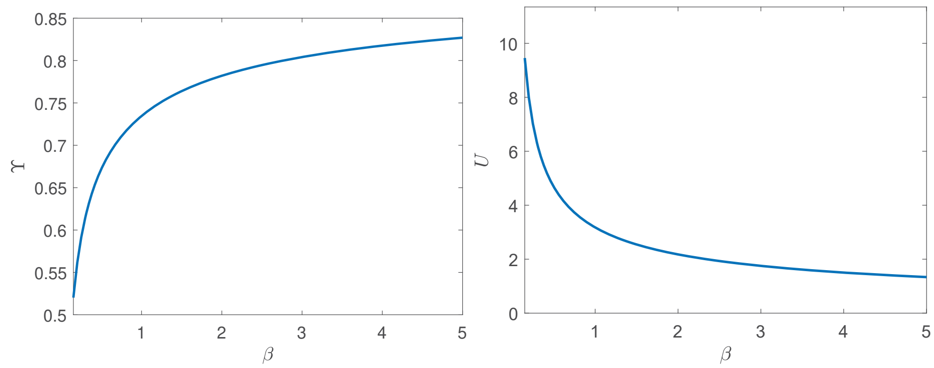
the value of β, the higher the risk; this is exactly why OSPs impose severe and more frequent throttling for larger values of β as a *risk-minimizing strategy*. In Section 6.2, we explain how we can use this observation to handle congestion management problems with capacity limitations.

Finally, the discount rate r basically represents the inflation rate. Thus, when the value of r increases, discounted future cost of throttling becomes cheaper. Hence, it is optimal for OSP to postpone throttling (see Figure 10(b)). Moreover, the intensity of these delayed throttling will be slightly less (see Figure 10(a)).

6. DISCUSSION AND ADDITIONAL REMARKS

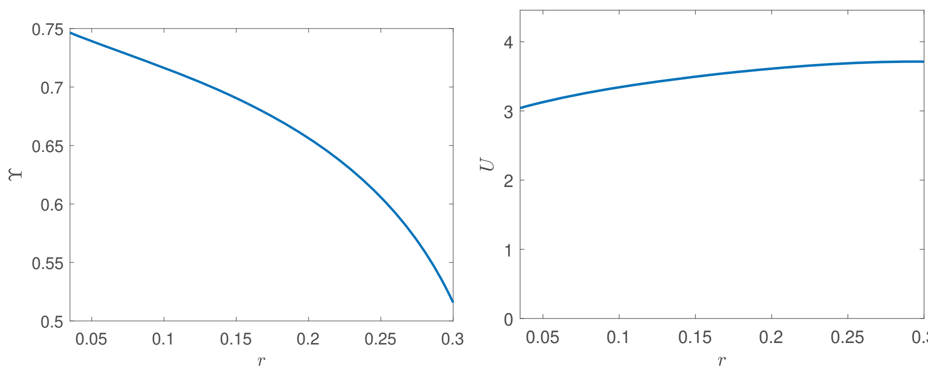
6.1 Long-term Effect of Demand Throttling

From Figure 6, we noticed that the growth rate of the demand during the throttling period (i.e., $\tilde{\mu}_g = \mu - \tilde{\mu}\Upsilon$) is negative. We further investigated this phenomenon using an extensive numerical study, and our study confirmed that $\tilde{\mu}_g < 0$ always holds at the *optimum*. We have presented a



(a) Υ versus β (b) U versus β

Figure 9: Optimal throttling factor and threshold level as β increases.



(a) Υ versus r (b) U versus r

Figure 10: Optimal throttling factor and threshold level as r increases.

part of our results in Table 1 below and included a comprehensive summary of other simulation results in Appendix D. Therefore, as suggested by our numerical study, it is optimal for the OSP to implement throttling so that the growth rate during the throttling period is negative. (Recall that from (2), it is feasible to have a positive growth rate during the throttling period.) Thus, OSP uses demand throttling not only to manage the congestion while throttling is in place, but also to force away unhealthy future demand from the system.

Table 1: Value of $-\tilde{\mu}_g$ (rounded to three decimal places) for different values of μ and $\tilde{\mu}$.

$\mu \downarrow \quad \tilde{\mu} \rightarrow$	0.02	0.04	0.06	0.08
-0.01	0.022	0.032	0.049	0.080
0.00	0.014	0.027	0.046	0.073
0.01	0.004	0.010	0.011	0.011

6.2 Incorporating Capacity Constraints

The capacity of a server is limited, and the risks associated with congestion is higher when the usage is higher and closer to the capacity. While our demand throttling model in Section 3 does not handle capacity directly, the capacity is incorporated in our definition of the maintenance cost. As we explained earlier, our maintenance cost represents the costs associated with risks such as server crashing risk, queueing delay risk, packet loss risk and the risk of blocking of new connections. If the capacity is lower, then the cost associated with these risks should be higher implying a higher value of β; moreover, infinite capacity does not mean zero risk either. For example, one could assume that

$$\beta(C) := \kappa_1 + \kappa_2/\eta(C), \tag{18}$$

where κ_1 and κ_2 are some constants such that $\beta(C) > 0$, and $\eta(\cdot)$ is a non-negative increasing function of C. However, for applications, properly estimating a closed-form expression such as the one in (18) could be challenging. Moreover, mathematical analysis of such an explicit formulation will be more intricate without adding further insights. Nevertheless, it should be noted that if the capacity C of the server is known, then one could incorporate the capacity as a constraint to our congestion management problem through β. We illustrate this idea by using a simple example.

Consider our baseline model introduced at the beginning of Section 5 with capacity C. We also redefine β as a function of capacity via $\beta(C) := 15/(C + 1)$. Thus, using the notation of (18) above, we have $\kappa_1 = 0$, $\kappa_2 = 15$ and $\eta(C) = C + 1$. Since $C = -1 + 15/\beta$ (with $\beta < 15$) decreases as β increases, from Figure 9(a), we can deduce that the throttling intensity Υ must be higher with tighter capacities. Moreover, Figure 11 compares optimal threshold (U) values with the corresponding capacities at each value of β; it is easily seen from the figure that it is optimal to stay below and follow the trajectory of the capacity in our setting.

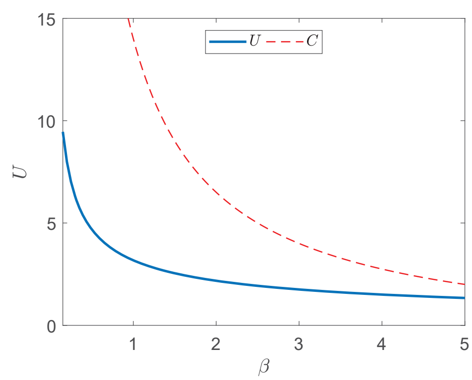


Figure 11: Optimal threshold values (U) with corresponding capacity constraints (C).

6.3 Trade-off Between Internal and External Costs

We also know, from Section 5.3, that higher fixed or proportional throttling costs always imply waiting longer before throttling. However, the throttling factor is increasing in K and decreasing in c. Thus, if the fixed cost is relatively higher than the proportional cost, then the optimal throttling factor should be higher. Consequently, whenever factors corresponding to fixed costs outweigh those associated with proportional costs, throttling should be stronger. For example, if the one-time cost of damaging the OSP's reputation outweighs the 'non-monetary' cost of making its current users unhappy, then the OSP should prefer throttling with higher intensities. Here, we should notice that factors associated with the fixed cost are *external* and do not have a direct connection with the user-experience of the current customers, whereas the proportional cost is *internal* and directly impacts the user-experience. In summary, our results suggest that the OSP should carefully evaluate the trade-offs between these external and internal costs, and the factors associated with these costs while making their throttling decisions.

7. CONCLUSION AND FURTHER DIRECTIONS

We have modeled a congestion management problem of an online service provider (OSP) as a stochastic impulse control problem and proved that the optimal strategy is to initiate throttling of a fixed intensity for a period of time whenever demand reaches a certain threshold. In particular, we have incorporated users' reactions to throttling in our model. Our numerical results suggest that service providers should employ users' willingness to exit the service, as a result of throttling, to regulate the future demand by always imposing negative growth rates during the throttling periods. This is the first stochastic model that handles the congestion management problem of an online service provider while preserving net-neutrality.

OSPs in some industries including energy distribution, cloud services, and video streaming (Fukuda et al. 2018, Deng et al. 2018, Dede et al. 2018, respectively) are more vulnerable to bulk and sporadic demand. Moreover, this type of demand could cause crashing more often than regular demand. To capture the dynamics of such demand, one could use a geometric jump diffusion process as the underlying stochastic process. While our analysis may be extended to handle this scenario, the corresponding analysis would be more intricate and will be a challenging problem for future research.

Although bandwidth throttling applies to OSPs only, demand throttling is a commonly observed practice in other businesses. For example, energy companies use programs such as 'Peak Power Saver'® (`https://peakpowersavers.com/`) program to manage their demand during the peak times in summer and winter (Consumers Energy 2018). Recently, under their Peak Power Savers® programs, Consumers Energy offered their customer $25 VISA gift card plus annual bill credit worth more than $30 for signing up for their 'AC Peak Cycling program'. When a customer signs up for this program, Consumers Energy installs a switch on (or near) the outside central air conditioning unit of the customer's household free of charge. On select summer weekdays during peak energy demand, the switch will be activated—reducing the output of the AC unit by a certain percentage for a short period of time. Our framework could be adopted to model this electricity throttling problem and other similar applications in the future. Thus, the scope of applications of our model is indeed broad.

A. VERIFICATION THEOREM AND ITS PROOF

A verification theorem for our congestion management problem is presented here. We first need to define the QVI-control associated with the QVI.

Definition 3 (QVI-control)**.** *Let ϕ be a solution of the QVI presented in Proposition 1 and*

$$M(\phi, y, \gamma) := K + c\gamma y + \int_0^\infty \mathbb{E}\left[\int_0^T e^{-rs} f\left((1-\gamma)\tilde{X}_y(s)\right) ds + e^{-rT}\phi\left(\tilde{X}_y(T)\right)\Big| T = t\right] dF_T(t).$$

Then, the QVI-control $\hat{\nu} = (\hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_i, \dots; \hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_i, \dots)$ (when it exists) is inductively defined by

$$\begin{aligned} \hat{\tau}_0 &= 0, \\ \hat{\tau}_{i+1} &= \inf\{t > \hat{\tau}_i + T_i; X_x^{(\hat{\nu}_i)}(t-) \notin \mathbb{C}\}, \text{ and} \\ \hat{\gamma}_{i+1} &= \arg\inf\left\{M(\phi, X_x^{(\hat{\nu}_i)}(\hat{\tau}_{i+1}-), \gamma);\ \gamma \in (0,1]\right\}, \end{aligned}$$

where $X_x^{(\hat{\nu}_i)}$ is the result of applying $\hat{u}_i = (\hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_i; \hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_i)$ to X_x.

Theorem 1 (Verification Theorem)**.** *Suppose that a solution of the QVI given in Proposition 1 also has continuous second order derivatives on $(0,\infty) \setminus \{x_0\}$ for some $x_0 > 0$, and the second order derivatives of ϕ are locally bounded near $x = x_0$. Moreover, assume that for $x > 0$, we have*

(I) QVI-control $\hat{\nu}$ corresponding to ϕ is admissible,
(II) $\lim_{t\to\infty} E\left[e^{-rt}\phi\left(X_x^{(\hat{\nu})}(t)\right)\right] = 0$,
(III) $\{\phi(X_x^{(\hat{\nu})}(\tau));\ \tau$ is a stopping time$\}$ is uniformly integrable, and
(IV) $\hat{\gamma} = \arg\inf\{M(\phi, x, \gamma);\ \gamma \in (0,1]\}$ is a Borel measurable selection,
then $\phi(x) = \Phi(x)$ and $\hat{\nu} = \nu^*$ for all $x > 0$.

Proof: Since $\phi \in C^1$ with continuous second order derivatives on $(0,\infty) \setminus \{x_0\}$ and the second order derivatives of ϕ are locally bounded near $x = x_0$, using a well-known approximation result (*cf.* Appendix D of Øksendal 2003), we can simply assume that $\phi \in C^2$ throughout our proof. Now, let

$$\phi_d\left(X_x^{(\hat{\nu})}(s)\right) := e^{-rs}\phi\left(X_x^{(\hat{\nu})}(s)\right)$$

for $s \geq 0$. Then, since T_i is independent of $X_x^{(\hat{\nu})}$ for $i \geq 1$, we have

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + T_i)\right)\Big| T_i = t\right] = E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + t)\right)\right], \tag{19}$$

where $t \geq 0$.

We use the Itô's formula (Øksendal 2003) to derive an expression for the right-hand side of (19). First, consider

$$g(s, X_x^{(\hat{\nu})}(s)) := e^{-rs}\phi\left(X_x^{(\hat{\nu})}(s)\right) = \phi_d\left(X_x^{(\hat{\nu})}(s)\right)$$

and note that $g(s,y)$ is C^1 w.r.t. $s \geq 0$ (i.e., for $y > 0$, $g(\cdot, y)$ is continuous on $[0,\infty)$ with continuous first order derivatives) and C^2 w.r.t. $y > 0$ (i.e., for $s \geq 0$, $g(s,\cdot)$ is continuous on $(0,\infty)$

with continuous second order derivatives). Thus, we can apply the Itô's formula for g from τ_i to $\tau_i + t$, and obtain [Note that, from τ_j to $\tau_j + t$, the effective-process $X^{(\hat{\nu})}$ follows $(1 - \gamma_j)\tilde{X}$.]

$$\begin{aligned}\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + t)\right) &= \phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right) - r\int_{\tau_i}^{\tau_i+t} \phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds \\ &+ \int_{\tau_i}^{\tau_i+t} \frac{\partial \phi_d}{\partial x}\left(X_x^{(\hat{\nu})}(s)\right)[(\mu - \tilde{\mu}\gamma_i)ds + \tilde{\sigma} dB(s)] \\ &+ \frac{1}{2}\tilde{\sigma}^2 \int_{\tau_i}^{\tau_i+t} \frac{\partial^2 \phi_d}{\partial x^2}\left(X_x^{(\hat{\nu})}(s)\right) ds. \qquad (20)\end{aligned}$$

For $x > 0$, we define the differential operator $\tilde{A}$ for the i-th throttling period as follows:

$$\tilde{A}\phi_d(x) := (\mu - \tilde{\mu}\gamma_i)\frac{\partial \phi_d}{\partial x}(x) + \frac{1}{2}\tilde{\sigma}^2\frac{\partial^2 \phi_d}{\partial x^2}(x) - r\phi_d(x).$$

Then, from (20), we deduce that

$$\begin{aligned}\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + t)\right) &= \phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right) + \tilde{\sigma}\int_{\tau_i}^{\tau_i+t} \frac{\partial \phi_d}{\partial x}\left(X_x^{(\hat{\nu})}(t)\right) dB(s) \\ &+ \int_{\tau_i}^{\tau_i+t} \tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds. \qquad (21)\end{aligned}$$

Since $E\left[\int_{\tau_i}^{\tau_i+t} \frac{\partial \phi_d}{\partial x}\left(X^{(\hat{\nu})}(t)\right) dB(s)\right] = 0$, taking the expectation of both sides of (21), we have

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + t)\right)\right] = E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)\right] + E\left[\int_{\tau_i}^{\tau_i+t} \tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right]. \qquad (22)$$

Using a similar argument from $\tau_j + t$ to τ_{i+1}, we can derive an expression for $E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_{i+1}-)\right)\right]$. Here, it must be noted that the process $X^{(\hat{\nu})}$ follows the original process X from $\tau_j + t$ to τ_{i+1}. Consequently, we can easily derive that

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_{i+1}-)\right)\right] = E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i + t)\right)\right] + E\left[\int_{\tau_i+t}^{\tau_{i+1}} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right]. \qquad (23)$$

Combining (22) and (23) yields

$$\begin{aligned}&E[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right) - \phi_d\left(X_x^{(\hat{\nu})}(\tau_{i+1}-)\right)] = \\ &\qquad - E\left[\int_{\tau_i}^{\tau_i+t} \tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right] - E\left[\int_{\tau_i+t}^{\tau_{i+1}} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right].\end{aligned}$$

Summing the above equation from $i = 1$ to $i = n$, we get

$$\begin{aligned}&\sum_{i=1}^{n} E[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right) - \phi_d\left(X_x^{(\hat{\nu})}(\tau_{i+1}-)\right)] = \\ &\qquad - E\left[\sum_{i=1}^{n}\int_{\tau_i}^{\tau_i+t} \tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right] - E\left[\sum_{i=1}^{n}\int_{\tau_i+t}^{\tau_{i+1}} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\right]. \qquad (24)\end{aligned}$$

Applying an argument similar to that leading up to (22), it can be easily shown that

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_1-)\right)\right] = E\left[\phi_d\left(X_x^{(\hat{\nu})}(0)\right)\right] + E\left[\int_0^{\tau_1} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right].$$

However, since $E[\phi_d\left(X_x^{(\hat{\nu})}(0)\right)] = E[e^{-r(0)}\phi(X_x^{(\hat{\nu})}(0))] = E[\phi(x)] = \phi(x)$, we deduce that

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_1-)\right)\right] = \phi(x) + E\left[\int_0^{\tau_1} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right]. \tag{25}$$

Combining (24) and (25), and rearranging the terms yield the following equation.

$$\begin{aligned}\phi(x) \quad &+ \quad \sum_{i=1}^{m} E\left[\phi_d(X_x^{(\hat{\nu})}(\tau_i)) - \phi_d\left(X_x^{(\hat{\nu})}(\tau_{i+1}-)\right)\right] - E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_1-)\right)\right] \\ &= \quad -E\left[\sum_{i=1}^{m}\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] - E\left[\int_0^{\tau_1} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] \\ &\qquad -E\left[\sum_{i=1}^{m}\int_{\tau_i+t}^{\tau_{i+1}} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right].\end{aligned}$$

Equivalently, we have

$$\begin{aligned}\phi(x) \quad &+ \quad \sum_{i=1}^{m} E\left[\phi_d(X_x^{(\hat{\nu})}(\tau_i)) - \phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right] - E\left[\phi_d(X_x^{(\hat{\nu})}(\tau_{m+1}-))\right] \\ &= \quad -E\left[\sum_{i=1}^{m}\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] - E\left[\int_0^{\tau_1} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] \\ &\qquad -E\left[\sum_{i=1}^{m}\int_{\tau_i+t}^{\tau_{i+1}} A\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right].\end{aligned} \tag{26}$$

Next, observe that the throttled process $\{X_x^{(\hat{\nu})}(s) : s \geq 0\}$ stays either within the continuation region $\mathbb{C}$ or on its boundary. Thus, for all $s \geq 0$, we have

$$f(X_x^{(\hat{\nu})}(s)) + A\phi(X_x^{(\hat{\nu})}(s)) = 0.$$

Multiplying the above equation by e^{-rs}, we obtain

$$f_d(X_x^{(\hat{\nu})}(s)) + A\phi_d(X_x^{(\hat{\nu})}(s)) = 0,$$

where $f_d(X^{(\hat{\nu})}(s)) = e^{-rs}f(X^{(\hat{\nu})}(s))$. Substituting this in (26) yields

$$\begin{aligned}\phi(x) \quad &+ \quad \sum_{i=1}^{m} E\left[\phi_d(X_x^{(\hat{\nu})}(\tau_i)) - \phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right] - E\left[\phi_d(X_x^{(\hat{\nu})}(\tau_{m+1}-))\right] \\ &= \quad -E\left[\sum_{i=1}^{m}\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] + E\left[\int_0^{\tau_1} f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right] \\ &\qquad +E\left[\sum_{i=1}^{m}\int_{\tau_i+t}^{\tau_{i+1}} f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right].\end{aligned} \tag{27}$$

From Condition (IV) and the definitions of $\mathcal{M}$ and $\hat{\nu}$, it follows that

$$\mathcal{M}\phi(X_x^{(\hat{\nu})}(\tau_i-)) = K + c\gamma_i X_x^{(\hat{\nu})}(\tau_i-) + \int_0^\infty \mathbb{E}\left[\int_0^t e^{-rs} f\left((1-\gamma_i)\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(s)\right) ds + e^{-rt}\phi\left(\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(t)\right)\right] dF_T(t)$$

Rearranging the right-hand side, we get

$$\begin{aligned}\mathcal{M}\phi(X_x^{(\hat{\nu})}(\tau_i-)) &= K + c\gamma_i X_x^{(\hat{\nu})}(\tau_i-) \\ &\quad + \int_0^\infty \left\{\int_0^t e^{-rs}\mathbb{E}\left[f\left((1-\gamma_i)\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(s)\right)\right] ds\right\} dF_T(t) \\ &\quad + \int_0^\infty \mathbb{E}\left[e^{-rt}\phi\left(\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(t)\right)\right] dF_T(t).\end{aligned} \tag{28}$$

We next derive alternative expressions for the last two terms of (28). For this, note that $X_x^{(\hat{\nu})}(s) = (1-\gamma_i)\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(s)$ for $\tau_i \le s \le \tau_i + t$. Then,

$$\begin{aligned}&\int_0^\infty \left\{\int_0^t e^{-rs}\mathbb{E}\left[f\left((1-\gamma_i)\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(s)\right)\right] ds\right\} dF_T(t) \\ &= \int_0^\infty \left\{\int_0^t e^{-rs} E\left[f\left(X_x^{(\hat{\nu})}(\tau_i+s)\right)\Big|\,\mathcal{F}_{\tau_i}\right] ds\right\} dF_T(t) \\ &= \int_0^\infty E\left[\int_0^t e^{-rs} f\left(X_x^{(\hat{\nu})}(\tau_i+s)\right) ds\Big|\mathcal{F}_{\tau_i}\right] dF_T(t) \\ &= \int_0^\infty e^{r\tau_i} E\left[\int_{\tau_i}^{\tau_i+t} f_d\left(X_x^{(\hat{\nu})}(s)\right) ds\Big|\mathcal{F}_{\tau_i}\right] dF_T(t),\end{aligned} \tag{29}$$

where the first equality above is due to the strong Markov property and the last equality is due to the $\mathcal{F}_{\tau_i}$-measurability of $e^{r\tau_i}$. Moreover, note that $X_x^{(\hat{\nu})}(s) = \tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(s)$ when $s = \tau_i + t$. Then, using the strong Markov property again, we obtain

$$\begin{aligned}\int_0^\infty \mathbb{E}\left[e^{-rt}\phi\left(\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(t)\right)\right] dF_T(t) &= \int_0^\infty E\left[e^{-rt}\phi\left(\tilde{X}_{X_x^{(\hat{\nu})}(\tau_i-)}(\tau_i+t)\right)\Big|\,\mathcal{F}_{\tau_i}\right] dF_T(t) \\ &= \int_0^\infty E\left[e^{-rt}\phi\left(X_x^{(\hat{\nu})}(\tau_i+t)\right)\Big|\,\mathcal{F}_{\tau_i}\right] dF_T(t).\end{aligned} \tag{30}$$

Substituting the alternative expressions derived in (29)-(30) into (28), and multiplying the resulting equation by $e^{-r\tau_i}$ yields

$$\begin{aligned}\mathcal{M}\phi_d(X_x^{(\hat{\nu})}(\tau_i-)) &= e^{-r\tau_i}\left\{K + c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\} \\ &\quad + \int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t} f_d\left(X_x^{(\hat{\nu})}(s)\right) ds\Big|\mathcal{F}_{\tau_i}\right] dF_T(t) \\ &\quad + \int_0^\infty E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i+t)\right)\Big|\,\mathcal{F}_{\tau_i}\right] dF_T(t).\end{aligned} \tag{31}$$

To derive an alternative expression for the last part of the above equation, we take the conditional expectation of the both sides of (21) and obtain that

$$\begin{aligned}E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i+t)\right)\Big|\,\mathcal{F}_{\tau_i}\right] &= E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)\Big|\,\mathcal{F}_{\tau_i}\right] + \tilde{\sigma}E\left[\int_{\tau_i}^{\tau_i+t}\frac{\partial\phi_d}{\partial x}\left(X_x^{(\hat{\nu})}(t)\right) dB(s)\Big|\mathcal{F}_{\tau_i}\right] \\ &\quad + E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right) ds\Big|\mathcal{F}_{\tau_i}\right].\end{aligned}$$

Since $\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)$ is $\mathcal{F}_{\tau_i}$-measurable and $E\left[\int_{\tau_i}^{\tau_i+t}\frac{\partial\phi_d}{\partial x}\left(X_x^{(\hat{\nu})}(t)\right)dB(s)\Big|\mathcal{F}_{\tau_i}\right]=0$, we deduce that

$$E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i+t)\right)\Big|\,\mathcal{F}_{\tau_i}\right]=\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)+E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\Big|\mathcal{F}_{\tau_i}\right].$$

Integrating the above equation with respect to the probability measure induced by T on $[0,\infty)$, and substituting the resulting equation in (31) yields

$$\begin{aligned}\mathcal{M}\phi_d(X_x^{(\hat{\nu})}(\tau_i-))=e^{-r\tau_i}&\left\{K+c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\\&+\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\Big|\mathcal{F}_{\tau_i}\right]dF_T(t)\\&+\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)+\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\Big|\mathcal{F}_{\tau_i}\right]dF_T(t).\end{aligned}$$

Subtracting $\phi_d(X_x^{(\hat{\nu})}(\tau_i-))$ from the both sides of the above equation and rearranging, we have

$$\begin{aligned}\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))&=\mathcal{M}\phi_d(X_x^{(\hat{\nu})}(\tau_i-))-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))\\&-e^{-r\tau_i}\left\{K+c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\Big|\mathcal{F}_{\tau_i}\right]dF_T(t)\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\Big|\mathcal{F}_{\tau_i}\right]dF_T(t).\end{aligned}$$

Now taking expectation and applying the double expectation formula, it can be deduced that

$$\begin{aligned}E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right]&=E\left[\mathcal{M}\phi_d(X_x^{(\hat{\nu})}(\tau_i-))-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right]\\&-E\left[e^{-r\tau_i}\left\{K+c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\right]\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right]dF_T(t)\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right]dF_T(t).\end{aligned}$$

Moreover, since $E\left[\mathcal{M}\phi_d(X_x^{(\hat{\nu})}(\tau_i-))-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right]=0$ on the boundary of the throttling region, we have

$$\begin{aligned}E\left[\phi_d\left(X_x^{(\hat{\nu})}(\tau_i)\right)-\phi_d(X_x^{(\hat{\nu})}(\tau_i-))\right]&=-E\left[e^{-r\tau_i}\left\{K+c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\right]\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}f_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right]dF_T(t)\\&-\int_0^\infty E\left[\int_{\tau_i}^{\tau_i+t}\tilde{A}\phi_d\left(X_x^{(\hat{\nu})}(s)\right)ds\right]dF_T(t).\end{aligned}$$

Finally, replacing the expression inside the second term on the left-hand side of (27) by the above expression and simplifying, we get

$$\phi(x) = E\left[\int_0^{\tau_{m+1}} e^{-rs} f(X_x^{(\hat{\nu})}(s))ds + \sum_{i=1}^{m} e^{-r\tau_i}\left\{K + c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\right] + E\left[e^{-r\tau_{m+1}}\phi(X_x^{(\hat{\nu})}(\tau_{m+1}-))\right].$$

Note that, Condition (II) implies that $E\left[e^{-r\tau_{m+1}}\phi(X_x^{(\hat{\nu})}(\tau_{m+1}-))\right] \to 0$ as $m \to \infty$. Thus, letting $m \to \infty$ in the above equation yields

$$\phi(x) = E\left[\int_0^{\infty} e^{-rs} f(X_x^{(\hat{\nu})}(s))ds + \sum_{i=1}^{\infty} e^{-r\tau_i}\left\{K + c\gamma_i X_x^{(\hat{\nu})}(\tau_i-)\right\}\right] = C^{(\hat{\nu})}(x).$$

Hence, $\phi(x) = \Phi(x)$ and $u^* = \hat{\nu}$. □

B. DERIVATION OF $R(X)$

Let

$$R(x;t) := \mathbb{E}\left[\int_0^t e^{-rs} f((1-\gamma)\tilde{X}_x(s))ds + e^{-rt}\Phi(\tilde{X}_x(t))\right],$$

then $R(x,\gamma) = \int_0^\infty R(x;t)\lambda e^{-\lambda t}dt$. We first compute $R(x;t)$:

$$\begin{aligned} R(x;t) &= \mathbb{E}\left[\int_0^t e^{-rs} f((1-\gamma)\tilde{X}_x(s))ds + e^{-rt}\Phi(\tilde{X}_x(t))\right] \\ &= \int_0^t e^{-rs}\mathbb{E}\left[f((1-\gamma)\tilde{X}_x(s))\right]ds + e^{-rt}\mathbb{E}\left[\Phi(\tilde{X}_x(t))\right] \\ &= \int_0^t e^{-rs}\mathbb{E}\left[\beta\left((1-\gamma)\tilde{X}_x(s)\right)^2\right]ds + e^{-rt}\mathbb{E}\left[a\tilde{X}_x(t)^{\alpha_1} + \left(\frac{\beta}{r-\sigma^2-2\mu}\right)\tilde{X}_x(t)^2\right] \\ &= \beta(1-\gamma)^2\int_0^t e^{-rs}\mathbb{E}\left[\tilde{X}_x(s)^2\right]ds + ae^{-rt}\mathbb{E}\left[\tilde{X}_x(t)^{\alpha_1}\right] + e^{-rT}\left(\frac{\beta}{r-\sigma^2-2\mu}\right)\mathbb{E}\left[\tilde{X}_x(t)^2\right] \end{aligned}$$

Now by applying (9), we get

$$\begin{aligned} R(x;t) &= \beta(1-\gamma)^2\int_0^t e^{-rs}x^2 e^{(2(\mu-\tilde{\mu}\gamma)+\tilde{\sigma}^2)s}ds + ae^{-rt}x^{\alpha_1}e^{\left((\mu-\tilde{\mu}\gamma)\alpha_1+\frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)\right)t} \\ &\quad + e^{-rt}\left(\frac{\beta}{r-\sigma^2-2\mu}\right)x^2 e^{(2(\mu-\tilde{\mu}\gamma)+\tilde{\sigma}^2)t} \\ &= \beta(1-\gamma)^2x^2\int_0^t e^{-\left(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2\right)s}ds + ax^{\alpha_1}e^{-\left(r-(\mu-\tilde{\mu}\gamma)\alpha_1+\frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)\right)t} \\ &\quad + \left(\frac{\beta x^2}{r-\sigma^2-2\mu}\right)e^{-(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2)t} \\ &= \beta(1-\gamma)^2x^2\frac{\left[1-e^{-\left(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2\right)t}\right]}{[r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2]} + ax^{\alpha_1}e^{-\left(r-(\mu-\tilde{\mu}\gamma)\alpha_1+\frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)\right)t} \\ &\quad + \left(\frac{\beta x^2}{r-\sigma^2-2\mu}\right)e^{-(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2)t}, \end{aligned}$$

where the last equality hold with $\left(r - 2(\mu - \tilde{\mu}\gamma) - \tilde{\sigma}^2\right) > 0$.

Finally, we can compute $R(x, \gamma)$ as follows:

$$\begin{aligned}
R(x,\gamma) &= \int_0^\infty R(x;t)\lambda e^{-\lambda t} dt \\
&= \int_0^\infty \beta(1-\gamma)^2 x^2 \frac{\left[1 - e^{-\left(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2\right)t}\right]}{[r - 2(\mu-\tilde{\mu}\gamma) - \tilde{\sigma}^2]} \lambda e^{-\lambda t} dt \\
&\quad + \int_0^\infty a x^{\alpha_1} e^{-\left(r-(\mu-\tilde{\mu}\gamma)\alpha_1 + \frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)\right)t} \lambda e^{-\lambda t} dt \\
&\qquad + \int_0^\infty \left(\frac{\beta x^2}{r - \sigma^2 - 2\mu}\right) e^{-(r-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2)t} \lambda e^{-\lambda t} dt \\
&= \frac{\beta\lambda(1-\gamma)^2 x^2}{[r - 2(\mu-\tilde{\mu}\gamma) - \tilde{\sigma}^2]} \int_0^\infty \left[e^{-\lambda t} - e^{-\left(r+\lambda-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2\right)t}\right] dt \\
&\quad + a\lambda x^{\alpha_1} \int_0^\infty e^{-\left(r+\lambda-(\mu-\tilde{\mu}\gamma)\alpha_1 + \frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)\right)t} dt \\
&\qquad + \left(\frac{\beta\lambda x^2}{r - \sigma^2 - 2\mu}\right) \int_0^\infty e^{-(r+\lambda-2(\mu-\tilde{\mu}\gamma)-\tilde{\sigma}^2)t} dt \\
&= \frac{\beta(1-\gamma)^2 x^2}{[r + \lambda - 2(\mu-\tilde{\mu}\gamma) - \tilde{\sigma}^2]} + \frac{a\lambda x^{\alpha_1}}{[r + \lambda - (\mu-\tilde{\mu}\gamma)\alpha_1 + \frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1-1)]} \\
&\qquad + \frac{\beta\lambda x^2}{(r - \sigma^2 - 2\mu)[r + \lambda - 2(\mu-\tilde{\mu}\gamma) - \tilde{\sigma}^2]},
\end{aligned}$$

where the last equality hold with $[r + \lambda - (\mu - \tilde{\mu}\gamma)\alpha_1 + \frac{1}{2}\tilde{\sigma}^2\alpha_1(\alpha_1 - 1)] > 0$.

C. PROOF OF PROPOSITION 2

Let $x > 0$ and consider a continuously differentiable solution $\Phi(x)$ of (12)-(15). Then, under the given conditions, we need to prove that (i) $A\Phi(x) + f(x) = 0$ and $\Phi(x) < \mathcal{M}\Phi(x)$ whenever $x \in (0, U)$, and (ii) $A\Phi(x) + f(x) > 0$ and $\Phi(x) = \mathcal{M}\Phi(x)$ whenever $x \in [U, \infty)$.

We first prove (i). For this, note that $\Phi(x)$ is derived from the equation $A\Phi(x) + f(x) = 0$ for $x \in (0, U)$. Thus, we only need to verify that $\Phi(x) < \mathcal{M}\Phi(x)$ for $x \in (0, U)$, which is also implied immediately by Condition (17).

Next for (ii), note that for $x \in [U, \infty)$, $\Phi(x)$ is derived such that $\Phi(x) = \mathcal{M}\Phi(x)$. Thus, we only need to validate that $A\Phi(x) + f(x) > 0$ for $x \in [U, \infty)$. We conclude the proof with this required step. Note that for $x \in [U, \infty)$, we have

$$\begin{aligned}
A\Phi(x) + f(x) &= \mu x \frac{\partial \Phi(x)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 \Phi(x)}{\partial x^2} - r\Phi(x) + f(x) \\
&= \mu x c\Upsilon - r\left(K + c\Upsilon x + R(U, \Upsilon)\right) + \beta x^2 \\
&= \beta x^2 + (\mu - r)c\Upsilon x - r[K + R(U, \Upsilon)].
\end{aligned}$$

Thus, $A\Phi(x) + f(x)$ is a convex and quadratic function of x; moreover, the value of this function at

$x = 0$ is negative. Therefore, $A\Phi(x) + f(x)$ should be clearly positive for all $x > x_0$, where

$$x_0 := \frac{-(\mu - r)c\Upsilon + \{(\mu - r)^2 c^2 \Upsilon^2 + 4\beta r[K + R(U, \Upsilon)]\}^{\frac{1}{2}}}{2\beta}$$

is the positive root of $A\Phi(x) + f(x) = 0$. Condition (16) yields that $x_0 < U$. Consequently, $x_0 < U \leq x$ and we have the desired result. □

D. NUMERICAL SIMULATIONS TO VALIDATE $-\tilde{\mu}_G > 0$.

Table 2: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and σ.

$\mu\downarrow$	$\tilde{\mu}\rightarrow$ $\sigma\downarrow$	0.02	0.04	0.06	0.08
-.01	0.01	0.030	0.050	0.070	0.090
	0.10	0.022	0.032	0.049	0.080
0	0.01	0.017	0.049	0.087	0.124
	0.10	0.014	0.027	0.046	0.073
.01	0.01	0.001	0.028	0.060	0.086
	0.10	0.004	0.010	0.011	0.011

Table 3: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and $\tilde{\sigma}$.

$\mu\downarrow$	$\tilde{\mu}\rightarrow$ $\tilde{\sigma}\downarrow$	0.02	0.04	0.06	0.08
-.01	0.10	0.027	0.046	0.070	0.099
	0.15	0.022	0.032	0.049	0.080
0	0.10	0.017	0.036	0.060	0.089
	0.15	0.014	0.027	0.046	0.073
.01	0.10	0.004	0.010	0.012	0.012
	0.15	0.004	0.010	0.011	0.011

Table 4: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and K.

$\mu\downarrow$	$\tilde{\mu}\rightarrow$ $K\downarrow$	0.02	0.04	0.06	0.08
-.01	10	0.020	0.027	0.036	0.069
	20	0.022	0.032	0.049	0.080
0	10	0.012	0.024	0.040	0.065
	20	0.014	0.027	0.046	0.073
.01	10	0.001	0.003	0.004	0.004
	20	0.004	0.010	0.011	0.011

Table 5: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and c.

$\mu \downarrow$	$\tilde{\mu} \rightarrow$ $c \downarrow$	0.02	0.04	0.06	0.08
-.01	1.0	0.024	0.038	0.060	0.092
	2.0	0.022	0.032	0.049	0.080
0	1.0	0.016	0.031	0.053	0.084
	2.0	0.014	0.027	0.046	0.073
.01	1.0	0.004	0.015	0.025	0.034
	2.0	0.004	0.010	0.011	0.011

Table 6: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and r.

$\mu \downarrow$	$\tilde{\mu} \rightarrow$ $r \downarrow$	0.02	0.04	0.06	0.08
-.01	0.05	0.022	0.032	0.051	0.083
	0.10	0.021	0.027	0.034	0.036
0	0.05	0.014	0.028	0.047	0.077
	0.10	0.013	0.025	0.042	0.064
.01	0.05	0.001	0.006	0.010	0.011
	0.10	0.004	0.018	0.034	0.056

Table 7: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and β.

$\mu \downarrow$	$\tilde{\mu} \rightarrow$ $\beta \downarrow$	0.02	0.04	0.06	0.08
-.01	1	0.022	0.032	0.049	0.080
	3	0.024	0.037	0.058	0.090
0	1	0.014	0.027	0.046	0.073
	3	0.016	0.030	0.042	0.053
.01	1	0.004	0.010	0.011	0.011
	3	0.003	0.016	0.029	0.042

Table 8: Value of $-\tilde{\mu}_g$ for different values of μ, $\tilde{\mu}$ and λ.

$\mu \downarrow$	$\tilde{\mu} \rightarrow$ $\lambda \downarrow$	0.02	0.04	0.06	0.08
-.01	0.5	0.022	0.032	0.049	0.080
	1.0	0.023	0.035	0.050	0.070
0	0.5	0.014	0.027	0.046	0.073
	1.0	0.014	0.027	0.043	0.064
.01	0.5	0.004	0.010	0.011	0.011
	1.0	0.005	0.013	0.015	0.015

References

Anthony, B. M., Chung, C. 2013. Data plan throttling: A simple consumer choice mechanism. *In Global Communications Conference (GLOBECOM), IEEE*, 3151–3156.

AT&T.com. 2018. AT&T Network Practices: Broadband Information. Available at: http://about.att.com/sites/broadband/network (Accessed March 21, 2019).

Bellman, R. 1957. *Dynamic Programming.* Princeton University Press, Princeton, NJ.

Bensoussan, A., B. Chevalier-Roignant. 2019. Sequential Capacity Expansion Options. *Operations Research* 67(1):33–57

Bensoussan, A., J.L. Lions. 1984. *Impulse Control and Quasi-Variational Inequalities*. Gauhiers-Villars, Paris.

Bensoussan, A., H. Long, S. Perera, S. Sethi. 2012. Impulse control with random reaction periods: A central bank intervention problem. *Operations Research Letters* 40(6):425–430.

Boyce, W. E., R. C. Di Prima. 1997. *Elementary Differential Equations*, John Wiley, New York.

Cadenillas, A., P. Lakner, M. Pinedo. 2010. Optimal Control of a Mean-Reverting Inventory. *Operations Research* 58(6):1697–1710.

Cadenillas, A., F. Zapatero. 1999. Optimal central bank intervention in the foreign exchange market. *Journal of Economic Theory* 87(1): 218–242.

Cadenillas, A., F. Zapatero. 2000. Classical and impulse stochastic control of the exchange rate using interest rates and reserves. *Mathematical Finance* 10(2):141–156.

Cho, S., L. Qiu, S. Bandyopadhyay. 2016. Should Online Content Providers Be Allowed To Subsidize Content?—An Economic Analysis. *Information Systems Research* 27(3):580–595.

Constantinides, G. M., S. F. Richard. 1978. Existence of optimal simple policies for discounted-cost inventory and cash management in continuous time. *Operations Research* 4:620–636.

Consumers Energy. 2018. Peak Power Savers® AC Peak Cycling program. Available at: https://peakpowersavers.com/ (Accessed March 21, 2019).

Dai, J. G., D. Yao. 2013a. Brownian Inventory Models with Convex Holding Cost, Part 1: Average-Optimal Controls. *Stochastic Systems* 3(2):442–499.

Dai, J. G., D. Yao. 2013b. Brownian Inventory Models with Convex Holding Cost, Part 2: Discount-Optimal Controls. *Stochastic Systems* 3(2):500–573.

De, P., L. Hao, Y. M. Li, Y. Tan. 2017. Quality of Service Based Pricing Schemes for Content Sharing in Peer-to-Peer Networks. *Production and Operations Management* 26(8):1427–1443.

Dede, G., G. Hatzithanasis, T. Kamalakis, C. Michalakelis. 2018. Brokering Cloud Computing: Pricing Models and Simulation Approaches. *in Perspectives, Trends, and Applications in Corporate Finance and Accounting, IGI Global*. 230–251.

Deng, T., J. Yao, H. Guan. 2018. Maximizing Profit of Cloud Service Brokerage with Economic Demand Response. *In IEEE INFOCOM 2018-IEEE Conference on Computer Communications, IEEE*. 1907–1915.

Deshpande, M., A. Amit, M. Chang, N. Venkatasubramanian, S. Mehrotra. 2007. Flashback: A peer-to-peer web server for flash crowds. *In Distributed Computing Systems, 2007. ICDCS 2007. 27th International Conference, IEEE*. 1–15.

Dolgui, A., Ivanov, D., Sethi, S. P., B. Sokolov. 2019. Scheduling in production, supply chain and Industry 4.0 systems by optimal control: fundamentals, state-of-the-art and applications. *International Journal of Production Research*, 57(2), 411–432.

Dyaberi, J. M., et al. 2012. Managing cellular congestion using incentives. *IEEE Communications Magazine* 50(11):100–107.

Easley, R. F., H. Guo, J. Krämer. 2018. Research Commentary – From Net Neutrality to Data Neutrality: A Techno-Economic Framework and Research Agenda. *Information Systems Research* 29(2):253–272.

Feng, H., K. Muthuraman. 2010. A Computational Method for Stochastic Impulse Control Problems. *Mathematics of Operations Research* 35(4):830–850.

Feuillet, M., M. Jonckheere, M., B. Prabhu. 2014. Bandwidth sharing networks with multiscale traffic. *Stochastic Systems* 4(2):449–478.

Fukuda, C., H. Pita, H. Grebel, R. Rojas-Cessa, A. Mohamed. 2018. A digital approach to energy networks: Allocation and distribution of energy requests. *In Fog and Mobile Edge Computing (FMEC), 2018 Third International Conference, IEEE*. 205–210.

Guo, H., H. K. Cheng, S. Bandyopadhyay. 2013. Broadband network management and the net neutrality debate. *Production and Operations Management* 22(5):1287–1298.

Guo, H., R. F. Easley. 2016. Network neutrality versus paid prioritization: Analyzing the impact on content innovation. *Production and Operations Management* 25(7):1261–1273.

Harrison, J. M., T. M. Sellke, A. J. Taylor. 1993. Impulse Control of Brownian Motion. *Mathematics of Operations Research* 8(3):454–466.

He, S., D. Yao, H. Zhang. 2017. Optimal ordering policy for inventory systems with quantity-dependent setup costs. *Mathematics of Operations Research* 42(4):979–1006.

Hecht, J. 2016. The Bandwidth Bottleneck that is Throttling the Internet. Available at: `https://www.scientificamerican.com/article/the-bandwidth-bottleneck-that-is-throttling-the-internet` (Accessed March 21, 2019).

Hildenbrand, J. 2017. *When do carriers start throttling you and what can you do about it?*. Available at: `https://www.androidcentral.com/when-do-carriers-start-throttling-you-and-what-you-can-do-about-it` (Accessed March 21, 2019).

Ivanov, D., Sethi, S., Dolgui, A., B. Sokolov. 2018. A survey on control theory applications to operational systems, supply chain management, and Industry 4.0. *Annual Reviews in Control*, 46, 134–147.

Johar, M., S. Menon, V. Mookerjee. 2011. Analyzing sharing in peer-to-peer networks under various congestion measures. *Information Systems Research* 22(2):325–345.

Jordan, S. 2017. Evaluating Zero-Rating and Associated Throttling Practices Under the Open Internet Order. *Journal of Information Policy* 7:450–507.

Kim, D., S. W. Lee, Y. B. Ko, J. H. Kim. 2015. Cache capacity-aware content centric networking under flash crowds. *Journal of Network and Computer Applications* 50:101–113.

Korn, R. 1997. Optimal impulse control when control actions have random consequences. *Mathematics of Operations Research* 22(3):639–667.

Kourandi, F., J. Krämer, T. Valletti. 2015. Net neutrality, exclusivity contracts, and internet fragmentation. *Information Systems Research* 26(2):320–338.

Krämer, J., L. Wiewiorra. 2012. Network neutrality and congestion sensitive content providers: Implications for content variety, broadband investment, and regulation. *Information Systems Research* 23(4):1303–1321.

Li, F., A. Razaghpanah, A. M. Kakhki, A. A. Niaki, D. Choffnes, P. Gill, A. Mislove. 2017. lib erate,(n): A library for exposing (traffic-classification) rules and avoiding them efficiently. *In Proceedings of the 2017 Internet Measurement Conference, ACM*,128–141.

Lumley, R. R., M. Zervos. 2001. A Model for Investments in the Natural Resource Industry with Switching Costs. *Mathematics of Operations Research* 26(4):637–653.

Muthuraman, K., S. Seshadri, Q. Wu. 2015. Inventory management with stochastic lead times. *Mathematics of Operations Research* 40(2):302–327.

Mitchell, D., H. Feng, K. Muthuraman. 2014. Impulse Control of Interest Rates. *Operations Research* 62(3):602–615.

O'Donoghue, R. 2014. You've been throttled, but don't stop browsing!. 'mobiForge' available at: `https://mobiforge.com/news-comment/youve-been-throttled-dont-stop-browsing` (Accessed March 21, 2019).

Øksendal, B. 1999. Stochastic control problems where small intervention costs have big effects. *Applied Mathematics and Optimization* 40:355–375.

Øksendal, B. 2003. *Stochastic Differential Equations. An Introduction with Applications,* 6th Ed., Springer-Verlag, Berlin, Germany.

Øksendal, B., A. Sulem. 2007. *Applied Stochastic Control of Jump Diffusions.* Springer-Verlag, Berlin, Germany.

Ohnishi, M., M. Tsujimura. 2006. An impulse control of a geometric Brownian motion with quadratic costs. *European Journal of Operational Research* 168(2):311–321.

Ormeci, M., J. G. Dai, J. Vande Vate. 2008. Impulse Control of Brownian Motion: The Constrained Average Cost Case. *Operations Research* 56(3):618–629.

Perera, S., W. Buckley, H. Long. 2018. Market-reaction-adjusted optimal central bank intervention policy in a forex market with jumps. *Ann. Oper. Res.* 262(1):213–238.

Perera, S., W. Buckley. 2017. On the existence and uniqueness of the optimal central bank intervention policy in a forex market with jumps. *J. Oper. Res. Soc.* 68(8):877–885.

Presman, E., S. P. Sethi. 2006. Inventory models with continuous and Poisson demands and discounted and average costs. *Production Operations Management* 15(2):279–293.

Spectrum.net. 2018. Spectrum.net Network Congestion Management. Available at: `https://www.spectrum.net/support/general/network-management/` (Accessed March 21, 2019).

Stavrou, A., D. Rubenstein, S. Sahu. 2004. A lightweight, robust P2P system to handle flash crowds. *IEEE Journal on Selected Areas in Communications* 22(1):6–17.

Stevens, M. 2018. *Verizon Throttled California Firefighters' Internet Speeds Amid Blaze (They Were Out of Data).* (Accessed March 21, 2019) Available at: `https://www.nytimes.com/2018/08/22/us/verizon-throttling-california-fire-net-neutrality.html`.

Sulem, A. 1986. Explicit solution of a two-dimensional deterministic inventory problem. *Mathematics of Operations Research* 11(1):134–146.

T-Mobile Binge On. 2018. *Binge On — Stream Unlimited Video without Using Your Data*. Available at: `https://www.t-mobile.com/offers/binge-on-streaming-video` (Accessed March 21, 2019).

Tada, H., M. Murata, M. Aida. 2017. Web Service Flash Crowd Mitigation Using Feedback to Users. *In Advanced Information Networking and Applications (AINA), 2017 IEEE 31st International Conference, IEEE*. 213–219

The Guardian. 2017. Game of Thrones fans furious after Foxtel streaming service crashes. Available at: `https://www.theguardian.com/australia-news/2017/jul/18/games-of-thrones-fans-furious-after-foxtel-streaming-service-crashes` (Accessed March 21, 2019).

Verizon.com. 2018. Verizon.com Network Management. Available at: `https://www.verizon.com/about/our-company/network-management` (Accessed March 21, 2019).

VOA.com. 2018. After Throttling Firefighters' Internet, Verizon Makes Changes. Available at: `https://www.voanews.com/a/after-throttling-firefighters-internet-verizon-makes-changes/4543478.html` (Accessed March 21, 2019).

Wu, J., X. Chao. 2014. Optimal Control of a Brownian Production/Inventory System with Average Cost Criterion. *Mathematics of Operations Research* 39(1):163–189.

Yamazaki, K. 2017. Inventory control for spectrally positive Lévy demand processes. *Mathematics of Operations Research* 42(1):212–237.

Zheng, L., C. Joe-Wong, M. Andrews, M. Chiang. 2018. Optimizing data plans: Usage dynamics in mobile data networks. *In IEEE Intl. Conf. on Comp. Comm. (INFOCOM).*

DECISION SCIENCES INSTITUTE

Behavioral Finance and the Architecture of the Asset Management Industry

Michel Verlaine
Michel.verlaine@icn-artem.com

ABSTRACT

The standard portfolio approach assumes that investors maximize Expected Utility functions and that the Markowitz Mean-Variance Standard Portfolio Optimization approach can be applied. Behavioral Research, however, indicates that investors' behavior with respect to risk or uncertainty is not consistent with EU. Notably, decision makers transform probabilities. Most of the academic literature integrating those aspects, however, focuses on so-called financial market anomalies but does not focus on the structural impact those behavioral aspects might have, notably in terms of industrial organization of the asset management industry. This paper analyzes the industrial organization of the asset management industry and argues that the architecture of the asset management industry is inconsistent with the standard EU framework. The origin of the industrial organization of the asset management industry stems from the convexity of the Flow-performance relationship, which is due to Prospect Theory effects. The convexity of the Flow-performance relationship explains the inflation of Mutual Funds and Fund families, a phenomenon inconsistent with the Mutual Fund Separation Theorem..

KEYWORDS:: Portfolio Choice, Behavior under Uncertainty, Behavioral Industrial Organization.

JEL classification: G11, D82, N2.

INTRODUCTION

Most of the academic literature in Behavioral Finance focuses on so-called financial market anomalies that cannot be explained by the standard rationality assumptions in Economics. An extensive literature has been developed over the last two decades. (Benartzi & Thaler, 2003) provide a somewhat older survey and a more recent account of the literature can be found in (Benartzi, 2018). As the reader of those surveys will readily notice, those surveys exclusively analyze the impact of deviations from rationality on asset markets. In a sense, this literature analyzes how behavioral aspects generate financial market anomalies. More recently, a new literature called Behavioral Corporate Finance has been developed. (Malmendier, 2018) surveys this literature and points out that until recently research on the impact of behavioral aspects on Corporate Finance decisions has been very low. Yet, recently, research on Behavioral Corporate Finance has become a strong sub-field.

Strangely, the academic research that focuses on the industrial organization of the asset management industry is weak or nearly non-existent. This is a pity as the understanding of the value chain of the financial industry is important to understand its Business Models, notably how they evolve through time and adapt to technological factors (FinTech) and behavioral shifts of stakeholders. With the growth of the AM Industry, the understanding of those value chains and eventual relocation strategies of sub-units of Financial Institutions has become even more relevant and eventually important for political decision makers. What we call the Architecture of Finance is the way the finance industry is structured with its different financial institutions. We thus distinguish different institutions such as Banks, Insurance companies, different types of funds (Mutual Funds, Alternative Investment funds) as well as the emergence of new intermediary structures such as Shadow Banks. The evolutionary process of those institutions can be analyzed through technological and behavioral aspects. A similar approach has been developed for financial markets. (Lo, 2017) has put forward the Adaptive Markets Hypothesis in comparison to the Efficient Markets Hypothesis. The idea is to analyze Financial Markets movements through evolutionary and adaptive mechanisms of “institutional traders” rather than rational and individual investors.

The underlying idea is that in order to understand the existence and the transformations of this Architecture we need to add frictions such as asymmetry of information, cognitive frictions and regulatory frictions. The latter, however, can be seen as derived from the two first types of frictions. Indeed, financial regulation exists in order to protect economic agents against others exploiting “irrationalities” and situations of asymmetry of information.

We are, here, interested in the “Architecture of the Asset Management Industry” an account of which in relation to behavioral aspects is provided in (Verlaine, 2016). Of central importance to understand why behavioral aspects matter to explain the Architecture of Finance is the Mutual Fund Separation Theorem (Tobin, 1958) and its questioning, namely the Asset Allocation Puzzle. Indeed, the Mutual Fund Separation Theorem states that “rational” investors with similar information sets should invest in a risky portfolio and mix this portfolio with the risk free asset. It means that whatever the risk aversions of investors, they should hold similar risky portfolios. This is a very strong prediction and it can be viewed as some kind of Modigliani-Miller Theorem for the Asset Management Industry, indicating what we should observe in a frictionless equilibrium. Basically, if we were to analyze the asset allocations of investors we should observe relatively balanced risky portfolios or funds combined with investments in non-risky investments, let’s say government bonds. This is not what we observe. Most small investors hold non-diversified portfolios of a few stocks. But recommendations of asset management firms as well are not in line with the Mutual Fund Separation Theorem (MFST). (Canner et al, 1997) were the first to document the inconsistency between asset allocation recommendations of investment advisors for different risk profiles of investors. This phenomenon is called the Asset Allocation Puzzle. Frictions such as background risk and transaction costs might explain some deviations from the MFST but are insufficient to explain the empirical differences. Strangely, not much has been published on that phenomenon.

The Mutual Fund Separation Theorem is not only important as a guide, and some professionals might say a bad guide, for portfolio construction. It is also a major building block of general equilibrium financial market models such as the Capital Asset Pricing Models and it hence also has an impact on valuation models in Corporate Finance. As far as equilibrium asset pricing is concerned, there has been some elaboration by adding further frictions to the CAPM model (Fama & French, 1993). The fundamental portfolio choice approach and its implications for the finance industry, however, have not been questioned. Recently, (Markowitz, 2014) analyzed the conditions for the Mean-Variance approach to portfolio choice. He points out that most Expected utility functions support Mean-Variance optimization but that models with probability weighting do not support Mean-Variance optimization. Empirical evidence on individual behavior indicates that most individuals act according to Prospect theory (Bruhin et al, 2010); Wakker, 2010). This implies that they are loss averse and transform observed probability distributions. Some tentative approaches to portfolio optimization have been developed notably by (Shefrin & Statman, 2000), but they are not very operational. Recent publications in operational research indicate an interest in decision-theoretic approaches, ambiguity and risk management (Borgonovo & Marinacci, 2015;

Borgonovo et al, 2018; Cerreia-Voglio et al, 2013). A first attempt to apply those concepts to the Fund Industry can be found in (Verlaine, 2018).

Focusing more on the industrial organization of the Asset Management Industry, the first attempt to analyze fund vehicles rather than individual investors seems (Berk & Green, 2004). Their approach, however, is based on standard neoclassical assumptions of rationality. The core element is the flow-performance relationship which indicates the sensitivity of in- and out-flows of the Fund as a function of abnormal performance. This abnormal performance is typically measured with the Carhart 4-factor model and for sophisticated investors the relationship should be linear. The neoclassical assumption is that costs in percentage should decrease with AUM, but alpha as well and faster. There hence exists an industry equilibrium where funds attract in-flows as long as alpha is higher than the fees, until the fees are equal to the alpha. The cross-section data of investment funds should thus indicate a positive relationship between alpha and fees, where in the long run equilibrium abnormal performance after fees is 0. The US data, however, indicates that the relationship between fees and alpha is negative. Funds with lower performance have higher fees (Gil-Basu & Ruiz-Verdu, 2009).

The flow-performance relationship is of central importance to understand that phenomenon. Indeed, the flow-performance relationship is not linear as presumed by neo-classical economic theory but non-linear and more or less convex. This means that inflows are much more sensitive to good performance than outflows to bad performance. This phenomenon can be linked to behavioral aspects such as the asymmetric behavior with respect to gains and losses implicit in Prospect Theory. This non-linearity is of central importance as it provides an explanation of the multitude of funds and families of funds, an empirical fact inconsistent with the MFST. Asset Management firms facing a convex flow-performance relationship can launch a family of funds that ex-ante have similar return profiles. As the realized returns ex-post will differ, the funds can then promote the best performing fund which then attracts higher inflows.

The paper is organized as follows. Section I presents the Investor Behavior and the portfolio delegation issues that go with it. Section II then analyses the relationship between fees and fund performance. The observed fee anomaly then leads us to analyse the impact of the Board composition on the fee setting process in Section III. Section IV focuses on the behavioral aspects behind the convexity of the flow-performance relationship. Section V analyses whether the distribution channels could have an impact on the fee-performance anomaly. Section VI and VII analyze the impact on the outsourcing and internalization. Finally, Section VIII explains how the behavioral features that lead to convex flow-performance relationships can be exploited within fund families.

INVESTOR BEHAVIOUR AND PORTFOLIO DELEGATION ISSUES

Standard Markowitz portfolio theory presumes that investors are EU maximizers with a concave Bernouilli utility function. The decision theoretic and behavioral economics literature, however, highlights that most individuals' behavior is not consistent with the EU hypothesis. Recent experimental evidence is provided in (Bruhin et al, 2010). (Kahneman & Tversky, 1979; 1992), document that individuals evaluate gains and losses with respect to a subjective reference level. Their so-called value function is concave on gains and convex on losses. Moreover, investors are strongly averse to losses. This has led to the following well-known specification:

$$v(w-\overline{w}) = \begin{cases} (w-\overline{w})^{\alpha} & if \quad w \geq \overline{w} \\ -\lambda(w-\overline{w})^{\beta} & if \quad w < \overline{w} \end{cases} \tag{1}$$

Where w and $\overline{w}$ respectively denote the realized wealth and the subjective reference wealth level. λ is the loss aversion parameter and is typically calibrated to 2.25. α and β are generally calibrated to 0.88.

Investors also transform observed probabilities and the way these probabilities are distorted is well-documented (Tversky & Wakker, 1995; Wakker, 2010; Prelec, 1998). Investors typically apply inverse S-shaped functions, called probability weighting functions, to the cumulative probability distribution. Recent research indicates that such behavior could stem from a kind of ambiguity aversion. Frameworks to treat ambiguity have been developed by (Gilboa & Schmeidler, 1989; Schmeidler, 1989). Those authors suggest that DMs maximize EU under the worst case scenario given sets of probability distributions. (Schmeidler, 1989) shows that relaxations of the independence axiom lead to a preference representation with respect to a non-additive set measure. This model is known as the Choquet Expected Utility (CEU) model. (Ghirardato et al, 2004) suggest to model behavior in the face of ambiguity with a weighted $(\alpha, 1-\alpha)$ combination of the worst and the best scenarios. This has led to the $\alpha - MEU$ approach where α is interpreted as a parameter of ambiguity aversion. Recently, (Amarante, 2009) shows that $\alpha - MEU$ can be mapped on CEU representations. Finally, a last dimension that is often forgotten in the literature, is that DMs also typically consider prospects in different mental accounts. This opens the door for multiple investment styles and funds. This provides an explanation of the asset Allocation Puzzle (Canner et al, 1997). The academic literature, however, has faced difficulties to integrate those behavioral aspects into an operational portfolio optimization framework.

(Shefrin & Statman, 2000) develop a portfolio optimization approach that is consistent with empirically documented behavior. They argue that there are three dimensions that matter for the DM: *potential*, *aspiration* and *security*. *Aspiration* indicates the level of wealth expected or desired by the DM. *Security* refers to the desire to limit losses with respect to the aspiration level. Finally, *potential* indicates that DMs want to invest a fraction in assets that have a very small probability of very high payoffs such as lottery tickets. This is in line with the well-documented fact that DMs buy insurance as well as lottery tickets. The three dimensions are linked to different behavioral aspects of prospect theory. Besides behavioral aspects, institutional features are also important to take into account and are actually conditioned by behavioral features. (Verlaine, 2016) provides a detailed analysis of this topic. (Polkovnichenko & Zhao, 2013) show that the risk neutral probability distribution implicit in option prices is consistent with probability weighting behavior. They estimate parameters for the (Prelec, 1998) probability weighting function. (Polkovnichenko, et al, 2018) show that those probability weighting features can explain the demand for active funds. This thus rationalizes the delegation of asset management to active asset managers and opens the door for information economics and principal-agent problems in asset management.

As argued in (Verlaine, 2018), the optimal contracts with the asset managers bear a close relationship with the risk management of the fund and behavioral aspects thus also influence the optimal contract. (Stracca, 2005) surveys the delegation issues in portfolio management. The delegation problem is typically modelled through a Principal-Agent relationship. The investor is the Principal and the Agent is the Asset Manager. The portfolio delegation literature focuses on contractual issues. In a first step, the investors face a set of Managers and have to select the good ones that have the skills to generate outperformance. In a second step, the contract and the compensation rules have to incentivize the Manager to spend maximal effort to generate outperformance. The investors are thus supposed to structure the contract in such a way that 3 conditions be respected. First, the managers with superior skills are selected. Second, the contract incentivizes the Asset Manager to generate risk-adjusted outperformance once the contract is signed. Finally, the risk-sharing between the investor and the Asset Manager should also be optimal.

Formally, the maximization problem takes the following form.
In the first order the Principal maximizes his expected utility given by :

$$\begin{aligned} &Max_f\, EU((1+R)-f(R)) \\ &with\, R = w(S(e))R_p \end{aligned} \tag{2}$$

He thus maximizes expected utility of gross return minus the fees. But in order to hire a manager he also needs to be sure that the expected utility of the return is higher than what he gets if he simply invests in a passively managed fund. Hence, the following participation constraint has to hold:

$$EU(1+R-f(R)) \geq EU(1+R_u) \tag{3}$$

So basically, the EU of the gross net of fees return has to be higher than what the investor gets when he invests in an Exchange Traded Funds, for instance.
The manager also maximizes utility and faces a participation contraint.

$$e, w(S) = \arg\max\left[EU(f(R)) - h(e)\right] \tag{4}$$

So the manager chooses an optimal effort level and portfolio combination so that he maximizes EU of fees perceived minus his disutility of effort *h(e)*. However, the following participation constraint

$$EU(f(R)-h(e)) \geq U^* \tag{5}$$

indicates that he is only working for the investors if the expected utility of the fee minus disutility of effort is higher than the utility he gets in another job.
Different fee functions have been suggested, but they all face the drawback that, due to the fact that there is no direct relationship between effort spent by the manager and portfolio return, it is extremely difficult to make sure the manager spends effort and that the returns are due to luck or hidden risks. Basically, there are two types of contracts, linear and non-linear, which we analyse below.

The linear compensation contract takes the following form:

$$\begin{aligned} &f(R) = C + B(R-b) \\ &with\ C, B > 0 \end{aligned} \tag{6}$$

b is a subjective benchmark that is known in advance and it can be either stochastic or fixed. A behavioral approach implies the evaluation of this subjective benchmark for a prospect theory type decision maker and then the choice of an investment fund benchmark that has the same expected return. Typically Mutual Funds are benchmarked with a stochastic market based benchmark, generally given by an index that reflects the strategy of the fund. Hedge Funds, however, are mostly benchmarked with an absolute target which is given by the Hurdle rate. Hedge Funds often also have so-called High-water mark clauses which specify that past losses with respect to the benchmark have to be made up before the variable fees are paid. As shown by (Brown & Goetzmann, 2003) this creates path dependent option like features which are difficult to analyse within this framework. Within the linear contract framework, a higher B induces the manager to spend more effort but also to take more risk which will increase the required flat payment C.

In the standard agency setting, there is a trade-off between risk sharing and effort inducement. However, in a portfolio delegation framework this kind of contract is insufficient to induce the manager to work harder as the manager has complete control over his response to signals (his information). This result is known as the 'irrelevance result'. This result holds irrespective of whether the contract contains a benchmark. A benchmark, unless it is the optimal conditional portfolio, is a distorting factor.

The non-linear compensation contract has also been suggested as an optimal compensation rule:

$$f(R) = C + B(R-b)^2 \quad with \;\; C, B > 0 \tag{7}$$

Here, compensation depends in a quadratic way on the relative performance compared to a benchmark. It can induce the manager to undertake more effort. However, if the manager is allowed to respond to the signal in a nonlinear way, the 'irrelevance result' still holds. Hence, restrictions have to be imposed on the manager's trading set. This is an interesting result in terms of industry regulation.
Finally, linear vs. non-linear asymmetric contracts

$$f(R) = \begin{cases} \mathrm{C + B(R - b)} \;\; \text{if} \;\; \mathrm{R > b} \\ \quad 0 \;\; \text{if} \qquad\quad \mathrm{R < b} \end{cases} \tag{8}$$

Symmetric contracts penalize negative outcomes in the same way as they reward positive outcomes, while for convex contracts the marginal reward is higher the better the performance compared to a benchmark. The "irrelevance result" and the fact that restrictions on the trading set have to be imposed rationalizes the recourse to risk management and risk limits. The Board's responsibility in terms of risk management is thus to select a benchmark that is in line with the risk profile of the fund and determine the risk limit or limits of the fund or positions. The foregoing developments imply that the Board has to develop a top-down approach to the risk management and measurement.

The function of the benchmark is also to distinguish performance related to systematic risk factors from performance due to superior selection skills. From an operational viewpoint, the role of the benchmark is also to control the deviation of the portfolio with respect to the benchmark. As pointed out in (CMRA, 2008), the main risk is that the performance will fall short of the benchmark. Finally, benchmarks are also often used to compare fund performance to the benchmark. Correct benchmarking is an issue and (Sensoy, 2009) suggests that approximately 35% of benchmarks are not in line with the risk profile of the fund.

Deviations from the benchmark are typically controlled through the trading error variance (TEV). This is in line with the quadratic compensation rule above. If we apply the expectation operator to the quadratic compensation rule we get:

$$\begin{aligned} E[f(R)] &= C + BE\left[(R-b)^2\right] \\ &= C + B \times TEV \end{aligned} \tag{9}$$

And the compensation thus depends on the TEV. Even though this approach is quite used in practice, the TEV evaluates positive and negative deviations from the benchmark in the same way.
The next sections focus on the different features of the asset management industry highlighting the behavioral and contractual features that seem inconsistent with the standard neoclassical EU assumptions.

ANALYSIS OF THE RELATIONSHIP BETWEEN FEES AND PERFORMANCE

In principle, there should be a positive relationship between before-fee performance and fees in the fund industry. According to some recent research, however, the relationship is negative (Gil-Bazu, & Ruiz-Verdu, 2009). It might be worthwhile highlighting that Fund management fees are typically computed as a fixed percentage of the value of AUM. These fees, together with other operating costs, such as custodian, administration, accounting, registration and transfer agent fees, comprise the fund's expenses. Expenses are usually expressed as a percentage of AUM known as the expense ratio. Fees paid to brokers in the course of the trading activity are detracted from the fund's assets but are not included in the fund's assets. Distribution fees also matter but, in general, they are included in the operating expenses. Funds often charge "loads", which are one time fees that are used to pay distribution. The "loads" are paid at the time of purchasing ("front-end loads") or redeeming ("back-end loads" or "deferred sales charge") fund shares and are computed as the fraction of the amount invested. Since 1980, funds may charge so-called "12b-1 fees", which are included in the expense ratio. Since 1990, many funds offer multiple share classes with different combinations of "loads" and "12b-1 fees" (f.i. high front-end loads and low annual 12b-1 fees or vice-versa).

In order to analyse the relationship between performance and fees, (Gil-Bazu, & Ruiz-Verdu, 2009) focus on a specific Mutual Fund sample. First, all funds that cannot be described as diversified domestic equity mutual funds are excluded from the sample. Second, observations with extreme values for returns and expenses are also removed from the sample (outliers). Finally, to ensure that results are not driven by differences between index and actively managed funds or between institutional and retail funds, passively managed (index) and institutional funds are excluded from the sample.

As the aim of the board members should be to care about shareholders' interests, Fund Governance issues matter for the fee setting process. The measure of the board quality is the sum of four equally weighted components that measure:

- The degree to which the board has taken action "in cases where the fund clearly has not served investors well".
- The significance of independent director's investments in the fund.
- Whether the board is "overseeing so many funds that it may compromise the ability to diligently protect interests of shareholders".
- The proportion of independent directors.

The authors are interested in comparing before-fee risk-adjusted performance to fees. In order to determine that performance, a model is needed. Different models could be used, but market standard has become the (Carhart, 1997) 4-factor or (Fama & French, 1993) 3-factor model. Carhart's model is used extensively as it encompasses the Fama-French model. The specification is the following:

$$r_{it} = \alpha_i + \beta_{rm,i} rm_t + \beta_{hml,i} hml_t + \beta_{smb,i} smb_t + \beta_{pr1y,i} pr1y_t + \varepsilon_{it} \tag{10}$$

Where smb and hml denote returns on portfolios that proxy for common risk factors associated with size and book-to-market, respectively. The term pr1yt is the return difference between stocks with high and low returns in the previous year. This term accounts for passive momentum strategies. r_{it} is the fund's before-expense return in month t in excess of the 30-day risk-free interest rate. The alpha is supposed to capture the fund's before-fee risk adjusted performance.

Suppose that this model were a good approximation for the data generating mechanism. What should be the fees of the funds in equilibrium? Well, in the absence of frictions, all funds should earn zero expected after-fee risk-adjusted returns. Otherwise, there would be excess demand for funds with positive after-fee risk-adjusted return (Berk & Green, 2004). Thus, if investors knew the fund alphas, in equilibrium we should observe $\alpha_i = f_i \ \forall_i$ and there should thus exist an approximately linear relationship of slope 1 between fees and alphas. If investors do not know the alphas then the equilibrium relationship takes the form of an expectation: $E(\alpha_i) = f_i \ \forall_i$.

What are the equilibrium adjustment mechanisms in place? Equilibrium in the MF market can be achieved through fee adjustments if funds with higher expected before-fee risk-adjusted returns increase their fees or underperforming funds lower theirs. (Berk & Green, 2004) show that in the market for MF's, market clearing can also be via quantity adjustments. If there are decreasing returns to scale in fund management, then flows of money into funds that are expected to perform better will reduce these funds' expected performance until after-fee expected returns are equalized across all funds in equilibrium. The equilibrium will require that expected after-fee risk-adjusted returns be zero for all funds and that there exists a linear relationship of slope 1 between fees and before-fee performance. In order to test this assumption, fund fees and before-fee risk-adjusted performance is regressed by pooled OLS regression:

$$\hat{\alpha}_{it} = \delta_{0t} + \delta_1 f_{it} + \varepsilon_{it},\ i = 1, \dots, N,\ t = 1, \dots, T \tag{11}$$

Where f_{it} is the fund's expense ratio expressed as a fraction of fund's assets. Alpha is the risk-adjusted before-fee performance measured according to Carhart's 4-factor model.

The results of the estimation indicate a slope coefficient that is between -1 and -0.5 and the null hypothesis of a unit slope is strongly rejected. This result is in stark contrast with the equilibrium predicted in a frictionless economy. In a market with frictions, however, we might observe deviations from unit slope and it is difficult to forecast whether the slope should be higher or lower. Managers should generate positive alpha and could over- or under-compensate for their respective performance. It might also be plausible that there exists no relationship between fees and alphas, which would lead to the slope being 0. However, the slope being negative, it suggests that worse-performing funds charge higher fees. It might be that the results are partly due to estimation errors. There might be estimation errors in the Carhart regression and these estimation errors might impact the estimates of the (alpha, fee) regression. (Gil-Bazo & Ruiz-Verdu, 2009) show by non-parametric techniques that the relationship is negative and robust to extreme cases. Moreover, it holds over different sub-periods.

The relationship warrants some explanations. There are 3 potential types of explanations:

- Cost-based explanations
- Strategic explanations
- Fund Governance explanations

Cost-based explanations:

Fees might simply reflect operating costs of the fund and low-operating cost funds might have high before-fee risk-adjusted performance. This would thus explain the negative relationship. Two rationales are suggested. First, economies of scale might lead to lower operating costs for larger funds and performance persistence might lead to better performing funds becoming larger. Second, better management skills might lead to better investment decisions and efficient management of fund operations.

Strategic explanations:

The relationship might be the result of some strategic fee setting.

First, (Christofferson & Musto, 2002) argue that fund's investors differ in their performance sensitivity. Also, funds with bad performance history face out-flows and mainly stay with performance insensitive investors. In order to compensate for losses, they might increase fees which would lead to a negative relationship between risk-adjusted performance and fees.

Second, (Gil-Bazo & Ruiz-Verdu, 2008) argue that competition in a market with asymmetric information might lead to an equilibrium in which funds with higher return expectations reduce their fees up to the point where they price lower return expectation funds out of the performance sensitive share of the market. Moreover, as the fee is a fraction of AUM and as the expected value increase is higher for "good" funds, they can expect a value increase in fees. "Bad" funds will then raise fees in order to extract rents from performance insensitive investors.

Third, low-performing funds might face higher marketing and distribution costs. Insensitive investors are less sensitive to marketing material and the costs and time spent to convince them to invest in the fund might be higher. Moreover, distributors might ask for a premium for the extra-effort and eventual reputation risk they face.

Fund Governance:

The strategic explanation implicitly assumed that fees are set by the management company, but fund fees are negotiated with the board of directors. The latter have the fiduciary duty to make sure that the fees are in line with the added value they provide investors. There are some conflicts of interest, however, as the board members are selected by the management company. The directors interests might thus be more aligned with that of the management company than with investors. It can be expected that boards of better governed funds might ensure that fees are in line with value added. It means that the relationship between performance and fees should be positive or flatter for good-governed funds.

To test the different suggested explanations the fees are regressed against fund characteristics, flow-to-performance sensitivity and performance. Fund fees are assumed to linearly depend on a vector x_{t-1} of variables supposed to determine fund operating costs, the performance sensitivity of fund flows, S_t, and the fund's expected before-fee performance in period t, α_t.

$$f_{it} = \gamma' x_{it-1} + \lambda_S S_{it} + \lambda_S \alpha_{it} + v_{it} \tag{12}$$

With $E(v_{it}) = 0 \;\; and \;\; \sigma_v^2 = c$

Most of the variables are only available on a yearly basis and *t* thus refers to years. The following variables are supposed to impact the operating costs of the fund:

- Size: defined as the log of the year-end total NAV
- Age: computed as the log of the fund's years since inception
- Volatility: The standard deviation of log returns
- Reported annual turnover
- Investment objectives: represented through dummies

In order to test whether the negative relationship between performance and fees stems from the omission of the variables determining operational costs, the above- mentioned relation is first estimated without the performance sensitivity. The results indicate that the coefficient on alpha does not change sign and is still negative (-14.65) and thus cost-based arguments cannot explain why fees and performance are negatively related. The relationship is then re-estimated with the performance-sensitivity variable. The latter is highly significant and this seems to confirm that funds strategically exploit a low elasticity of demand with respect to performance to increase their fees. However, the coefficient on alpha is still negative and this is thus not sufficient to explain why underperforming funds have higher fees. Does fund governance have an influence on the fee structure? In order to test this assumption, the regression is implemented for different governance-rated sub-groups. The results indicate that better-governed funds resist the attempts to exploit performance-insensitivities of investors, as the relationship between performance and fees is flatter. This leads us to the question of the impact of the board structure on fees.

THE IMPACT OF THE BOARD STRUCTURE ON FEES

It is interesting to focus more in detail on the fee-setting process. Note that the fees have to be approved by the Board and the latter might thus have an impact on the % of fees charged to service providers. (Tufano & Sevick, 1997) analyse the relationship between Board composition and fees in more detail. As in any business, Boards are legally charged with the protection of shareholders' interests. In the fund business, this implies that they have to negotiate and approve the contracts with the management company, the distributor and other eventual service providers. Note, however, that in the fund industry there is a major specificity that is non-existent in other businesses. Indeed, the shareholder structure can rapidly change with different investors investing and redeeming their shares on a daily basis. Moreover, many funds are launched by fund sponsors, firms that manage and distribute a whole set of funds called fund families. Those sponsors typically choose the dependent as well as the independent board members (directors) at the inception of the fund. Those board members often sit in different funds of the fund family. Such a constellation might lead to conflicts of interest as the board members might not be representative of the shareholders when the shareholder structure of the fund changes.

A simple example will help illustrate the problem. Assume that a sponsor launches a fund by putting 10 Mio $ in the fund in order to kick-start it. At the same time the sponsor selects the board members. At inception, as the sponsor is also the only investor (shareholder), the interests of the sponsor, the Board and the shareholders are aligned. Now, assume that new investors step in and that the fund grows to 100 Mio $ AUM. Those investors are not affiliated with the sponsor but the board composition is still the same as at inception. Potential conflicts of interest might arise as for most of the shareholders, lower fees are in their interest, whereas for the fund sponsor it might be interesting to increase fees in order to increase income. As board members might be dependent on the compensation provided by the sponsor, notably on other funds distributed by the sponsor, they might align their interests with the ones of the sponsor.

In order to evaluate this potential governance issue, (Tufano & Sevick, 1997) analyse a sample of US open-end mutual funds which account for 69% of all US open-end mutual fund assets. The following characteristics of mutual fund boards are analysed:

- Board size
- The fraction of independent directors in the board
- The number of boards on which a board member represents a particular sponsor
- The compensation of the board members

In the US, the regulator requires that a certain percentage (around 40%) of fund board members are not affiliated with the sponsor, underwriter or broker. Also, amongst other responsibilities the fund directors are legally responsible to

- Evaluate and approve contracts and fees with the fund sponsor, the distributor and other service providers.
- Select the accountants
- Oversee the compliance of the securities transactions
- Check the NAV valuation process

We focus, here, on the responsibilities that are the most relevant from the economic viewpoint. The rationality and the asymmetry of information play a major role. If investors were well-informed and rational, they would disinvest of funds that are not well-governed and that have high fees compared to their performance. However, investors are subject to many cognitive frictions and there exist many layers of informational asymmetry.

In the respective dataset analysed by the authors, the average board has approximately 9 members with about 6 independent members. Still, there is considerable variation amongst boards and all funds have more independent directors than what is required by the regulator. On average directors sit on 16 boards within a fund family and, in principle, work exclusively for one fund sponsor. In order to evaluate the effect of board concentration, Tufano and Sevick measure the so-called independent director's asset share. It is calculated by dividing the assets under oversight by the board member, by the total assets under management by the sponsor. For instance, if total AUM of the sponsor amount to 100 Mio $ and the board member sits on 4 boards of funds having 20 Mio $ AUM, then the asset share of the board member is 0.8.

This measure is then used to calculate the board and sponsor concentration measures. The board concentration measures the fraction of assets overseen by a typical board and sponsor concentration the average fraction overseen by the average director of the sponsor. In the respective sample, the average sponsor concentration is 0.8. Nevertheless, there is strong variation amongst fund sponsors as the average concentration shifts between 0.1 and 1! A typical independent director earns around 60,000$ and is compensated through different board memberships provided by one sponsor. As already mentioned, he is thus partly dependent on the sponsor's decision to renew his board membership.

The question is whether and how this will impact the level of fees. Still, fund fees might be impacted by other factors as well. The following factors are known to impact the level of fees.

- Economies of scale which might exist at the fund and sponsor level. There is evidence (Figure 3 to Figure 6 in the Appendix) that fund costs in percentage decrease with AUM. At the fund family level there might exist synergies in terms of research.
- The age of the fund. The literature documents a relationship between age and fees, even though the mechanisms are not clear yet.
- Performance. In Principle a positive relationship could be expected, but empirically the relationship seems negative.
- Investment styles. As Figure 7 indicates, total production costs differ for different kinds of Funds.
- The distribution channel might impact fees as the cost of marketing the fund varies for different investment vehicles.

Those variables are typically used to statistically isolate their impact on fees independently of board governance issues. After careful analysis, Tufano and Sevick find economies of scale at the fund as well as the sponsor level. They also find a positive relationship between fund age and fees, which might be explained by more experienced funds charging higher fees. We know already that there is no or a negative relationship between performance and fees. After isolating all those factors, the authors find out that board structure matters. Notably, larger boards charge higher fees, whereas for a given size, boards with a higher fraction of independent directors charge lower fees. The share of assets overseen by independent directors is also negatively related with fees. The funds whose independent directors sit on many boards of the sponsor have lower fees.

BEHAVIOURAL ASPECTS OF THE FLOW-PERFORMANCE RELATIONSHIP AND MARKET SHARES

From the foregoing section, it is evident that the Flow-performance relationship plays a major role for the (Berke & Green, 2004)) rational adjustment mechanisms to work. As the Flow-Performance relationship is not linear, the so-called rational industry equilibrium is unlikely to be realized. In order to better understand the industry adjustment mechanisms it is thus important to better understand the behavioural aspects underpinning the Flow-performance relationship. A first explanation that comes to mind is that the relationship could be explained by the Kahneman-Tversky Prospect Theory Framework. More recently, some authors have also analysed learning effects to explain the empirically observed convexity of the Flow-Performance relationship.

In a recent paper, (Chang et al, 2016) analyse how the Flow-Performance relationship could be related to the so-called disposition effect. One of the largely documented phenomena in Behavioural Finance is the disposition effect (Shefrin & Statman, 1985). Individual investors tend to sell winning stocks but stick (too long at least) to losing stocks. Even though the phenomenon seems well-documented there doesn't seem to exist a commonly accepted explanation. (Chang et al, 2016) suggest an explanation together with a clever approach to test it through the analysis of selling and buying decisions of stocks and funds. They first note that the disposition effect is reversed in the case of funds as investors have a greater tendency to sell losing funds and invest in good performing funds. As already alluded to earlier, this is evident from the performance-flow relationship, even though the relationship is not linear but convex. More generally, this positive relationship seems to hold for existing as well as new investors.

The above-mentioned authors note that the disposition effect can be explained with Festinger's Cognitive Dissonance effect (Festinger, 1957). Cognitive Dissonance typically arises when a Decision Maker takes decisions that are inconsistent with its self-image. For instance, for an investor to make losses on his investment is inconsistent with his self-image of being a good investor. The point is that as the long as the investor doesn't sell the asset which is losing in value, the losses are not realized and he can still convince himself that this is a short term dip and the market is wrong! This, let's say subjective conviction, thus enables him not to take the negative blow to his self-image. Cognitive Dissonance hence rationalizes the aversion for realized losses and the tendency to realize gains, the latter providing a positive blow to the self-image.

If we consider the fund industry, however, investments are typically delegated to a manager. This relationship is analysed through the Principal-Agent model. The investor (the Principal) can thus blame the manager (the Agent) if the performance is not good and/or the fund loses in value. In the case of a delegated investment there is thus no cognitive dissonance involved on the side of the investor. Those insights lead the authors to the following predictions.

- Non-delegated assets will exhibit a disposition effect and delegated assets display a lesser or reverse disposition effect. This difference is stemming from behavioural biases.
- If investor's cognitive dissonance is strong, it leads to a larger disposition effect in non-delegated assets and a larger reverse disposition effect in delegated assets (funds). Inversely for low levels of cognitive dissonance.
- For investors that focus more on the fund manager's role, the reverse disposition effect is stronger.

(Chang et al, 2016) analyse the (Barber & Odean, 2000) individual investor trading data and come up with the following three stylized facts.

- The disposition effect in stocks and the reverse disposition effect for funds is observable for the same investors at the same time.
- Across asset classes, a positive disposition effect is observed for non-delegated assets whereas a reverse disposition effect is observed for delegated assets.
- Passive funds, namely index funds, display a small positive disposition effect.

In a recent paper, (Khorana & Servaes, 2012) analyse Market Shares in the fund industry through the degree of competition and fees. The two main mechanisms of price competition at work are:

- The response of the investor to fund fees
- The way economies of scale feed into a reduction of fees

Recall that in the rational (Berk & Green, 2004) framework, expected abnormal performance should be equalized with fees across funds. Even performance of skilled managers, in equilibrium, should respect this condition as inflows will increase AUM and abnormal performance will decrease with AUM till the condition is fulfilled. Second, it is well-known that the performance-flow relationship is non-linear (Lynch & Musto, 2003) and this leads to a situation where AUM can be increased by launching a set of funds whose future performance is random, but where some by chance will outperform. Figure 9 in the Appendix illustrates the idea .
4 funds are randomly launched (the blue arrows) and the probability distribution, represented in green, is the same for all the funds. Their future random performance realization will be one of the points on the performance axis. Due to the convexity of the Flow -Performance relationship this leads to a non-linear impact on inflows.

This, let's say anomaly, creates incentives for fund sponsors to inflate their product offerings, apart from any product innovation. The demand-side mechanism that might explain the creation of differentiated products is narrow framing. This mechanism might actually lead to product innovation by the fund family, even though such differentiation is difficult to reconcile with the Efficient Market Hypothesis and Rational Expected Utility maximization. Anyway, it is important to

distinguish product innovation through differentiation and the random launch of funds that might not even have different investment styles. Another interesting question is how marketing and distribution fees impact the market share. As Figure 1 and 2 in the Appendix show, distribution costs cover approximately 60-70% of total costs of European Equity Funds.
Finally, should the fund family specialize in specific strategies or should it diversify the fund family. Diversification of the fund family enables investors to more easily diversify whereas specializing in an investment style might lead to higher risk-adjusted performance. Both might lead to an increase in market share.

In the dataset analysed by (Khorana & Servaes, 2012), the number of fund families increases from 134 to 584 from 1976 till 2009. During the same time, the average family increased the mean number of funds offered from 2.87 to 19.61! Fund families also seem to have become less focused in terms of investment objectives. The industry has become slightly more fragmented, but fees as a percentage of AUM have increased. Empirical results indicate that price competition is at play as funds that significantly increase fees subsequently observe a significant decrease in market share. As far as the effect of performance on market share is concerned, its effect is mainly mediated through the presence of a top 5% performing fund in the fund family. The authors measure that the presence of a top 5% fund in a given family leads to an increase of market share of 42%.

The diversity of product offerings also has a bigger impact in terms of market share than the focus on specific investment styles. The innovation of new products, which is measured through the launch of new funds typically has a positive impact on the market share. The opening of a new fund leads to an increase in the market share of 8.6%. The effect, however, decreases with higher levels of innovation.

INCENTIVES TO GENERATE FUND PERFORMANCE AND THE DISTRIBUTION CHANNELS

Recent research also points to the fact that the distribution channels might have an impact on fund performance, in the sense that funds distributed through more intermediaries perform less. Basically, if funds are sold directly to investors they face more incentives to perform than if they are sold through brokers. Let's focus more in detail on the process.
Typically, standard retail funds that are sold to non-sophisticated investors can be classified into two broad categories:

- Unbundled Portfolio Management Services without involvement of investment advice
- Bundled Portfolio management and investment advice

In most advanced economies, at least 50% of mutual fund shareholders keep an ongoing relationship with financial advisors. Surveys indicate that quite many investors value face to face contact and a trust relationship with an investment advisor. A recent paper even indicates that investment advisors can be viewed as Money Doctors (Gennaioli et al. 2015)) in the sense that they help investors feel more safe about their investments and thus the latter feel safe in taking more risks. Investors using full brokerage services also seem to be less experienced (Chalmers & Reuter, 2012).

This potentially leads to differences in competition within two market segments, namely the direct-sold and the broker-sold channels. The broker-sold segment, in principle, should be less subject to competitive pressure as the investors are less experienced and less likely to grasp the difference between raw returns and risk adjusted returns. This leads (Del Guercio & Reuter, 2016) to the assumption that inflows of broker-sold retail funds should be less sensitive to alpha. They classify funds as direct-sold or broker-sold when at least 75% of their assets are sold through share classes focussing on that segment.

As the fund industry is typically structured around fund families it is important to understand how fund families market their funds. They might use different channels for different funds and/or different customers. The authors use the (Del Gercio et al, 2010) approach to evaluate the fraction of assets each family distributes through the respective channels. Most Mutual Fund Families mainly focus on one of both distribution channels and are quite specialized. A few distinguishing features can be observed across market segments.

- There are more actively managed funds in the broker-sold segment
- The average broker-sold fund manages fewer assets
- Broker-sold funds have significantly fewer assets.
- Broker-sold funds have higher average expense ratios, more or less equivalent to 12b-1 fees.
- Average monthly after-fee returns of broker-sold funds underperforms those of direct-sold funds.

In order to understand those facts, the Performance-flow relationship is of highest importance. Remember that as AM fees are fixed, the revenues of the Asset Manager increase with the inflow to the fund. Basically, Investment Funds are in competition to attract inflows. As investors in the broker-sold segment are less sophisticated, they are likely to focus on raw returns rather than on different estimates of alpha. Del Guercio and Reuter thus test for difference in performance-flow relationships between the direct and broker-sold segment. They use data on US domestic equity funds from January 1993 till December 2004. They indeed find that fund flows in the direct-segment channel are significantly more sensitive the risk-adjusted returns than fund flows in the broker-sold channel. The estimates imply that a standard deviation in alpha leads to an increase in inflows over the next year of 6.18% in the direct-sold segment, whereas as the increase of inflows is only 0.59% in the broker-sold segment.

In order to evaluate the impact in terms of revenue consider the following example. To make things straightforward to calculate. Imagine that the AUM are 1bn $ and the AM fees make up 1% of AUM. For the direct sold fund, the increase in inflows would be 61.8 million $ whereas for the broker sold fund it would be 5.9 million $ only. The increase in revenue from fees would be 618,000 $ for the direct sold funds and 59000$ for the broker-sold funds. The incentives to hire teams and invest in technology are thus much weaker for broker-sold funds. If at all, the incentives to generate alpha are much lower in the broker-sold segment. Conversely, broker-sold funds have incentives to increase exposure to systematic risk as non-sophisticated investors are focussing on raw returns.

Del Guercio and Reuter’s research indeed indicates that direct-sold funds exhibit higher after-fee risk-adjusted returns. The difference is approximately 10 basis points per month and increases to approximately 17 bps for small-cap funds. The authors suggest the difference between more balanced and small-cap funds can be viewed as an indication of more active management. They test this assumption more formally through the (Kacperczyk et al, 2008) return gap measure. The latter measures the difference between the fund’s actual gross return and the gross return implied by the fund’s lagged reported holdings. The authors find that a great part of the differences in performance can be explained through more active management as measured through the return gap. For the overall sample, the difference in return gap between direct-sold and broker-sold funds is approximately 6% and for small cap-funds it is approximately 12%. In order to evaluate whether direct sold funds are more actively managed the authors also use the (Cremers & Petajisto, 2009) active share measure. The latter is the fraction of fund *i*'s assets that have to be traded to obtain a portfolio that mirrors fund *i*'s benchmark. Evidence points out that direct-sold funds are more actively managed than broker-sold funds. The differences in returns between direct-sold and broker-sold funds is much higher for more actively managed small-cap funds. What about systematic risk exposures such as beta? In principle, broker-sold fund managers have a higher temptation to increase systematic risk as their investors are more sensitive to raw returns. Indeed, broker-sold funds seem to have higher betas than direct-sold funds.

Another issue concerns the number of different styles provided by fund families. In principle this fund variety is inconsistent with the standard rationality paradigm of Expected Utility maximization and the Mutual Fund Separation Theorem but can be explained within the Behavioural Economics framework, notably through Mental Accounting. There is evidence that quite many investors value "one-stop shopping" in the sense that they prefer to invest in a large fund family which offers a variety of investment styles. However, (Ciccorello et al, 2006) indicate that this convenience costs investors through lower risk-adjusted returns. According to (Massa, 2003), such lower performance is the result of diseconomies of scope in the generation of fund variety and fund performance. Given the greater sensitivity to risk-adjusted performance of investors in the direct-sold channel, it can be expected that the families will offer a narrow range of investment styles. Indeed, the study indicates that direct-sold funds offer on average 2.2 different Morningstar styles whereas broker-sold funds offer 3.5. Families specializing in direct-sold funds are also more specialized by investment styles than broker-sold funds. Direct-sold funds thus take better organizational decisions in order to generate risk-adjusted performance for investors.

A further well-documented fact is that funds outsourced to subadvisors are less performant (on a risk-adjusted basis). They should thus be less attractive for direct-sold funds. Indeed, broker-sold funds are much more likely to delegate to a subadvisor. The research indicates that around 22% of broker-sold funds are delegated to subadvisors as compared to 12% for direct-sold funds. These results are thus consistent with the theoretical prediction that families concentrating in providing direct-sold funds do not increase the scope of offered styles at the cost of risk-adjusted performance, conversely to broker-sold funds.

OUTSOURCING OF MUTUAL FUND MANAGEMENT: RISKS AND PERFORMANCE

Most of the investors delegate the management of their wealth to intermediaries that manage investment funds. Hence, from 1980 to 2007, the percentage of American households owning mutual funds rose from 5.7 to 43.6% (Investment Company Institute, 2007). The fraction of total equity owned by those investment vehicles has been increasing steadily and has recently increased to 32.4% by the end of 2007 (French, 2008). The fact that those investments are delegated to third parties leads to an agency problem. In that respect, the optimal compensation rules have received a lot of attention (Stracca, 2005). The role of the organizational structure, however, has not received that much attention. A paper by (Chen et al, 2010) analyzes how outsourcing of the fund management may impact managerial incentives and thus the performance of investment funds.

Although, in many cases the management company and the investment advisor are one unit, there are situations where the management of the fund is outsourced to an external investment advisor. The latter is then not involved in the operational day to day work and distribution. The management company retains the marketing and distribution fees and the external advisor obtains the pure management or advisory fees. In principle, the board of directors of the management company then monitors the fund's activities and risks. The Management company can at any time close down the fund or replace the investment advisor. The latter can either advise or manage funds. It is worthwhile noting that the investor typically does not know that the management is outsourced. The question is whether the outsourcing of the fund impacts the performance and, if, for what reason? (Chen et al, 2010) study the impact of fund outsourcing and consider a fund to be outsourced if one of the investment advisors is not affiliated with the Management Company.

An emerging literature also analyzes the influence of mutual fund families on mutual funds. In that respect, investors seem to choose a fund family first and then invest in one of the funds of the fund family (Massa, 2003). This, let's say behavioral bias, leads fund families to offer a greater degree of product differentiation. This excessive degree of product differentiation is actually a puzzle for the standard finance theory as the risk aversion of the investor should not influence the composition of the risky portfolio. This fact is known in the literature as the asset allocation puzzle (Canner *et al,* 1997). Of course, this product differentiation can, at least, partly be explained by behavioral aspects. (Gasper et al, 2006) note that fund families might subsidize the performance of a favorable fund in the family at the expense of another fund. What is worth highlighting, here, is that the organizational structure of the fund industry matters. Even though, this might seem evident for any practitioner, it has not been addressed by academics, except by (Chen et al, 2004). The fundamental issue is the boundary of the firm in a multi-task agency framework such as the one in (Holmstrom & Milgrom, 1991). How does the nature of the organization, notably the degree of integration, impact performance?

(Chen et al, 2010) aim at analyzing the performance of outsourced funds, they first analyze whether outsourced funds significantly differ from in-house funds in terms of characteristics. In their respective dataset, they distinguish outsourced funds by the fact that the investment advisor's name differs from that of the family complex. They identify roughly 56% of fund-year observations as being in-house managed and 44% as being outsourced. Some funds might be unidentified. Basically, there is a growing trend toward more outsourcing, but the degree of outsourcing does not seem to depend on the investment style that is followed. Note, also, that the number of funds has more than doubled from 1994 to 2007, which again is difficult to reconcile with the standard rationality paradigm in Finance. Figure 10 in the Appendix reports the characteristics of the fund families analyzed by (Chen et al, 2010).

The number of families has increased from 1994 to 2000, to slightly decrease afterwards. This goes together with a rather similar increase and decrease in the number of funds per family, a topic that would be interesting to analyze in more detail. Still, the average number of funds per family is roughly 8. The percentage of families that outsource to some degree is around 43 with again the same peak around the years 2000-2004. Interestingly, the average fraction of funds outsourced has increased steadily. The last column indicates the concentration in their respective investment styles of the families of funds. The table indicates that fund families specialize and follow the styles in which they have expertise.

Let's focus for a moment on the characteristics of the Mutual Funds in the sample. The expense ratios as a fraction of the year-end total net assets (TNA) do not differ from in-house funds and amount to approximately 1.3%. Fund turnover, which is defined as the minimum of purchases and sales over average TNA for the respective calendar year, is around 87.6% for in-house funds and 81.4% for outsourced funds. This doesn't seem to be a significant difference. Yet, outsourced funds tend to be somewhat younger and the total load they charge is somewhat lower. Fund in- and out-flows also do not depend on the outsourcing status. This seems quite straightforward as the typical investor doesn't know whether the fund is outsourced or not. Note that as smaller funds are more likely to be outsourced, it is important to adjust for fund size when analyzing performance.

In order to get a grasp of the performance as a function of risk factors, the authors benchmark fund performance using the famous Carhart 4-factor model (Carhart, 1997). As we already know, the latter is a combination of the Fama-French 3-factor model and the momentum factor discovered by (Jegadeesh & Titman, 1993). The authors also add two other factors: 1) The Morgan Stanley Capital International Index (MSCI) return covering Europe, Australia and Far East, 2) the Lehmann Aggregate Bond Index (LABI) return. Each month funds are then sorted into two respective portfolios dependent on whether they are outsourced or managed in-house. Equity funds are also sorted from non-equity funds as their return drivers differ. Funds are then sorted as a function of TNA, in order to isolate the effect due to the size of the fund. The time series of monthly returns is then used to estimate factor loadings. Figure 10 in the appendix provides an overview of the results.

Interestingly, the factor loadings are more or less the same for out-sourced and in-house funds. Outsourced funds, however, have smaller alphas accounting for around 60-80 basis points (bps) per annum lower than in-house funds. Here, it is noteworthy to highlight that the average alpha for in-house funds is already negative. As fund performance might be correlated with other observable fund characteristics, the authors also provide a regression between fund performance, outsourcing indicators as well as other characteristics supposed to eliminate other effects. They are the so-called control variables. The regression methodology is the one suggested by Fama and McBeth (1973). The regression specification takes the following form:

$$FUNDRET_{i,t} = \mu + \phi OUTSOURCED_{i,t-1} + \gamma X_{i,t-1} + \varepsilon_{i,t} \tag{13}$$

Where $FUNDRET_{i,t}$is the alpha of the fund *i* in month *t* adjusted by the typical Carhart and Fama-French factor adjustments. $OUTSOURCED_{i,t-1}$ is an indicator variable that indicates whether or not the fund is outsourced. The set of control variables in $X_{i,t-1}$ consist of the following: size, expense ratio, age, total load, flows.

Outsourcing has a negatively statistically significant impact on performance, whatever the benchmark adjustments used. The underperformance of outsourced funds compared to other funds is between 40 and 60 basis points per year. A few questions arise. How does performance differ for in-house and outsourced funds when the funds belong to the same family? Furthermore, does fund performance of an investment advisor differ when he manages on his own behalf rather than for an outsourced fund. (Chen et al, 2010) test this assumption by first testing whether the number of funds in the family impacts the probability to outsource. They find a positive impact using logistic regressions. If this impact is included as variable in the regression discussed above, the negative impact on performance of outsourcing is even stronger. Outsourcing reduces performances by 1.03% to 1.64% per year!

Theories of the firm (Holmstrom, 1999)) suggest that outsourced funds, not being under the direct control of the family, should face stronger incentives. First, the probability of closure following poor past performance should be higher. Second, risk taking should be monitored more closely than for in-house funds. The authors analyze the sensitivity of fund closures to past performance. This is typically done with the following logit regression specification:

$$Prob\left(Closed_{i,t} = 1\right) = \Lambda\left(\mu + \lambda Z_{i,t+1}\right) \tag{14}$$

Where $Closed_{i,t}$ is a dummy variable that takes value 1 if fund *i* is closed in year *t* and 0 otherwise. $\Lambda(.)$ is the logistic cumulative distribution function and is the vector of coefficients multiplied by the vector of fund characteristics. A fund is considered closed if it does not have the full set of twelve month returns. The fund characteristics are the standard ones considered above but an indicator whether the fund is outsourced and whether it is in its model style is added.

The aim is to test the (Chevalier & Ellison, 1999) hypothesis which leads to expect that fund closures are more sensitive to poor past performance for outsourced funds than for in-house funds. Indeed, the results indicate that the odds for an outsourced fund to be closed are much higher than for an in-house managed fund. Still, other types of factors might impact the results. For instance, outsourced funds are smaller and younger and thus face steeper incentives due to those characteristics. Even after controlling for those characteristics, econometric evidence still indicates that outsourcing leads to steeper incentives.

Contractual theories also predict that families should more closely monitor the risk taking of outsourced funds. Risk–taking deviation measures are suggested in (Chevalier & Ellison, 1999). The first variable if the deviation of the fund's beta from the average beta of funds in the same class. (Chen et al, 2010) calculate beta's using six factor models. The beta deviation risk measure is then calculated by taking the square-root of total squared deviations from the style means of the factor loadings. The second measure is the idiosyncratic risk calculated from the 6-factor model. The authors than re-estimate the probability to close with the risk-taking variables included and risk-taking is more penalized for outsourced funds. This also leads to less risk-taking from outsourced funds.

INTERNATIONALIZATION AND OUTSOURCING OF MUTUAL FUND MANAGEMENT

In a recent paper (Chuprinin et al, 2015) analyze outsourcing in a more international framework. Their approach is based on a distinction between two types of fund management status and they distinguish:

- Funds that are managed and marketed by the same financial group, called in-house.
- Funds managed on behalf of other financial groups called outsourced.

Basically, the management company is supposed to perform the management of the portfolio and the fund family the distribution and marketing of the fund. For in-house managed funds, the management company and the fund family are thus the same. The authors analyze so-called open-end actively managed equity funds from 2001 till 2008.

In order to detect outsourced funds, the difference in reporting between two datasets is exploited. One dataset, namely FactSheet/LionShare report the management company and the financial group that manages the fund, whereas the other dataset, Morningstar Direct, report the fund families of each respective fund. If the fund family in the second dataset differs from the financial group disclosed in the first dataset, the fund is diagnosed as outsourced. Other sources are then checked to verify that the supposedly outsourced fund is not managed by a subordinated structure. This leads to an indicator (dummy) variable that can be used in an asset pricing model to evaluate the impact of outsourcing. The asset pricing model used by the authors is the standard Carhart 4-factor model, but they also add a specification with an added liquidity risk factor based on the Amihud (2002) concept of illiquidity. Country-specific illiquidity risk factors are based on the methodology of (Liang & Wei, 2012). As usual, abnormal performance is measured by comparing the excess gross return to the one forecasted by the asset pricing model. However, in order to distinguish fund performance related to outsourcing from performance stemming from other characteristics, the following standard control variables are added:

- FundSize which is measured through the log of the fund's TNA
- MgmtCompanySize as the log of 1 plus the TNA of all funds excluding fund *j*.
- Expenses as the annual expense ratio of fund *j*.
- Fund Age which is the time since inception
- Past Retrun which is the last 12 month return
- Volatility measured trough the SD of the returns over the last 12 months
- Shareclasses, namely the number of share classes in the respective fund
- Inst Shareclass, indicating whether the fund had a share class open to only Institutinal Share holders.

Broadly speaking a quarter of funds is outsourced and different sources such as (Chen et al, 2013) and (Del Guercio & Reuter, 2014) seem to broadly accord with that number. The characteristics of the control variables also seem to be consistent with those of foregoing studies. (Chuprinin et al, 2015) analyse differences in performance as a function of whether the fund is outsourced or not. The impact of outsourcing is measured through a dummy variable and the following panel regression is implemented.

$$Performance_{j,t} = \beta_1 Outsourced_{j,t} + \beta_2 Controls_{j,t} + \alpha_t + \alpha_S(+\alpha_m) + \varepsilon_{j,t} \qquad (15)$$

Where the *j* indexes the funds, $Outsourced_{j,t}$ is the dummy variable and $Controls_{j,t}$ are the control variables discussed earlier. $\alpha_t,\ \alpha_S\ and\ \alpha_m$ indicate time, style and management company fixed effects.

Estimates indicate that outsourced funds underperform in-house funds by approximately 7 bps per month if we compare performance with the Carhart 4-factor model. On an annual basis this makes up approximately 84 bps and corresponds to approximately 57% of fund's annual expenses. Investment companies thus seem to treat their in-house funds more favourably than the outsourced funds they manage.

What are the mechanisms that might lead to such preferential treatment?
Two broad mechanisms are suggested in the literature:

- The strategic allocation of IPO stocks or the privileged use of information for in-house funds
- Cross-trading activity between in-house and outsourced funds

Let's focus on the IPO allocation, a mechanism already documented in (Gaspar et al, 2006). As management companies have long-term relationships with underwriters, they potentially have information about interesting IPO stocks. Moreover, IPO stocks often outperform at the beginning, a phenomenon known as the IPO Puzzle. It is highly likely that management companies might allocate IPOs to their in-house funds rather than their co-managed outsourced funds. In order to evaluate the impact of IPO allocations, the authors measure the weight of IPOs in the respective in-house and outsourced funds. The differences between those weights indicates that outsourced funds are allocated approximately 44% fewer IPO stocks than in-house funds. The preferential treatment due to IPO allocations seems to explain around a third of the return gap between funds of different status.

Is there any favouritism in terms of exploited investment opportunities?
Suppose that a management company has some information about future potential stock movements. The question is whether it is exploiting these opportunities systematically for in-house funds. This is checked by estimating the correlation between buy and sell positions at time *t* and

future stock returns between *t* and *t+1*. For quarterly data, the correlation between buy trades and future stock returns is around 35% larger for in-house funds, confirming the assumption that the best investment opportunities are mostly allocated to in-house funds.

Last but not least, let's consider cross-trading between affiliated funds. (Gaspar et al, 2006) discuss how cross-trading can affect performance either through favourite prices between affiliated funds (instead of through exchanges) or liquidity provision in case of distress. According to (Chuprini et al, 2015) the cross trading between affiliated funds is twice higher than the unconditional cross trading (between unrelated funds). Moreover, the cross trading is especially higher during times when funds are distressed.

Here the (Holmstrom & Milgrom, 1987) Principal-Agent model can be applied, where the fund family is viewed as the Principal and the Management Company the Agent. The question is whether outsourced funds just get treated less favourable than in-house funds or whether preferential treatment can be seen as incentive provision by the Principal. The first assumption is called the "inferior product hypothesis", the second the "incentive hypothesis". Moral Hazard plays a key role and the eventual degree of Moral Hazard is directly related to the monitoring ability of the Principal and the bargaining power of the Agent. The mechanisms are the following: If the Principal cannot adequately monitor, the Agent is shirking and the former has to provide a preferential treatment to the Agent. Second, the more market power the Agent has, the more bargaining power he has to get a preferential treatment. The authors measure monitoring ability through the use of a common language by the Principal and the Agent. Note that the country location of the Principal and the Agent matter in this case. Bargaining power of the Agent is measured through the market share of the subcontractor firm in the outsourcing market.

(Chuprinin et al, 2015) develop the following empirical approach to discriminate between both hypotheses. First, they note that the task of managing in-house and outsourced funds is complementary. Indeed, the agent acting as a subadvisor can use his skills and investment research to manage other in-house funds. This is the spillover effect and this positive effect increases with the number of in-house fund managed by the subcontractor. Given that the positive spillover effect increases, the preferential treatment should be less under the "incentive hypothesis" for subcontractors having a higher fraction of in-house funds. Inversely, the "inferior product" hypothesis predicts that for subcontractors having a higher fraction of in-house funds, the latter are more important and this raises their preferential treatment. A more extreme version of the "inferior product" hypothesis leads to profit reallocation between in-house and outsourced funds. If that were the case, we should observe negative contemporaneous correlations between the two types of funds. Basically, superior performance of in-house funds is subsidized by outsourced funds. Under the incentive hypothesis however, preferential treatment can be seen as a kind of in-kind compensation with the tacit agreement of the Principal and contemporaneous returns between the two types of funds should be positively correlated. Finally, preferential treatment can be seen as a substitute for subadvisory fees. The incentive hypothesis entails a positive relationship between subadvisory fees and outsourced performance. .Moreover, the relationship between affiliated in-house and outsourced funds is expected to be negative.

The results are the following.

- Fund families that share a common language with the management perform better and the difference is about 2.5% per year! The impact is event stronger if we focus on cross-border relationships between fund families and management companies. This might be related to higher information asymmetry due to cultural and geographical distance.

- The market power is measured through the market share of the respective management company in the outsourcing segment of a respective style. More exactly, the ratio of aggregate TNA of outsourced funds run by the management company in a given style to the total TNA of all outsourced funds in that given style. When a management company has a higher market share in a given outsourced investment style, the preferential treatment tends to be higher. A one standard deviation increase in market power approximately lowers the performance of outsourced funds by 1.8% per year.
- If we measure the fraction of outsourced funds as the ratio of the aggregate TNA of outsourced funds run by the management company to the aggregate TNA of all funds run by the management company, there is a negative relationship between the performance gap and the fraction of own account activity of the management company.
- Superior performance of in-house funds is not associated with contemporaneous poor performance of affiliated outsourced funds. Subadvisory fees bear a positive relationship with outsourced fund performance. A percentage point increase of subadvisory fees goes together with a 2% performance increase. Broadly, the results suggest that preferential treatment and subadvisory fees are substitutes.

Liquidity Provision in Fund Families

The fund industry is typically structured around so-called fund families. Fund sponsors then launch a few types of funds available to investors. Two arguments can be suggested to explain this fact. First, from a demand-side perspective, different kinds of investors have different kinds of tastes. Second, from an industry supply-side perspective, it might be interesting for a sponsor to launch a few funds, which maximizes the chance that one particular fund has outstanding abnormal performance. It is well know that this fund then attracts huge inflows. Indeed, the non-linearity of the inflow/outflow relationship as a function of performance is well documented in the literature. Good abnormal performance has a high impact on inflows, whereas under normal market circumstances, fund outflows are not very reactive to bad performance. This leads to a convex flow relationship with respect to performance.

Still, the supply-side as well as the demand-side explanation implies a relaxation of the standard behavioral paradigm in financial economics. Standard rationality postulates have to be relaxed. This doesn’t mean that investors are necessarily irrational, but rather that they are less rational or rational in a different way, than suggested by standard textbooks. For instance, they could be inattentive or evolutionary rational. For recent contributions on evolutionary rationality the reader can consult (Lo ,2017) and (Smith, 2009). The deviations with respect to full rationality, however, are potential weaknesses that can be exploited by financial decision makers. As already mentioned, fund families potentially exploit such weaknesses by launching a whole set of funds. They thus exploit individuals’ cognitive biases with respect to probabilities. But fund sponsors might exploit investors in other different ways.

In that sense, a recent research paper by (Bhattacharya et al, 2012) analyze whether fund families provide insurance to other funds belonging to the same family. Fund families are typically organized as a collection of legally independent entities sponsored by a management company. From a legal perspective, the different investment vehicles are different legal vehicles. The temptation, however, is strong to provide insurance to funds that are facing redemptions due to bad performance. Better performing funds would thus provide insurance to less performing funds, at the cost of well performing funds’ shareholders. (Bhattacharya et al, 2012) examine the assumption by analyzing the so-called affiliated funds of funds, the latter being funds that only invest in other funds in the fund family.

Funds in families that have affiliated funds of mutual funds (AFoMFs) have two groups of investors: the AFoMFs and the outside investors. In order to understand how and if the family provides insurance, we have to distinguish the flows. The total flow to each fund *j* in family *k* during period *t* is measured in the standard way.

$$Flow_{j,k,t}^{Total} = TNA_{j,t} - TNA_{j,t-1} * \left(1 + r_{j,t}\right) \tag{16}$$

r is the net of fees return. Note that the equation presumes that the cash flows arrive at the end of the respective period. In order to isolate the impact of the AFoMFs on fund flows, the authors calculated the change in the number of shares held by AFoMF *i* in fund *j* multiplied by the Net Asset Value (NAV) of fund *j*. If the values are aggregated for all AFoMFs within the respective family, we get the total flow going from the family to fund *j*.

$$Flow_{j,k,t}^{AFoMF} = \sum_{i=1}^{n_k} \Delta shares_{i,j,t} * NAV_{j,t} \tag{17}$$

Where $n_k$is the numver of AFoMFs in family *k*. Delta shares is the change in the number of shares of fund *j* held by AFoMF *I* between date t-1 and t. The flow from outsiders can now be calculated by removing the flow of AFoMFS from te total flow:

$$Flow_{j,k,t}^{Outside} = Flow_{j,k,t}^{Total} - Flow_{j,k,t}^{AFoMF} \tag{18}$$

Funds can now be sorted in each family into deciles according the outside flows. Funds in the lowest decile 1 are considered distressed and experience severe redemption requests. Decile calculations are adjusted each period. For each decile, the authors calculate the average flows from AFoMFs (scaled by TNA) and the fraction of their positions. Figure 11 in the Appendix indicates the flows analyzed by (Battacharya *et al,* 2012).

The dashed line divides the funds into positive and negative flow subgroups. It is clear from the Figure that, generally, there is a positive relationship between outside flows and AFoMF flows. For the decile of distressed funds this is not the case anymore and Figure 11 provides first evidence that AFoMFs invest in distressed funds.

The authors also implement an econometric analysis to evaluate a model consistent with (Ellison, 1997; Sirri & Tufano, 1998; Dela Guercio & Tkac, 2002). In those models, the flow performance relationship is positive even though non-linear. In order to test the impact of funds being distressed, they add an indicator variable for the first decile. In general, outside flows have a positive impact on AFoMF flows, except for the first decile, as indicated by the sum of beta1 and beta2. This is clear indication that AFoMFs provide liquidity to distressed funds. Yet, if AFoMFs provide support to funds facing liquidity problems, this behavior should not be observed for funds that are less in need of liquidity support.

Some might argue that the flow movements are a result of asset allocation decisions. Imagine that the asset class value has decreased below target value, then an asset manager will have to inject inflows to keep the target weight. To test this assumption (Bhattacharya et al, 2012) construct an "asset allocation" variable which is defined as the difference between the weight of the AFoMFs in the asset class of fund *j* at the beginning of the respective portfolio period and the target rate. The target weight is presumed equal to the average weight over time. A regression analysis using this variable indicates, even though somewhat significant, does not remove the strong impact of liquidity provision. What kind of funds does the family help? According to (Gaspar et al, 2006) fund families strategies tend to help high value or high fee funds in order to maximize total revenue. This seems rational in the sense that star funds attract flows to other funds of the family as well (Nanda et al, 2000). Is the liquidity provision really beneficial to the respective funds? In order to test this assumption, the following approach developed in (Edelen, 1999) is used. Fund alphas, either determined by a Carhart 4-factor model or a (Fung & Hsieh, 2004) 7-factor model, are regressed with the following specification:

$$\alpha_{j,t} = \beta_0 + \beta_1 * I_{j,t} + \beta_2 * I_{j,t} * Flow_{j,t}^{AFoMF} + controls + \varepsilon_{j,t} \quad (19)$$

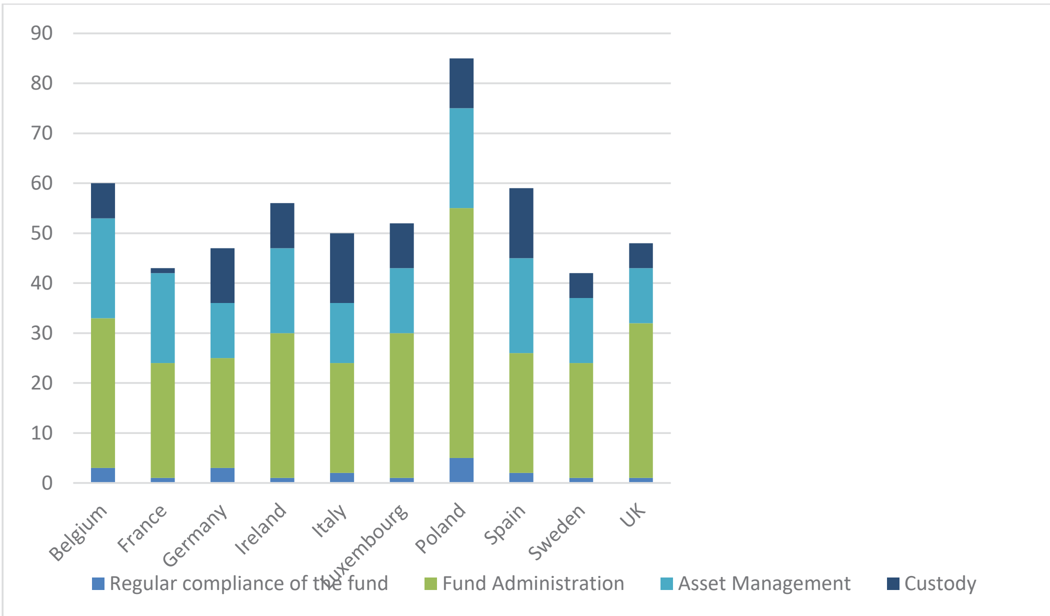
Where $\alpha_j$denotes the abnormal performance of fund j, $Flow_{j,t}^{AFoMF}$is the flow that fund j receives from the AFoMFs in its family. I_j is an indicator that takes value 1 if fund is distressed. The control variables are diverse standard control variables used in this literature. The results indicate that $\beta_1$is significantly negative and amounts to -0.0009. This means that, indeed, huge redemptions have a negative impact on performance. This is probably due to costly fire sales. β_2, however, is statistically significant and positive which implies that AFoMF flows are counteracting the negative impact on performance. The estimated $\beta_2$equals 0.0481 and a 1% increase in AFoMF flow during a stress period lowers the negative impact by 4.8 bps.

CONCLUSION

The industrial organization of the fund industry differs a lot from the standard portfolio theory predictions. The Mutual Fund Separation Theorem suggests that most investors would hold diversified portfolios and that fund industry would consist of big diversified funds. The Behavioral Finance literature mostly focuses on market and pricing anomalies but does not analyze the industrial organization of the fund industry. The common thread of the paper is that behavioral aspects explain the origin of the complexity of the fund industry. More broadly, there is a need for a kind of industrial organization of the finance sector based on behavioral aspects.

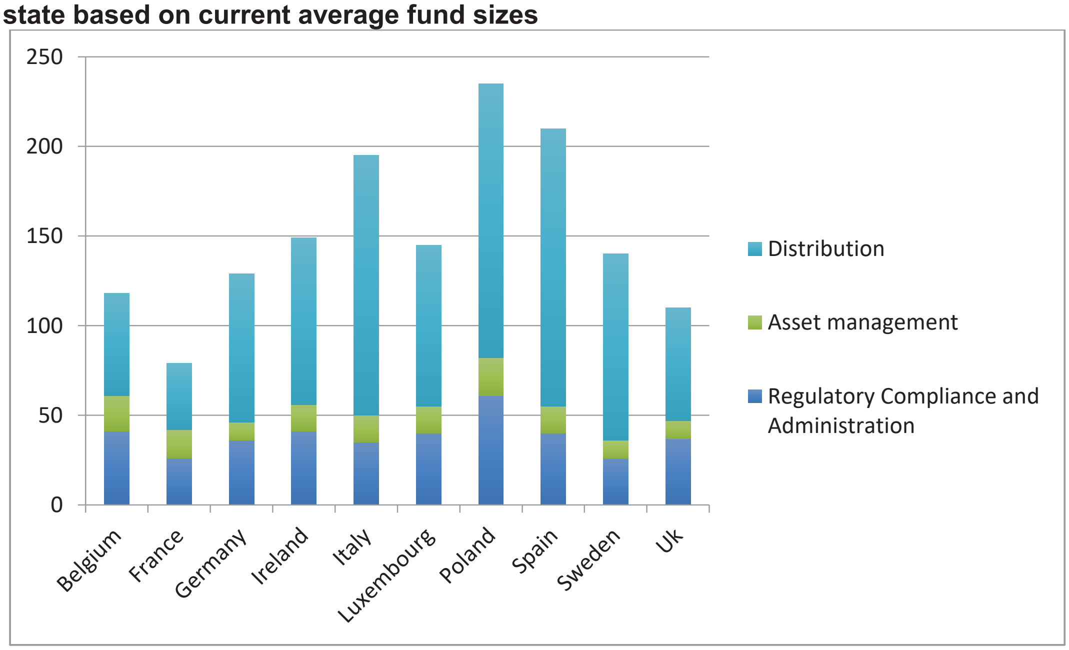
APPENDIX: FIGURES AND GRAPHS

Figure 1: Comparison of total production costs for equity funds by member state based on current average fund sizes



Based on CRA data

Figure 2: Comparison of total production and distribution costs for equity funds by member state based on current average fund sizes



Based on CRA data

Figure 3: Cost curves as a function of AuM and efficiency versus scale effects

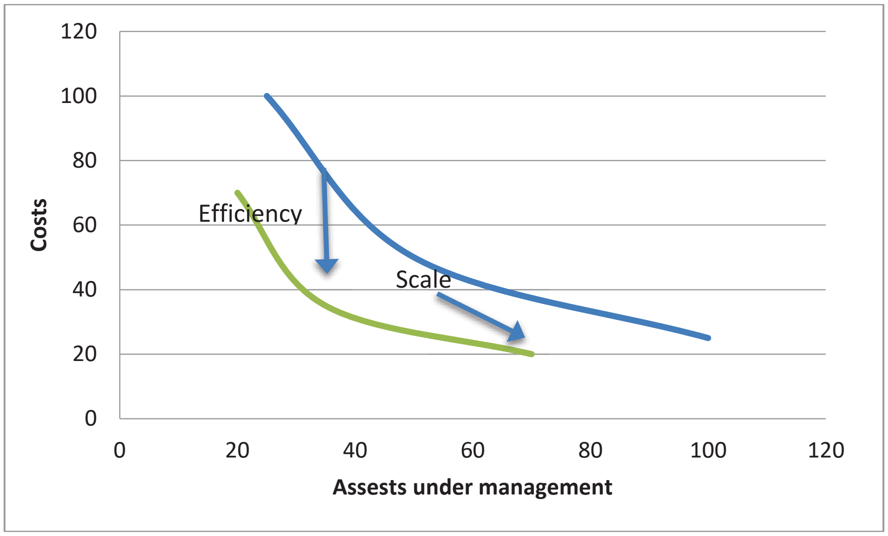
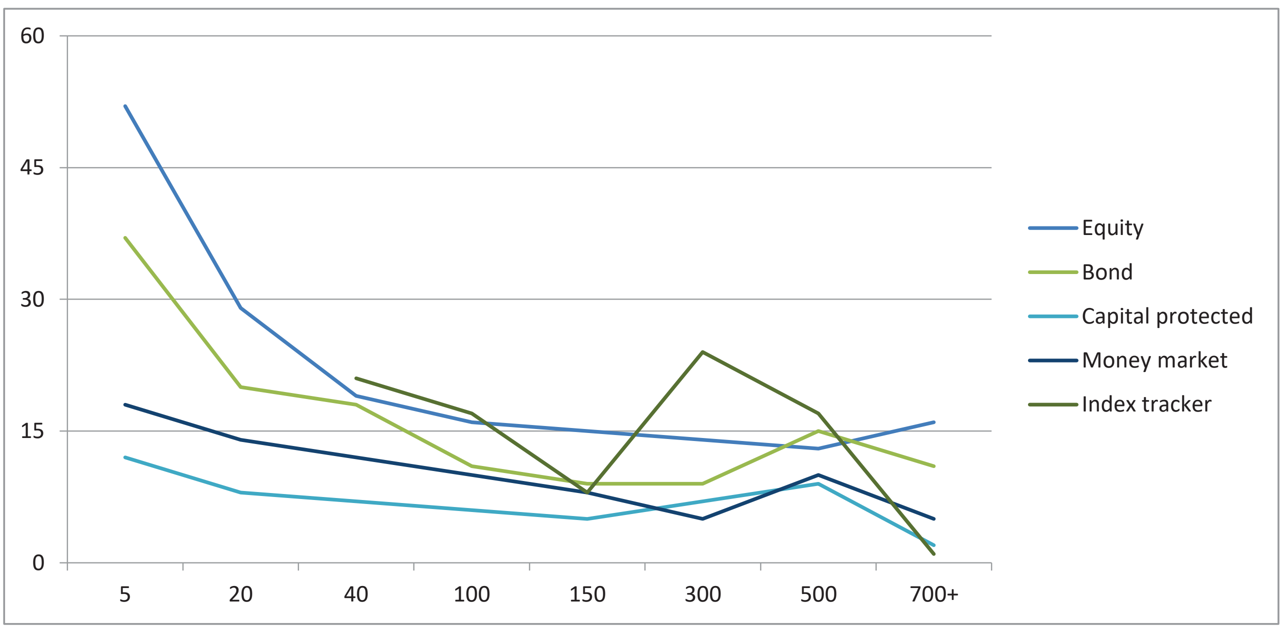
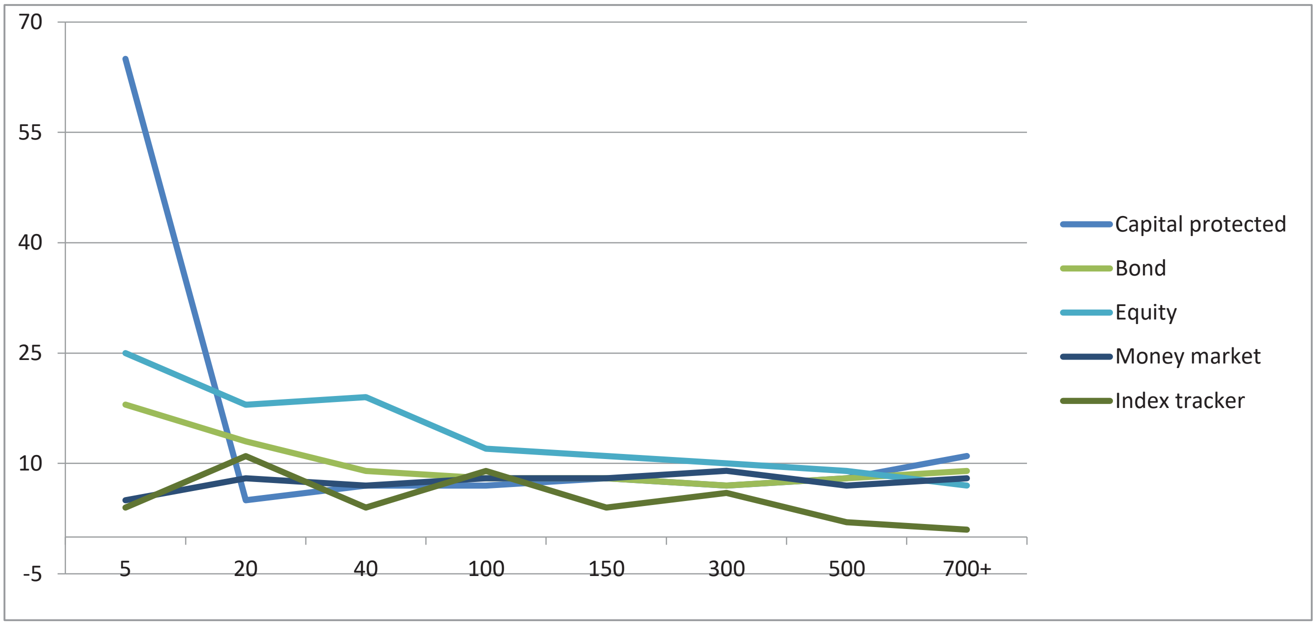


Figure 4: Fund Accounting costs for different Fund types



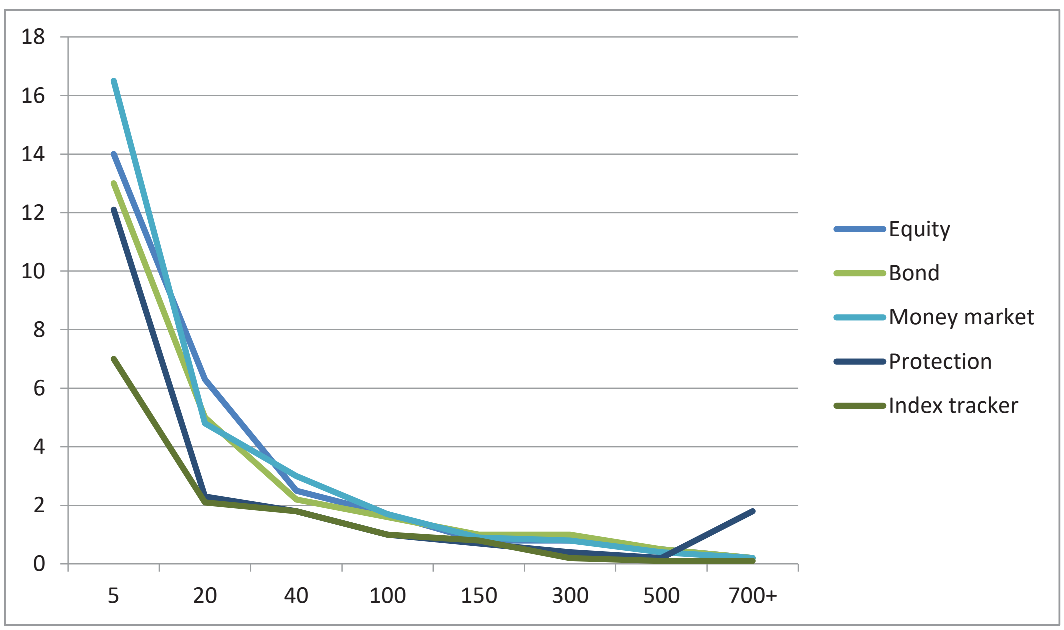
Based on CRA Analysis

Figure 5: Custody costs for different fund types



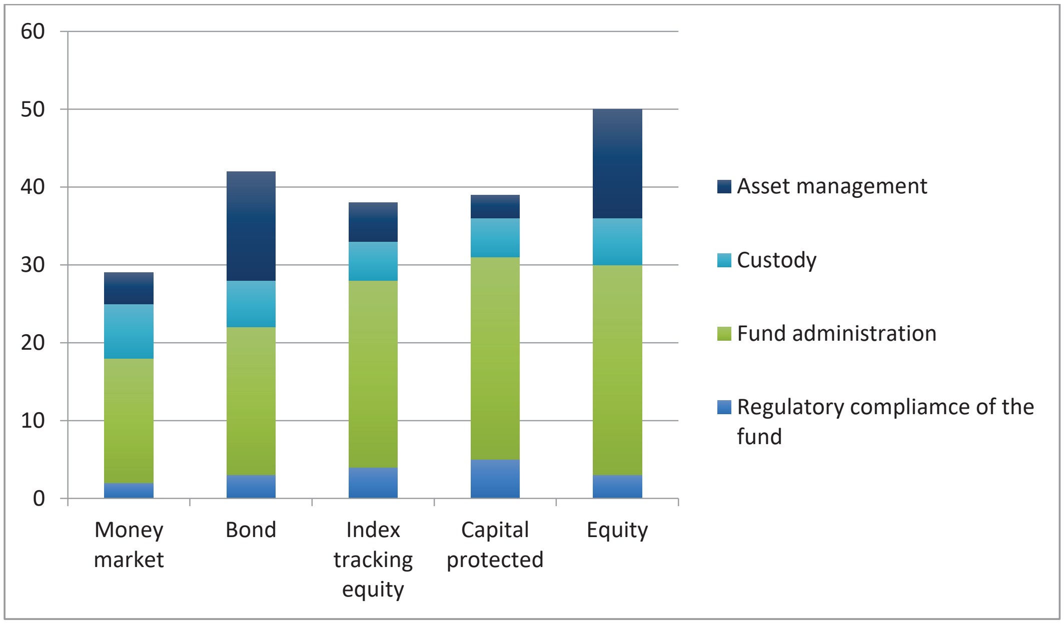
Based on CRA Analysis

Figure 6: Audit costs for different Fund types



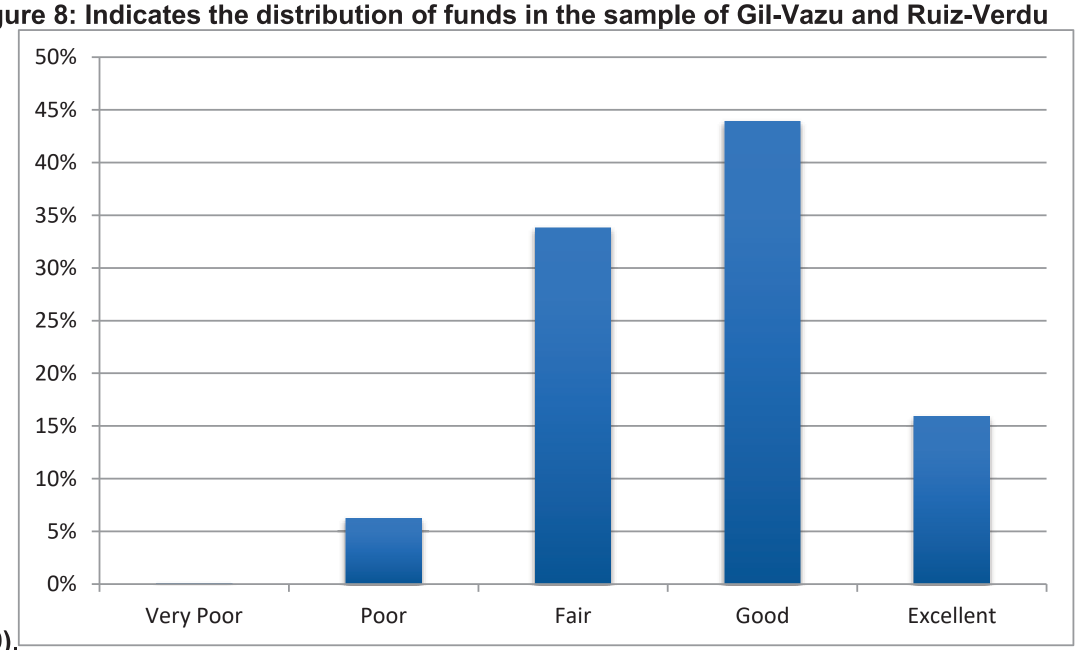
Based on CRA Analysis

Figure 7: indicates the different total production costs for different kinds of Funds.



Based on CRA Analysis

Figure 8: Indicates the distribution of funds in the sample of Gil-Vazu and Ruiz-Verdu (2009).



Graph based on data in : Gil-Bazu, J. and Ruiz-Verdu, P. (2009)

Figure 9: Non-linear Flow-Performance and effect on fund families

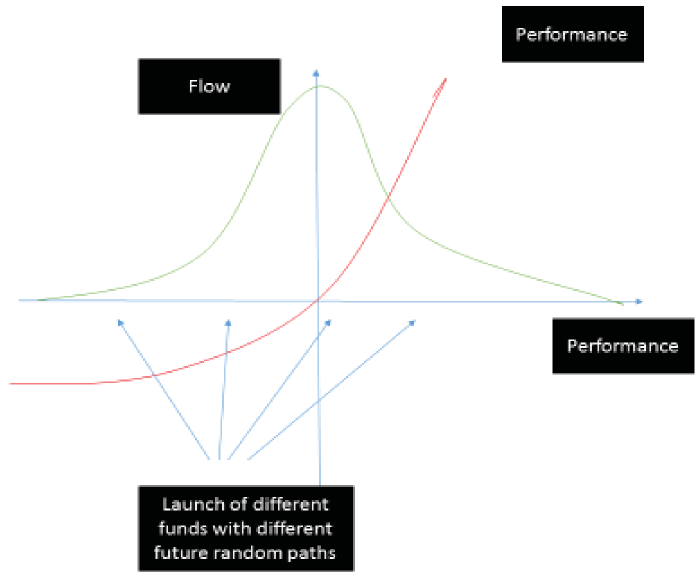
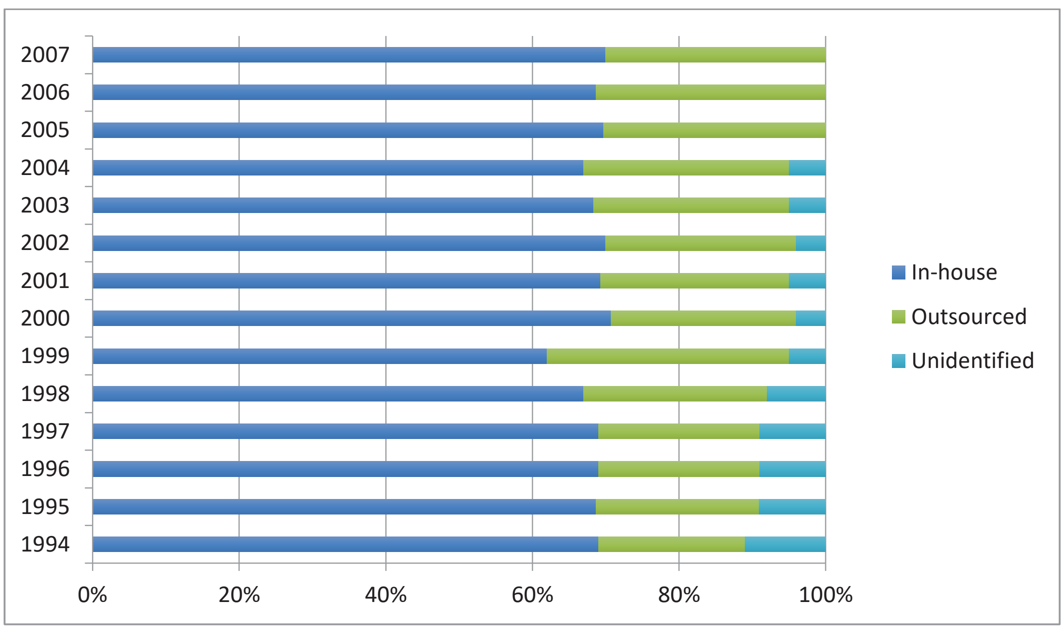
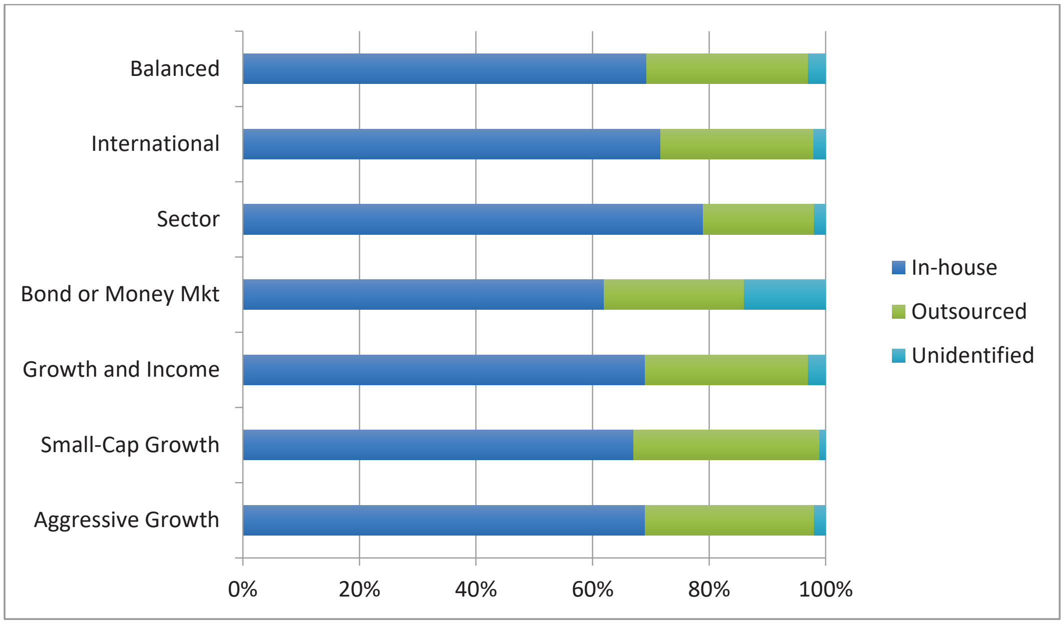


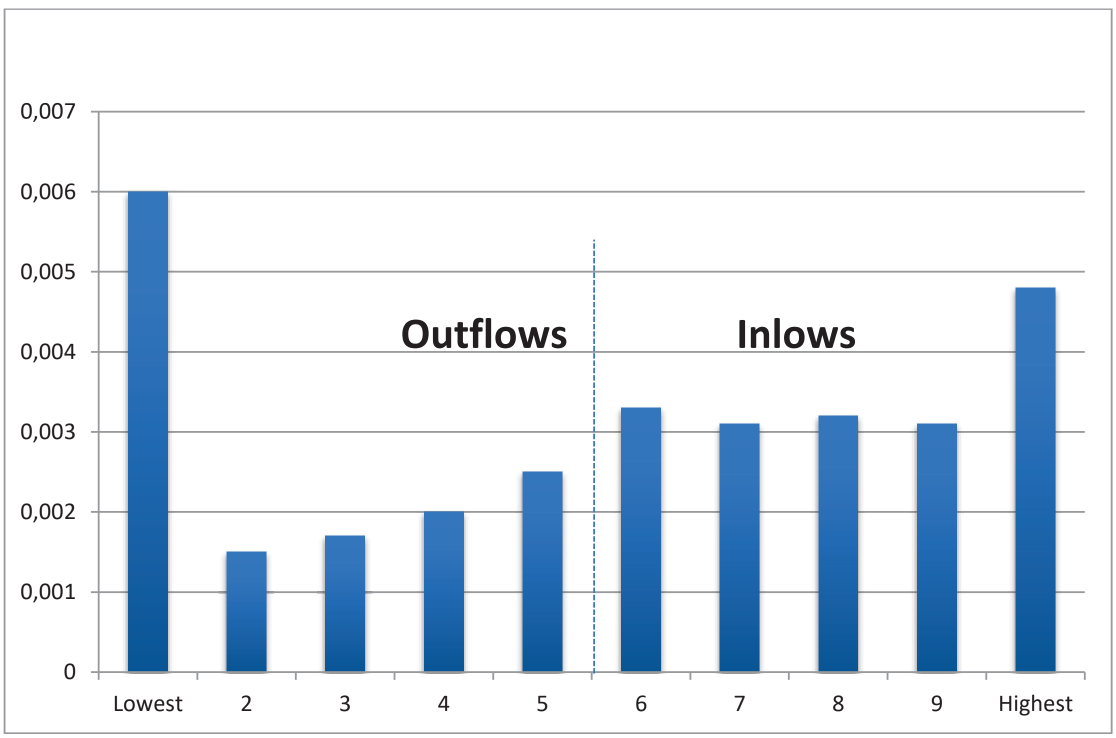
Figure 10: Identification of Mutual Fund Management





Based on tables in Chen *et al*. (2010)

Figure 11: Indicates the average AFoMF flow per decile of outside-flows.



REFERENCES

Admati, A. & Pfleiderer, P. (1997). Does it all add up? Benchmarks and the compensation of active portfolio managers,.*The Journal of Business*, Vol. 70, 323-350.
Agarwal, V., Daniel, N. & Naik, N. (2009). Role of Managerial Incentives and Discretion in Hedge Fund Performance. *The Journal of Finance*, Vol. LXVI, Oct. 2009.
Alexander, G. & Baptista, A.M. (2008). Active portfolio management with benchmarking: Adding a value-at-risk constraint., *Journal of Economic Dynamics and Control*, Vol. 32, 779-820.
Almazan A., Brown K., Carlson M. & Chapman, D. (2004). Why constrain your mutual fund manager. *Journal of Financial Economics* 73, 289-321.
Amarante, M. (2009). Foundations of neo-Bayesian statistics. *Journal of Economic Theory*, Vol. 144, 146-2173.
Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time series effects. *Journal of Financial Markets 5*, 31–56.
Arya, A. & Mittendorf, B. (2005). Offering stock options to gauge managerial talent. *Journal of Accounting and Economics* 40, 189-210.
Barber, B. & Odean, T. (2000). Trading is hazardeous to your wealth; The common stock investment performance of individual investors. *The Journal of Finance*, Vol. 116, 261-291.
Barberis, N. (2018). Psychology-based Models of Asset Prices and Trading Volume, *Yale School of Management Working Paper*.
Barberis, N & Thaler, R. (2003). A survey of behavioral finance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*, vol. 1. North-Holland, Amsterdam.
Berk, J. & Green, R. (2004). Mutual Fund flows and performance in rational markets. *Journal of Political Economy 112*, 1269-1295.
Bhattacharya, U., Jung, L. & Pool, V. (2012). Conflicting Family Values in Mutual Fund Families. *The Journal of Finance*, Vol. 68(1), 173-200.
Bollen, N. & Pool, V. (2009). Do Hedge Fund Managers Misreport Returns: Evidence from Pooled Distributions. *The Journal of Finance*, Vol 64(5), 2257-2288.
Borgonovo, E. and Marinacci, M. (2015). Decision Analysis under ambiguity. *European Journal of Operational Research* 244, 823-836.
Borgonovo, E., Cappelli, F., Maccheroni, F. & Marinacci, M. (2018). Risk analysis and decision theory : A bridge. *European Journal of Operational Research* 264, 280-293.
Brown, S. & Goetzmann, W. (2003). Hedge Funds with style. *Journal of Portfolio Management 29,* 101-112.
Bruhin, A. , Fehr-Duda, H. & Epper, Th. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distorsion. *Econometrica*, Vol. 78, 1375-1412.
Capital Market Risk Advisors (2008). *Risk Principles for Asset Managers*
Cerraiera-Voglio, S., Maccheroni, F., Marinacci, M. and Montruccio, L. (2013) « Classical Subjective Expected Utility », Proceedings of the National Academy of Science USA, 110, 6754-6759.
Canner, N, Mankiw, G. & Weil, D. (1997). An Asset Allocation Puzzle. *The American Economic Review*, Vol. 87.
Carhart, M. (1997). On the persistence of Mutual Fund performance. *Journal of Finance* 52, 57-82
Chalmers, J. & Reuter, J. (2012). What is the impact of financial advisors on retirement portfolio choices and outcomes. *NBER Working Paper 18158*
Chang, T., Solomon, D. & Westerfield, M. (2016). Looking for Someone to Blame: Delegation, Cognitive Dissonance and the Disposition Effect. *The Journal of Finance*, Vol. LXXI., February 2016, 267-301.
Charles River Associates (2006). Potential cost savings in a fully integrated European investment fund market. *Report of the Directorate General for Internal Market and Services*

Chen, J., Hong, H., Huang, M. & Kubik, J. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review* 94, 1276-1302.
Chen, J., Hong, H & Jiang, W. (2013). Outsourcing mutual fund management: Firm boundaries, incentives and performance. *Journal of Finance* 68, 523-558.
Christofferson, S. and Musto, D. (2002) "Demand curves and the pricing of money management", review of Financial Studies, Vol. 15.
Chuprinin, O, Massa, M. & Schumacher, D. (2015). Outsourcing in the International Mutual Fund Industry: An Equilibrium View. *Journal of Finance 70*, 2275-2308.
Chevalier, J, & Ellison, G. (1999). Career concerns of mutual fund managers. *Quarterly Journal of Economics 114*, p. 389-432.
Ciccotello, C., Miles, J. & Walsh, L. (2006). Should investors choose funds from focused families. *Financial Services Review* 15, 247-264.
Cornell, B. & Roll, R. (2005). A delegated-agent asset pricing model. *Financial Analyst Journal*, Vol. 61, p.57-69.
Cremers, M. & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies 22*, 3329-3365.
Cremers, M., Petajisto, A. & Zitzewitz, E. (2012). Should benchmark indices have alpha? Revisiting performance valuation. *NBER Working Paper 18050*
Del Guercio, D., Reuter, J. & Tkac, P. (2010). Broker incentives and mutual funds market segmentation. *NBER Working Paper Series*.
Del Guercio, D. & Reuter, J. (2014). Mutual Fund Performance and the incentive to generate alpha. *Journal of Finance* 69, 1673-1704.
EC (2006). *White Paper on enhancing the Single Market framework for investment funds*.
Fama, E. & McBeth, J. (1973). Risk, return and equilibrium: Empirical tests. *The Journal of Political Economy* 81, 607-636.
Fehr-Duda, H., Epper, Th. & Schubert, R. (2011). Risk and rationality: The ffects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, Vol 78(1), 14-24.
French, K. (2008). The cost of active investing. *Journal of Finance* 63, p. 1537-1573.
Fama, E. & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*, 3-56
Festinger, L. (1957). *A Theory of Cognitve Dissonance*, Stanford University Press.
Fung, W. & Hsieh, D. (2004). Hedge Fund benchmarks: A risk based approach. *Financial Analysts Journal 60*, 60-80.
Gasper, J.-M., Massa, M. & Matos, P. (2006). Favoritism in Mutual Fund Families? Evidence on Strategic Cross-Fund Subsidization. *The Journal of Finance* 61, 73-104.
Gennaioli, N., Shleifer, A. & Vishny, R. (2012). Money Doctors, *The Journal of Finance,* Vol. 70(1), 91-114.
Getmansky, M. (2012). The life cycle of hedge funds: Fund flows, size, competition, and performance. *Quarterly Journal of Financial Economics 2,* 1-53.
Getmansky, M. , Lo, A.W. & Makarov, I. (2004). An econometric model of serial correlation and illiquidity in Hedge Fund returns. *Journal of Financial Economics 74,* 529-609.
Ghirardato, P., Maccheroni, F. & Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory,* 118, 133-173.
Ghirardato, P., Maccheroni, F. & Marinacci, M. (2005). Certainty Independence and the separation of utility and beliefs. *Journal of Economic Theory,* 118, 133-173.
Gil-Bazu, J. & Ruiz-Verdu, P. (2009). The Relation between Prize and Performance in the Mutual Fund Industry. *Journal of Finance*, Vol 64(5), 2153-2183.
Gilboa, I. (2009). *Theory of Decision under Uncertainty*, Econometric Society Monographs, Cambridge University Press.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-additive Probabilities. *Journal of Mathematical Economics*, Vol. 16, p. 65-88.
Gilboa, I. & Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, Vol. 18, p. 141-153.
Goetzmann, W., Ingersoll, J. & Ross, S. (2003). High-water marks and hedge fund management contracts. *Journal of Finance 58,* 1685-1718.
Gremillion, L. (2005). *Mutual Fund Industry Handbook: A comprehensive guide for investment professionals*. Wiley.
Holmstrom, B. (1999). The firm as subeconomy. *Journal of Law, Economics and Organization* 15, 74-102.
Holmstrom, B. & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica 55,* 303-328.
Holstrom, B. & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization* 15, 74-102.
Investment Company Institute (2007), *Mutual Fund Fact Book*
Jegadeesh, N. & Titman, Sh. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48, 93-130.
Jurek, J. & Stafford, E. (2015). The Cost of Capital for Alternative Investments. *The Journal of Finance*, Vol 70, No 5.
Kahneman, D. & Tversky, A. (1979). Prospect Theory : An Analysis of Decision under Risk. *Econometrica*, Vol. 47, p. 263-291.
Kahneman, D. & Tversky, A. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty 5,* 297-323.
Khorana, A., Servaes, H. & Tufano, P. (2005). Explaining the size of the mutual fund industry around the world. *Journal of Financial Economics.* Vol 78, 145–185
Khorana, A. & Servaes, H. (2012). What Drives Market Share in the Mutual Fud Industry. *Review of Finance*, Vol. 16, p. 81-113.
Kuhnen, C. (2009). Business Networks, Corporate Governance, and Contracting in the Mutual Fund Industry. *The Journal of Finance*, Vol LXIV, No 5
La Porta, R., Lopez-de-Silanes, F, Shleifer, A., & Vishny, R. (1998). Law and Finance. *Journal of Political Economy*, vol. 106, p. 1113-1150.
Liang, S.X. & Wei, J.K. (2012). Liquidity risk and stock returns around the world. *Journal of Banking & Finance 36*, 3274–3288
Lo, A. (2017). *The Adaptive Markets Hypothesis: Financial Evolution at the Speed of Thought* Princeton University Press.
Malmendier, U. (2018). Behavioral Corporate Finance. *NBER Working Paper No*. 25162
Markowitz, H. (2014). Mean-variance approximations to Expected Utility. *European Journal of Operational Research,* Vol. 234, p. 346-355.
Massa, M. (2003), “How do family strategies affect fund performance? When performance maximization is not the only game in town”, *Journal of Financial Economics 67*, 249-304.
Nanda, V., Narayanan, M. & Warther, V. (2000). Liquidity, investment ability and mutual fund structure. *Journal of Financial Economics* 57, 417-443.
Oxera (2006).Current Trends in the Asset management Industry. prepared for the EU Commission, DG Internal Market and Services.
Panageas, S. & Westerfield, M. (2009). High-water marks: High Risk Appetites? Convex compensation, long horizons, and portfolio choice. *Journal of Finance* 64, 1-36.
Polkovnichenko, V. & Zhao, W. (2013). Probability weighting functions implied in option prices. *Journal of Financial Economics 107,* 580-609.
Polkovnichenko, V., Wei, K. & Zhao, F. (2018). Cautious Risk Takers: Investor Preferences and Demand for Active Management. *The Journal of Finance*, Vol 74(2), 1025-1075.
Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, Vol. 66, p. 497-527.

Roll, R. (1992). A mean/variance analysis of tracking error. *Journal of Portfolio Management*, Vol. 18, p. 13-22.
Savage, L. J. (1954). *The Foundations of Statistics*, New York: John Wiley and Sons, Second Edition 1972, Dover.
Schmeidler, D.(1989). Subjective probability and expected utility without additivity. *Econometrica*, 57, 571-587.
Sensoy, B. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, Vol. 92, p. 25-39.
Sharpe, W. (1992). Asset Allocation: Management Style and Performance Measurement. *The Journal of Portfolio Management 18(2)*, 7-19..
Shefrin & Statman (1985). The disposition to sell winners too early and ride losers too long: Theory and Evidence. *The Journal of Finance*, Vol. 40, p. 777-790.
Shefrin, H. & Statman, M. (2000). Behavioral Portfolio Theory. *Journal of Financial and Quantitative Analysis*, Vol. 35.
StatMan, M. (2004). The Diversification Puzzle. *Financial Analyst Journal 60(4),* 44-53.
Stoughton, N., Wu, Y. & Zechner, J. (2011). Intermediated Investment Management. *The Journal of Finance*, VOL LXVI, No3.
Stracca, L. (2005). Delegated Portfolio Management: A survey of the Theoretical Literature. ECB Working Paper Series No. 520.
Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies 25*, 68-85.
Tversky, A. & Wakker, P. (1995). Risk Attitudes and Decision Weights. *Econometrica*, 63,1255-1280.
Tufano, P. & Sevick, M. (1997). Board structure and fee-setting in the US mutual fund industry. *Journal of Financial Economics*, Vol. 46, p 321-355.
Verlaine, M. (2016) *The Economics of the Asset Management Industry: A Behavioral Economic Approach.* Editions Universitaires Européennes
Verlaine, M. (2018) “Risk Management for Funds: A decision Theoretic Approach”, Proceedings of the Decision Science Institute, Annual Meeting, Chicago 2018.
Wakker, P. (2010) *Prospect Theory for Risk and Ambiguity*, Cambridge University Press.
Yin, Ch. (2016) “The Optimal Size of Hedge Funds: Conflict between Investors and Fund Managers”, *The Journal of Finance*, Vol 71(4), 1857-1894.

DECISION SCIENCES INSTITUTE

Earnings Managements in Banks: An Analysis of Reduced Interest Rate Policies

José Luis Modena
Universidade Federal do Paraná
Email: jotamodena@gmail.com

Marcos Wagner da Fonseca
Universidade Federal do Paraná
Email: marcos.w.fonseca@gmail.com

Jorge Eduardo Scarpin
Concordia College
Email: jscarpin@cord.edu

Edilson Paulo
Universidade Federal do Rio Grande do Sul
Email: e.paulo@uol.com.br

ABSTRACT

This article's main objective is to identify the relationship between the reduction of the basic interest rate in Brazil (represented by Selic) and the management of results in banks. For the development of this analysis, this study utilizes a specific accrual model to estimate what portion of loan loss provisions (LLPs) can be managed by banks. The results of our survey reveal that when the Selic rate has been low, Brazilian banks have increased their use of discretionary loan loss provision (LLP) accruals, thus reducing their overall earnings from credit operations.

KEYWORDS: Earnings Management, Loan Loss Provisions and Interest Rate.

INTRODUCTION

In light of Accounting Choice Theory, exogenous factors may encourage banking organization managers to display opportunistic behavior. In this conception, considering that incentives in terms of practices related to earnings management may be derived from the external environment, where regulatory policy, contracts and capital markets may exert an influence (Watts & Zimmerman, 1978). Studies have already found, for example, that the earnings management is related to economic development (Riahi-Belkaoui, 2004; Baatour & Othman, 2016) and the quality of auditing (Tendeloo & Vanstraelen, 2008). On the other hand, given the wide range of external variables that can influence decisions related to accounting procedures, there still remain gaps in the literature that need to be explored. One of these concerns low Selic rates which, given this rate's relevance to Brazilian economic policy, as well as its impact on the banking spread, may encourage the management of bank earnings.

In Brazil, the Selic (Special Settlement and Custody System) represents a basic interest rate practiced in the economy, being used as reference for loans and financial investments. Part of the mandatory bank deposits, as well as some Brazilian public debt securities, are remunerated by this indicator. For this reason, interest rates applied by financial institutions are parameterized by the Selic rate.

The Selic rate plays a fundamental role in the macroeconomic sphere and is published for monetary policy purposes. This has been true ever since the adoption of inflation targets in 1999, in which decisions related to this indicator came to influence the money supply and, as a result, investment, consumption and production. It should be emphasized, however, that this prerogative is only possible due to the relationship that exists between the Selic rate and bank credit, which in turn, according to Carvalho (2005), is a mechanism for the transmission of monetary policy. In this mechanism, changes in the macro rate affect production through credit, inducing changes in banking reserves which can be driven by expansionary or contractionary monetary policy. In other words, depending on how it is applied through changes in the Selic rate, there will be a change in the operational volumes experienced by banking institutions. Due to this macroeconomic influence, changes in the Selic rate, in addition to affecting issues related to monetary policy, also affect the operational volume of entities that participate in financial intermediation.

Based on these macroeconomic influences on banking credit, it is assumed as a consequence that changes in the prime interest rate can also influence the results of financial institutions. Studies of this subject have found that the prime interest rate appears to be one of the determinants of earnings (Demirgüç-Kunt & Huizinga, 1999). In addition, it has been noted that alterations in this indicator are statistically related to the growth in the earnings of institutions linked to financial intermediation (Ho & Saunders, 1981; Saunders & Schumacher, 2000). Or in other words, it has been concluded that macroeconomic oscillations of interest rates directly influence banking margins in the granting of credit.

However, even though this empiric evidence assumes that alterations in the Selic rate have the potential to influence the reported earnings of financial institutions, there still have been no studies that make it possible to infer their capacity to encourage opportunistic management of the reporting of banking assets. In studies of the management of financial institutions conducted by Zendersky (2005), Santos (2007) and Dantas, Medeiros, Galdi and Costa (2012), the Selic rate was employed as the only control instrument in their statistical models. In general, however, few studies have sought to identify macroeconomic factors that may encourage the practice of earnings management (Almeida, Lopes & Corrar, 2011; Rodrigues, 2013). To Watts (1992), decisions related to accounting choices are influenced by economic issues, which indicates the need to explore these instigating factors which may lead to opportunistic behavior on the part of managers.

Based on these considerations, as well as the theoretical gap regarding the influence of the prime interest rate on the making of accounting decisions, we have formulated the following research question: **what is the relationship between a low Selic rate and earnings management in Brazilian banks?** Based on this question, this study seeks to analyze the relationship between low Selic rates and the use of discretionary LLP accruals. To do this, we have used a specific accrual approach based on the existing literature. To perform these tests, we use accounting data from financial institutions in the Brazilian market between the first quarter of 2010 through the third quarter of 2017 which have been classified as commercial and multiple banks and credit unions.

LITERATURE REVIEW

Earnings Management in Banks

In the banking sector, studies have identified asset accounting which offer the manager discretion, making it possible to manage earnings in an opportunistic manner. In this sphere, loan loss provisions (LLPs) play a prominent role. This type of accounting has been tested statistically rendering results that testify to its use in the manipulation of accounting reports.

The findings of McNichols and Wilson (1988) obtained results that make it possible to make inferences about the use of LLPs in earnings management. The study examined American companies that used large LLPs in relation to the profits they obtained. To capture evidence of earnings management, the authors used a specific accrual approach which later was used in other studies (Zendersky, 2005; Lobo & Yang, 2001). To estimate the non-discretionary portion of an LLP, the author used the initial balance of the LLP, the losses during period t and the losses during period t + 1. The results were significant, suggesting the use of these accruals in earnings management.

Lobo and Yang (2001) also deal with the use of accounting choices related to LLP reporting. To perform their statistical tests, the authors used time series regressions, obtaining three results: significant evidence in terms of income smoothing practices; the sign and coefficient were better specified by time series for each bank; and banks managed their earnings to conform with regulatory capital requirements.

The findings of Kanagaretnam, Lobo and Mathieu (2003) suggest that the discretionary nature of LLPs is used to smooth income. To do this, the authors divided financial institutions into groups, discriminating them based on their pre-income tax and pre-LLP earnings. The evidence indicates that these organizations increased their LLPs during good times to reduce their published earnings, thus “saving” earnings for subsequent financial reports.

Curcio, Simone and Gallo (2016) found that banks which were subject to stress tests by a supervisory authority demonstrated a greater propensity for smoothing income. To reach this conclusion, the authors developed a study of financial institutions that were operating in Europe during the financial crisis. The results revealed exogenous stimuli which could affect the decisions made by managers in terms of loan loss provisions (LLPs).

Moyer (1990), in turn, found that earnings management practices were used to conform with regulatory capital requirements. The author pointed out that in the United States, banks that do not fulfill these minimum capital requirements receive larger fines, which influences institutional earnings. Within this context, the author found evidence that suggests that LLPs and assets were being used to manipulate earnings to meet capital adequacy norms.

However, Beatty, Ke and Petroni (1999) sought to identify evidence of the existence of differences between discretionary bank accounting between public and private banks. The authors identified indices of smoothing income for banks that are listed on the stock exchange. They observed that private financial institutions, with earnings that were close to zero, were more likely to report losses than public financial institutions. This evidence suggests that financial institutions which are listed on the stock exchange have a greater incentive to report market gains.

To Tomy (2018), earnings management in banks can be useful in inhibiting new competitors from entering the market. The author argues that community banks increase their provisions for questionable credit in order to appear less profitable. The statistical evidence found helped reinforce the hypothesis put forward in this study.

Within the Brazilian context and using a specific accrual approach, Zendersky (2005) identified that the discretionary portions of LLPs are used by financial conglomerates in Brazil to manage earnings. To give greater robustness to the results found through his employed methodology, the author also performed tests with an alternative stage model. The results found in the two methodologies were not divergent, and thus corroborated his findings.

Goulart (2007) analyzed the fifty largest banks operating in Brazil between June 2002 and December 2006 in order to identify income smoothing practices. After performing statistical tests, he found evidence that banks were managing earnings through LLPs, which were the result of derivatives and revenues positively adjusted on the stock market. The results also make it possible to infer that according to the metric utilized, the LLP is the most often used instrument to manage bank earnings.

Prime Interest Rate and Bank Earnings

There is evidence that the prime interest rate is related to the bank spread. In Demirguç-Kunt and Huizinga's study (1999), for example, the authors sought to identify factors related to the bank margins practiced in financial intermediation. Based on a sample that included banks from 80 countries, it found a significant relationship between the macro rate and the bank spread, suggesting that banks demand greater margins in countries with higher interest rates. In Brazil, a study by Afanasieff, Lhacer and Nakane (2002) revealed that the factors that present the greatest explanatory power in terms of the bank spread are macroeconomic variables, represented by the Selic rate and the GDP. Studies by Jiang, Tang and Law (2003) and Dick (1999) also presented empirical evidence which observed that increasing interest rates resulted in elevated bank earnings. Based on these studies, it was found that the macro rate established in a country influenced the margins demanded by banks.

It should be emphasized that some studies have found an elevated financial margin in more underdeveloped nations. To Claeys and Vander Vennet (2008), this condition is a measure to compensate for exposure to risks due to scenarios characterized by greater instability. In an analysis of the earnings of financial intermediaries in some countries in Latin America, Brock and Suarez (2000) stated that macroeconomic uncertainty as well as the financial reserves in these countries are reflected in the practices of elevated banking earnings.

Another justification for the inclusion of the Selic as an indicator of the bank spread is interest rate risk. Oreiro, Paula, Silva and Ono (2006) explain that this risk associated with banking activity occurs due to the chance that the remuneration of the credit portfolio will not be sufficient to cover a positive oscillation in the prime interest rate, resulting in losses for the organization. This being so, they suggest that even with a low interest rate policy, the uncertainties associated with macro-instability may lead institutions to maintain an elevated bank spread to cover macroeconomic uncertainty.

In addition to these considerations in terms of the makeup of the bank spread, some studies deal with the passing on of interest rate increases by financial institutions. To Castro and Melo (2010), in pricing bank loans, Brazilian financial institutions respond more quickly to increases in the Selic rate than to decreases. Hannan and Berger (1991) have found that American banks display greater rigidity in relation to passing along interest rate hikes in terms of fundraising costs. That being so, this evidence suggests that in function of the velocity of the passing on of changes in the macro prime interest rate to the market, banks end up obtaining positive results, because they pass along these changes to financial applications, but delay reducing interest rates for financial loans. In this way, it is suggested that a decrease in the Selic rate produces positive effects on bank earnings.

However, it should also be noted that considerable oscillations in the Selic rate can also be reflected in bank spreads. This hypothesis was tested in the studies of Ho and Saunders (1981), which identified that volatility in the macroeconomic interest rate has an impact on bank margins. Saunders and Schumacker (2000), in turn, using a similar approach, investigated this relationship in six European countries and the United States and observed that a 1% variation in the prime interest rate generated a growth of 0.2% in the margin of financial intermediation practiced by banks. The authors suggested that volatility in the macro rate influences banks' net margin on interest rates, and that policy directives that seek to inhibit abrupt oscillations can have a positive effect on reducing bank spreads.

Nonetheless, considering positive factors in terms of the growth of institutional earnings from credit operations, opportunistic behavior by managers is expected in trying to reduce current earnings, transferring a portion of the earnings to subsequent periods. In this strategy, at a certain level of earnings it is considered appropriate to "store" earnings for more difficult periods (Paulo, 2007). Based on this assumption, it is expected that banks will increase their use of discretionary

LLP accruals, reducing bank earnings when the Selic rate is low. This is why we have established the following test hypothesis

> *H0a – There is a positive relationship between lower Selic rates and discretionary LLP accruals.*

METHODOLOGY

Delimitation of the Period and the Sample

The time period for this study spans from 2010 to September 2017. This initial restriction is based on the publication of BACEN Resolutions 3786/09 and 3853/10, which determined that financial institutions need to publish consolidated accounting records beginning in 2010, based on international standards emitted by the International Accounting Standards Board (IASB). The convergence of these standards makes a better comparison of these records possible, which should yield a greater homogeneity in the standards of the data collected, which explains the restriction determined in terms of the beginning of our collection period for the accounting information used in this study.

The delimitation of the period also serves as a parameter for the selection of the institutions examined in this study. Therefore, we initially defined the population examined in this study as financial institutions operating in Brazil which are classified as commercial banks, multiple banks with commercial operations, or credit unions that are not members of a conglomerate. A non-representative sample was selected from these institutions based on the information available from the BACEN.

In addition, to arrive at a representative sample for this study, we also used the criterion of the availability of accounting information, in order to make the operationalization of the employed model viable. Given this, of the 99 financial institutions that constitute this universe, we excluded 32 of them because they did not possess data for the variables under examination which thus left us with a sample of 67 banks. This selection, according to the accounting data published by BACEN, is representative of this larger universe, given that this group performed 99.36% of its credit operations and possessed 98.14% of its total assets. That being so, this sample is appropriate in terms of the established objectives.

Analytical Model

The operationalization of the current study was carried out by applying a statistical model in two stages. This methodology consists of first estimating discretionary accruals through the use of a statistical equation. The results of this equation were used as the dependent variable of a second equation to test the study's hypothesis. It should be emphasized that this methodological approach has been widely applied in both the national (Santos, 2007; Zendersky, 2005; Dantas *et al*., 2012), and international literature (McNichols & Wilson, 1988; Lobo & Yang, 2001; Beaver & Engel, 1996). The results of these studies have been discussed in the course of this research and support the methodology which we have adopted to test this study's hypotheses.

In terms of the first stage, we have opted to estimate accruals by focusing on identifying practices that indicate the management of earnings through specific accounting. To McNichols (2000), this methodology is an alternative to aggregate accumulations, and has the advantage of making it possible to identify key factors that influence accruals. Martinez (2001) points out another advantage which would be the chance to develop a model more appropriate for the problem under examination. Nonetheless, the author warns that this depends on the clarity of the accumulation which can be used to manage earnings.

In this sense, to address this condition, we have estimated accruals for loan loss provisions (LLP). To Beatty, Ke and Petroni (2002), this figure is one of the most manipulated areas of bank earnings. Corroborating the authors' inference, various studies in the literature have found evidence of the management of earnings through this accumulation (Mcnichols & Wilson, 1988; Moyer, 1990; Beaver & Engel, 1996; Lobo & Yang, 2001; Goulart, 2007; Santos, 2007; Zendersky, 2005). These findings suggest that this choice has the potential to meet the needs of this study.
In light of these considerations, we first estimated LLP accruals through Equation 1, whose structure is based on models that have already been used in other studies (Beaver & Engel, 1996; Lobo & Yang, 2001; Kanagaretnam *et al.*, 2003; Zendersky, 2005)

$$LLP_{it} = \beta_0 + \beta_1 CO_{it} + \beta_2 LLP_{it-1} + \beta_6 RCO_{it} + \varepsilon_{it} \qquad (1)$$

in which:
i = Bank
t = Period
LLP_{it}= Total LLP accruals.
CO_{it} = Credit operations.
LLP_{it-1}= LLPs during the previous period.
RCO_{it}= Revenues from credit operations.
ε_{it}= Standard error, which represents the discretionary portion of the LLP.

With the results of the equations used in the first stage, we will begin the procedure that seeks to test this study's hypothesis. To do this, we will use the 2 model, whose structure is represented below.

$$DLLP_{it} = \beta_0 + \beta_1 LWSEL_{it} + \beta_2 ROA_{it} + \beta_3 SIZ_{it} + \beta_4 PUB_{it} + \beta_5 BAS_{it} + \beta_6 PTC_{it} + \mu_{it} \qquad (2)$$

in which:
$DLLP$= Discretionary LLP Accruals
$LWSEL_{it}$ = Low Selic rate.
ROA_{it}= Return on assets.
SIZ_{it}= Variable which represents size.
PUB_{it}= Variable which indicates whether a bank is public or private
BAS_{it}= Basel Index
PTC_{it}= Companies that are publicly traded on the BM&F Bovespa stock market

The above mentioned model also makes it possible to test income smoothing practices performed by financial institutions in our sample. The confirmation of these practices is dependent on the statistical significance of the variable ROA_{it}, with positive parameters for Equation 2. Concerning their signs, we would like to emphasize that the expectation is justified by LLPs producing a reduction in accounted earnings. In this manner, a positive coefficient for the employed variable is interpreted as a smoothing of income, where in the face of elevated earnings, managers broaden the utilization of the discretionary portion of this figure.
In terms of the validity of the models, we have used tests designed to verify whether they meet our statistical assumptions. To do this, we confirmed the normality of the results of the equations, diagnosing multi-collinearity through the Variation Inflation Factor (VIF), homoscedasticity through White's test, and autocorrelation through the Durbin Watson test. These procedures give the inferences made greater robustness in terms of our results.

Variables Used

In Equation 1, the dependent variable represents the value listed as an LLP in the Earnings Statement. This value which is used as an estimate of discretionary accruals has been employed as a proxy for the management of earnings, and has also been used in other studies (McNichols & Wilson, 1988; Beaver & Engel, 1996; Lobo & Yang, 2001; Zendersky, 2005).

In terms of explanatory variables, the balance of credit operations (CO_{it}), registered as assets in financial institutions, was also used in the model. To Beaver and Engel (1996), credit operations from previous quarters also represent a risk to the financial institution, which suggests that an increase in the credit portfolio should follow the direction of LLPs. The findings of Kim and Kross (1998) confirm this expectation, revealing a positive coefficient in the correlation between these variables. That being so, it is expected that this explanatory variable will present a positive parameter in the employed regression model.

In relation to the variable LLP_{it-1}, which represents the initial balance of the LLP balance, which reduces assets, we would expect a positive relationship with the dependent variable. Lobo and Yang (2001) suppose that if a bank presents an elevated history of LLPs, it is less likely that the non-discretionary portion of this provision will be elevated in this period, due to the functioning of regulatory bodies, which will find it strange to have an increase in these provisions. However, the positive relationship observed in the model used by the authors contradicted this supposition, making them adherents to the arguments of Wahlen (1994), who suggests that the data presented by LLPs is used by investors to formulate future expectations. In Brazil, Zenderski's study (2005) also perceived a positive relationship between these two variables. This is why, we expect a positive coefficient for this explanatory variable.

The third variable of the equation (RCO), which represents the revenues from financial intermediation in credit operations registered in the earnings statement, is intended to express a bank's elevated exposure to risk. According to the banking regulations emitted by the National Monetary Board through Resolution 2682/99, banks are required to classify credit operations by order of increasing risk, with reviews being required in only a few special cases, such as, for example, a default on credit granted. In this manner, we suggest that the greatest impact of LLP expenses comes from operations that are contracted during the period, a fact that justifies the utilization of this variable in this model. In addition, it is expected that the presented coefficient will be positive, because according to previously realized studies (Lobo & Yang, 2001; Kanagaretnam, Lobo, & Mathieu, 2003; Tomy, 2018), where an elevated exposure to risk has been found through the credit portfolio, a positive relationship has been obtained.

In terms of Model 2, the $LWSEL_{it}$ is designed to test this study's hypothesis. To operationalize it, we use a dummy variable, using 1 for quarters which demonstrated low values for this Brazilian economic indicator and 0 for other cases. The information related to the Selic rate was obtained directly from the Central Bank website, employing the last rate before each quarter's close. Nonetheless, it is expected that this variable will be positive in Equation 2, keeping in mind the expectation that a decrease in this variable will have a positive influence on bank earnings.

In terms of the ROA_{it} variable, which represents the return on total assets, it was obtained by dividing net profits by the total assets registered at the beginning of the quarter. As mentioned in the previous section, this variable is intended to capture income smoothing practices employed by financial institution managers. To Kanagaretnam *et al.* (2001), banks tend to elevate the discretionary elements of LLPs in favorable times, economizing earnings to be later used in times of trouble. In line with the expectation mentioned by the authors, it is expected that this variable will be positively correlated with the discretionary portion of the LLP.

To represent the size of the financial institutions used in this statistical model, we use the SIZ_{it} variable which is obtained through the natural logarithm of the total assets. It is expected that this variable will be significant in the proposed model, indicating that the size of banking institutions influences earnings management practices. According to Curcio et al (2016), we would expect a positive coefficient for this variable.

We also have employed the PUB_{it} variable in this statistical model in order to distinguish public from private banks. To do this, we have operationalized this indicator through a dummy variable, using 1 for public banks and 0 for other types. In Santos (2007), even though there were no expectations in terms of the sign of this coefficient, this variable presented positive parameters and statistical significance. Nonetheless, it should be emphasized that the theoretical basis for this is still too incipient to indicate prospects for the sign of this variable.
The BAS_{it} variable was used to capture the relationship that exists between the need to conform to the regulatory capital requirements instituted through the Basel Accord and earnings management practices. In this sense, we have used a similar approach to that used by Shieves and Dahl (2003), in which they divided financial institutions into quartiles, using the Basel Index as a criterion for this stratification. In this manner, the first quartile selected the banks with the lowest indices of complying with regulatory capital requirements. To operationalize this segmentation of the statistical model, we used a dummy variable, with a value of 1 indicating organizations in the first quartile and 0 for the others. A negative relationship for this variable is expected in Equation 2, because the institutions with the greatest prominence in non-compliance with this regulatory parameter hypothetically would be more likely to reduce the use of the discretionary portion of accruals, as suggested by studies of Shieves and Dahl (2003) and Santos (2007).
The model also uses the control variable PTC_{it} in order to capture a possible relationship between companies that are publicly traded on the Brazilian stock market and earnings management. To Beatty *et al.* (1999), private companies are more likely to report smaller losses as compared to publicly traded companies. A possible justification suggested by the authors would be less asymmetry of information between the owners and managers of private banks, since contracts signed between these parties rarely are defined in terms of benchmarks. In relation to these arguments, it is expected that publicly traded companies will exhibit greater use of discretionary accruals. To operationalize the use of this indicator, we have used a binary variable, with a value of 1 used for companies which are publicly traded on the stock market and 0 for private companies.
Finally, we would like to mention that to verify the normality of the indicators used in these regressions, the variables LLP_{it}, CO_{it}, LLP_{it-1}, RCO_{it} and ROA_{it} were scaled by the amount of total assets at the beginning of quarter t (At_{it-1}). A summary of all the variables, their operationalization and their expected relationships is displayed in Text Table 1.

Variable	Operationalization	Expected relationship
LLP_{it}	$\frac{LLP_{it}}{At_{it-1}}$	
CO_{it}	$\frac{CO_{it}}{At_{it-1}}$	Positive (Beaver and Engel, 1996; Kim and Kross, 1998)
LLP_{it-1}	$\frac{LLP_{it-1}}{At_{it-1}}$	Positive (Lobo and Yang, 2001; Zendersky, 2005)
RCO_{it}	$\frac{RCO_{it}}{At_{it-1}}$	Positive (Lobo & Yang, 2001; Kanagaretnam, Lobo & Mathieu, 2003; Tomy, 2018)
$LWSEL_{it}$	Dummy variable: 1 for low and 0 for other cases	Positive

		(Ho & Saunders, 1981; Saunders & Schumacher, 2000)
ROA_{it}	$\frac{NR_{it}}{At_{it-1}}$	Positive (Kanagaretnam *et al.* , 2003)
SIZ_{it}	InTA	Positive (Curcio, Simone, Gallo, 2016)
PUB_{it}	Dummy variable: 1 for public banks and 0 for others	? (Santos, 2007)
BAS_{it}	Dummy variable: 1 for first quartile and 0 for others	Negative (2) / (Shrieve & Dahl, 2003; Santos, 2007)
PTC_{it}	Dummy variable: 1 for publicly traded institutions and 0 for others	Positive (Beatty, Ke & Petroni, 1999)

Text Table 1. Variables Used
Note: LLP_{it} is the expense associated with doubtful loan loss provisions; CO_{it} is the balance of portfolio credit operations; LLP_{it-1} is the balance for doubtful loan loss provisions at the beginning of the quarter, which reduces the assets; RCO_{it} is total revenues from credit operations; $LWSEL_{it}$ represents a low Selic rate; ROA_{it} indicates the return on assets; SIZ_{it} is the natural logarithm of total assets; PUB_{it} indicates 1 for banks with public participation and 0 for others; BAS_{it} indicates 1 for banks in the first quartile (which consists of the worst banks in terms of the Basel Indices) and 0 for others; PTC_{it} for banks with stocks that are traded publicly on the BM&F Bovespa stock market and 0 for others; At_{it-1} represents total assets at the beginning of the quarter.
Source: The authors (2017)

RESULTS

The information pertinent to the descriptive statistics and the normality tests of the variables used in the study is described in Table 1. Below are listed the averages, medians and standard deviations for the variables employed in Models 1 and 2.

Table 1:
Descriptive Statistics of the Variables Used

	Average	**Median**	**Standard Deviation**	**Jarque Bera Test**	
				Coefficient	**p-value**
Panel A – LLP					
LLP_{it}	-0.0067	-0.0033	0.0137	261022.3	0.0000
CO_{it}	0.3573	0.3598	0.2362	61.3262	0.0000
LLP_{it-1}	-0.0204	-0.0152	0.0249	67394.43	0.0000

RCO_{it}	0.0286	0.0206	0.0285	3086.853	0.0000
Panel B - DLLP					
SIZ_{it}	15.3254	15.2450	2.4264	47.7882	0.0000
ROA_{it}	0.0014	0.0034	0.0239	1061199	0.0000

Where: LLP_{it} is the expense for doubtful loan loss provisions; CO_{it} is the balance of the credit operation portfolio; LLP_{it-1} is the balance of the doubtful loan loss provisions at the beginning of the quarter, which reduces the assets; RCO_{it} represents the revenues from credit operations; ROA_{it} is the return on assets; SIZ_{it} is the natural logarithm of total assets

Source: This study

The Jarque-Bera test was used to test the normality of the variables employed in the statistical models. The results described in the table above indicate a rejection of the normality of the variables.

Nonetheless, to apply this panel data to Equation 2, the Ordinary Least Square model was first used (OLS). Based on these results, autocorrelation, homoscedasticity and multicollinearity tests were realized. The results are displayed in Table 2.

Table 2:
Results of the Estimates of Discretionary Accruals

	OLS	**OLS with robust standard error**	**Random Effect (robust standard errors)**
Constant	0.0002 (0.6896)	0.0002 (0.6675)	0.0009 (0.0386)
CO_{it}	0.0136 (0.0000)***	0.0136 (0.0006)***	0.0123 (0.0386)**
LLP_{it-1}	0.2361 (0.0000)***	0.236115 (0.0000)***	0.241815 (0.0001)***
RCO_{it}	-0.2366 (0.0000)***	-0.2366 (0.0000)***	-0.2422 (0.0001)***
VIF	≤ 1.9413	Chow Test 0.0000	
White Test	1044.9150 (0.0000)	Breusch-Pagan Test 0.0000	
Durbin-Watson Test	(0.0000)	Hausman Test 0.2625	
F-Sig	0.0000	0.0000	0.0000
R^2	0.4230	0.4230	0.3016
Adjusted R^2	0.4222	0.4222	0.3006
Observations	2077		

Where: CO_{it} is the balance of portfolio credit operations; LLP_{it-1} is the balance of doubtful loan loss provisions at the beginning of the quarter, which reduces assets; RCO_{it} represents revenues from credit operations;
Note: ***, ** and * levels of significance of 1%, 5% and 10%.

Source: Study results

Looking at the presented results, note that the Durbin-Watson test, as well as the White test presented p-values below 0.05, revealing problems of autocorrelation and heteroscedasticity. In light of this, it was necessary to use robust standard errors. In terms of diagnosing multicollinearity, the statistical measure VIF (*variance inflation factor*) presented values less than or equal to 1.9413 for the variables of Equation 1. Thus, we can see that there is an absence of problems in terms of multicollinearity, because the estimated values are less than 3 (Wooldridge, 2002).

After testing our statistical assumptions, the Chow, Breusch-Pagan and Hausman tests were performed to identify the most appropriate model for the proposed equation. The Chow test, which is designed to compare the OLS model with the fixed effect model, presented a low p-value, which rejects the null hypothesis that the OLS model is appropriate, and thus validates the alternative hypothesis of the existence of fixed effects. To test whether the OLS model is more appropriate than the random effect model, Breusch-Pagan's LM test was performed and it presented a low p-value, contradicting the null hypothesis that the OLS model is appropriate. Finally, the Hausman test compared the appropriateness of the random effects model with the fixed effects model, and the result of 0.2625 did not contradict the null hypothesis that the random effects model is consistent. That being so, the tests indicate that the random effects panel model is more consistent for Equation 1, which thus guided the realization of the regression, whose results are presented in Table 3.

The equation presents an adjusted R^2 of 0.3016. This determinant coefficient indicates that the variables employed in the model explain 30.416% of the doubtful loan loss provisions (LLPs). The results of this regression were used as a proxy for Equation 2 (DLLPs).

In terms of the explanatory variables employed in Model 1, it may be observed that CO_{it} presents a positive coefficient at a 1% level of significance. In keeping with the findings of Beaver and Engel (1996) and Kim and Kross (1998), this result which the banks use for the balance of their portfolio credit operations is a measurement of risk that is reflected in the LLP figures reported.

In relation to the explanatory variable LLP_{it-1}, a positive relationship was observed with a significance level of 1%. This is in keeping with the findings of Lobo and Yang (2001), which suggests that the balance of LLPs serves as a parameter for future reported LLP figures, as suggested by Wahlen (1994).

The RCO_{it}, which represents credit operations that originated during the quarter, results in a negative coefficient using this regression with a 1% level of significance. This result would contradict the initial expectation based on the findings of Beaver and Engel (1996) and Tomy (2018). A possible justification for this relationship is in the improvement of the credit portfolio of Brazilian financial institutions, which could be a consequence of economic factors or greater rigor in granting credit. The results of Beaver and Engel (1996), even though they found a positive relationship, observed that the increments of the credit portfolio presented a negative relationship. The authors suggest that this fact occurred due to several banking regulations which were implemented in the United States during the 1980s, which led to a greater concern about the quality of credit, and which resulted in greater scrutiny by bank examiners. In the case of Brazil, the reasons for this negative relationship are not within the scope of that work, which represents a limitation to this study.

Analysis of Earnings Management

After estimating the discretionary accruals through Equation 1, models from Equation 2 were used to test the hypothesis of the present study. To do this, the OLS regression model was used to test its statistical assumptions in terms of autocorrelation, heteroscedasticity and multicollinearity.

Table 3 displays the results generated by the tests realized on Equation 2. For the Durbin-Watson test, the result was a p-value less than 0.05, which validates the alternative hypothesis of autocorrelation, violating the statistical assumption. The White test in turn presented a value below 0.05, which thus rejected the null hypothesis of an absence of heteroscedasticity. In terms of the results generated by the VIF statistic, no multicollinearity problems were detected, because the values presented were less than 3 (Wooldridge, 2002). In light of the violation of the assumptions of autocorrelation and heteroscedasticity, robust standard errors were used, whose results are also presented in Table 3.

Table 3:
Results of the Regression – Second Stage

	Regression	**Regression with robust standard errors**	**Fixed Effect (robust standard errors)**
Constant	0.0152	0.0152	-0.0023
	(0.0000)***	(0.0000)***	(0.8094)
$LWSEL_{it}$	0.0008	0.0008	0.0007
	(0.0445)**	(0.0054)***	(0.0153)**
ROA_{it}	-0.1138	-0.1138	-0.0804
	(0.0000)***	(0.0016)***	(0.0288)**
SIZ_{it}	-0.0007	-0.0007	0.0004
	(0.0000)***	(0.0000)***	(0.4736)
PUB_{it}	0.0011	0.0011	
	(0.1075)	(0.3713)	
BAS_{it}	0.0016	0.0016	1.21e-05
	(0.0002)***	(0.0843)*	(0.9778)
PTC_{it}	0.0007	0.0007	-0.0007
	(0.1718)	(0.2797)	(0.3264)
VIF	≤ 1.4334	Chow Test	
		0.0000	
White Test	389.1220	Breusch-Pagan Test	
	(0.0000)	0.0000	
Durbin-Watson Test	(0.2519)	Hausman Test	
		0.0000	
F-Sig	0.0000	0.0000	0.0000
R^2	0.1339	0.1339	0.0618
Adjusted R^2	0.1314	0.1314	0.0596
Observations	2077		

$LWSEL_{it}$ represents a low Selic rate; ROA_{it} indicates the return on assets; SIZ_{it} is the natural logarithm of total assets; PUB_{it} indicates 1 for banks with public participation and 0 for others; BAS_{it} indicates 1 for banks in the first quartile (which registers the worst banks in terms of the Basel Indices) and 0 for others; and PTC_{it} for banks with stocks that are traded publicly on the BM&F Bovespa stock market and 0 for others.
Note: ***, ** and * levels of significance of 1%, 5% and 10%.

Source: Study data

After testing the statistical assumptions, the panel data was diagnosed to determine the correct type of balanced panel data regression. In terms of Chow's test, we may observe the statistical significance of the tested model (p-value < 0.05), which validates the alternative hypothesis of the existence of fixed effects. Breusch-Pagan's LM test also produced significant results which validates the alternative hypothesis of the existence of random effects. To compare the appropriateness of the random effect and fixed effect models, the Hausman test was performed, which generated significant results only for Equation 2, and did not reject the null hypothesis of the existence of fixed effects. Thus, based on the performed tests, the regression was conducted using fixed effect panel data.

The results of the regression of the equation are reported in Table 3, with the determinant coefficients (adjusted R^2) of the 2 Model explaining 5.96% of discretionary LLP accruals. It should be emphasized that its low explanatory power is a possible consequence of having a two-stage model, which as Zendersky remembers (2005), absorbs all of the heterogeneity of the transversal data. In Brazil, for example, Zendersky (2005) presented a model that explains 5% of the discretionary portion of the LLP. Meanwhile the model proposed by Dantas *et al.* (2012), explains 13.8% of the discretionary accruals of stock earnings. In the international sphere, in the studies of Beaver and Engel (1996), the equation used in the second stage to explain the discretionary portion of the LLP between the period from 1977 to 1984 resulted in a determinant coefficient of 11%. Thus, this suggests that the low determinant coefficient found in the study does not discredit the findings of the proposed model.

In terms of the results found in Equation 2, the most notable are the positive and statistically significant parameters (p-value<0.05) of the variable that represents a low Selic rate ($LWSEL_{it}$), which thus does not reject H0. These results are consistent with the initially defined expectation, which was traced in consonance with the allusions made based on the studies of Demirgüç-Kunt and Huizinga (1999), Ho and Saunders (1981), Saunders and Schumacher (2000) and Hannan and Berger (1991). It has therefore been verified that during times of economic stimulus, when there are low prime interest rates, financial institutions are increasing their use of discretionary LLP accruals, in order to reduce bank earnings during the period, saving them for future use.

It may further be observed, that in the model employed to explain discretionary LLP accruals (2), that the ROA_{it} variable presents a negative coefficient and a significant value (p-value<0.05). These findings suggest that banks raise the manageable portion of LLPs when they have lower earnings, which therefore does not confirm the hypothesis of income smoothing. Thus, the intentional reduction of these results, based on the statistical evidence, is not being realized to minimize the volatility of the results over the long term. The studies of Ahmed *et al.* (1999), also obtained similar parameters to these variables, which are not in line with the presented literature. The authors argue that the discrepancy in these findings is due to the difference in the sample period utilized. Wall and Koch (2000) in turn, agree with these justifications remembering that banks are subject to auditing procedures that can change the focus of managers during any given period. Nonetheless, the fact that income smoothing has not been found to occur does not contradict the hypothesis that banks manage earnings to “save” profits, because according to Martinez (2001) there is also the “*big bath”* explanation, whose applicability is beyond the scope of the present study.

CONCLUSION

This study seeks to analyze the relationship that exists between a low Selic rate and earnings management practiced by banks that operate in Brazil. This purpose was mainly motivated by historical data that demonstrates that when there have been lower Selic rates ever since this has become an indicator for economic policy. In addition, theoretical evidence argues that this macro-change can positively influence bank earnings. In light of this scenario, it has been suggested that managers use discretionary accounting to reduce earnings, saving them for later times of difficulty.

To achieve this objective, as well as test this study’s hypotheses, we have used a representative sample of Brazilian banks, classified as commercial banks, multiple banks with commercial operations, and credit unions that do not belong to conglomerates. The criterion used to select the sample was the availability of the information necessary to operationalize the econometric model during the period ranging from January 2010 to September 2017.

In terms of earnings management, we have opted to use the specific accruals model, electing to use reported LLPs to estimate the portion that can be managed by banks. These equations have been constructed based on theoretical fundamentals which justify the use of the explanatory variables. To verify the robustness of the model, we have employed tests designed to address statistical assumptions. Based on these criteria, we have determined that the estimates related to the discretionary portion of these figures can appropriately be used as proxies for earnings management.

Given the estimates of the discretionary portion of LLPs obtained in the first stage, we used these results in a second equation to test this study's hypothesis. We also used the ROA variable to verify the use of income smoothing practices by managers. The model also considers other explanatory variables which represent the size of the financial institution, whether it is publicly or privately controlled, Basel indicators, and whether they are publicly traded on the BM&F Bovespa stock market.

The reported earnings, when this last model is used, reveal that decreases in the prime interest rate are related to the utilization of the discretionary portion of LLPs. This evidence testifies to the importance of macroeconomic changes in the business climate, and includes elements that make it possible to infer whether exogenous factors can influence managers in terms of opportunistic behavior regarding accounting choices.

This statistical evidence suggests that bank managers are reducing earnings using discretionary elements of LLPs during economic times when Selic rates are low. This finding is consistent with our initial expectation which was based on allusions to studies that reported the influence of the prime interest rate on bank earnings (Ho & Saunders, 1981; Saunders & Schumacher, 2000; Hannan & Berger, 1991). In other words, considering that low interest rates are propitious for achieving greater earnings through financial intermediation, it was expected that managers would have a tendency to exhibit opportunistic behavior, increasing expenses to reduce reported quarterly gains to be later used in subsequent years. The results for this first stage offered aspects that confirmed this behavior.

However, this practice of income smoothing was not perceived statistically in the model that was used to explain discretionary LLP accruals. The negative relationship between the variable used to express the return on assets and earnings management suggest that banks are reducing their earnings even when this indicator is decreasing. These statistical parameters have already been perceived in the literature (Ahmed *et al*., 1999) and have been justified as a limitation of the sample period. In addition, this situation does not contradict the hypothesis of earnings management designed to "save" earnings.

Based on this evidence, this study contributes to the existing literature by offering findings regarding the use of discretionary accounting to perform earnings management. These findings confirm the need of normative elements that seek to restrain practices that seek to report opportunistic and therefore asymmetric information. In addition, in line with Accounting Choice Theory, these findings offer evidence of factors outside organizations that can influence manager decision making. Nonetheless, we would like to emphasize the relevance of policy decisions concerning low Selic rates, which in addition to influencing macroeconomic issues, have been statistically revealed to have the power to instigate opportunistic behavior on the part of business managers. That being so, we consider the evidence described throughout this study to be relevant to the existing literature.

In terms of this study's limitations, the period of analysis stands out, given that it only addresses low rates between 2010 and 2017 for the reasons mentioned above. In addition, the sample used is limited only to Brazilian commercial banks and does not encompass other economic activities. There are also limitations in terms of the econometric model utilized, which can present a non-discretionary portion of the total accumulations within the error term used to express discretionary accruals.

Finally, in terms of future research, we would recommend that other periods of low Selic rates be tested in order to provide new scientific evidence related to the influence of economic indicators on earnings management. In addition, other samples can be used to test earnings management in other economic sectors.

References

Afanasieff, T. S., Lhacer, P. M., & Nakane, M. I. The Determinants of Bank Interest Spread in Brazil. Working Paper Series, 2002.

Almeida, J. E., Lopes, A. B., & Corrar, L. J. Gerenciamento de resultados para sustentar a expectativa do mercado de capitais: impactos no índice market-to-book. Advances in Scientific and Applied Accounting, 2011.

Beatty, A., Ke, B., & Petroni, K. Differential earnings management to avoid earnings declines and losses across publicly and privately-held banks . Working Paper, Penn State University, 1999.

Beaver, W. H., Eger, C., Ryan, S., & Wolfson, M. Financial reporting and the structure of bank share prices. Journal of Accounting Research, 1989.

Brock, P. L., & Suarez, L. R. (2000). Understanding the behaviour of bank spreads in Latin America. Journal of Development Economics, 2000.

Baatour, Kais e Othman Hakim. “Legal origin, economic freedom and earnings management practices: MENA evidence”. Internationa Journal of Accounting, 2016.

Carvalho, F. J. Uma contribuição ao debate em torno da eficácia da política monetária e algumas implicações para o caso do Brasil. Revista de Economia Política, 2005.

Castro, P., & Mello, J. M. Há assimetria no repasse dos juros bancários de variações na taxa Selic? . Departamento de Economia PUC-Rio Working Paper, 2010 .

Claeys, S., & Vennet, R. V. Determinants of Bank Interest Margins in Central and Eastern Europe: A Comparison with the West. Economic Systems, 2008.

Curcio, Domenico; Simone, Antonio De; Gallo, Angela. Financial crisis and international supervision: New evidence on the discretionary use of loan loss profisions at Euro Area comercial banks. The British Accounting Review, 2016.

Dantas, José Alves, Otávio Ribeiro Medeiros, Fernando Caio Galdi, e Fábio Moraes Costa. “Gerenciamento de Resultados em Bancos com Uso de TVM: Validação de Modelo de Dois Estágios .” *Revista Contabilidade & Finanças* , 2012.

Demirgüç-Kunt, Asli, e Harry Huizinga. “Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence.” *World Bank Economic Review*, 1999.

Dick, Astrid. “ Banking Spreads in Central America: Evolution, Structure and Behavior.” *Development Discussion Papers Central America Project Series* , 1999.

Goulart, André Moura Cintra. “Gerenciamento de resultados contábeis em instituições nanceiras no Brasil.” *Tese de doutorado: Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, Brasil.* 2007.

Hannan, Timothy, e Berger Allen. “The Rigidity of Prices: Evidence from the Banking Industry.” *The American Economic Review*, 1991.

Ho, T., e A. Saunders. “The determinants of bank interest margins: theory and empirical evidence.” *Journal of Financial and Quantitative Analyses* , 1981.

Jiang, Guorong, Nancy Tang, Eve Law, e Angela Sze. “Determinants of Bank Profitability in Hong Kong.” *Hong Kong Monetary Authority Quarterly Bulletin*, 2003.

Kanagaretnam, Kiridaran, Gerald J., Lobo, e Mathieu, Robert. “Managerial Incentives for Income Smoothing through Bank Loan Loss Provisions.” *Review of Quantitative Finance and Accounting*, 2003.

Kim, Myung-Sun; Kross William. The impact of the 1989 change in bank capital standards on loan loss provisions and loan write-offs. Journal of Accounting and Economics, 1998.

Lobo, Gerald J., e Dong-Hoon Yang. “Bank Managers' Heterogeneous Decisions on Discretionary Loan Loss Provisions.” *Review of Quantitative Finance and Accounting*, 2001.

Martinez, Antonio Lopo. “Gerenciamento dos resultados contábeis: estudo empírico das companhias abertas brasileiras.” *Tese de doutorado: Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, Brasil.* 2001.

McNichols, Maureen, e Peter G. Wilson. “Research design issues in earnings management studies.” *Journal of Accounting and Public Policy*, 2000.

McNichols, Maureen, e Peter G. Wilson. “Evidence of Earnings Management from the Provisionfor Bad Debts.” *Journal of Accounting Research*, 1988.

Moyer, Susan E. “Capital adequacy ratio regulations and accounting choices in commercial banks.” *Journal of Accounting and Economics*, 1990.

Oreiro, José Luís da Costa, Luiz Fernando de Paula, Guilherme Jonas Costa da Silva, e Fábio Hideki Ono. “Determinantes macroeconômicos do spread bancário no Brasil: teoria e evidência recente.” *Economia Aplicada*, 2006.

Paulo, Edilson. “Manipulação das Informações Contábeis: Uma Análise Teórica e Empírica sobre os Modelos Operacionais de Detecção de Gerenciamento de Resultados.” *Tese de doutorado: Universidade de São Paulo, São Paulo, Brasil.* 2007.

Riahi-Belkaoui, Ahmed. “Is Earnings Management Affected by Human Development and Economic Freedom?” *Review of Accounting and Finance*, 2002.

Rodrigues, L. F. Evidência de gerenciamento de resultados em ofertas públicas de aquisição de ações de companhias abertas: análise empírica das operações realizadas no Brasil no período de 2007 a 2012. 2013. Dissertação (Mestrado em Ciências Contábeis) – Programa Multi-institucional e Inter-Regional de Pós-Graduação em Ciências Contábeis da Universidade de Brasília, da Universidade Federal da Paraíba e da Universidade Federal do Rio Grande do Norte, Brasília, 2013.

Santos, Eliana Calixto. “Capital Regulatório e Gerenciamento de Resultados nas Instituições Brasileiras que atual no Brasil.” *Dissertação de Mestrado: Fundação Instituto Capixaba de Pesquisa em Contabilidade, Economia e Finanças.* 2007.

Saunders, Anthony, e Liliana Schumacher. “ The determinants of bank interest rate margins: an international study.” *Journal of International Money and Finance*, 2000.

Shrieves, R. E., e D. Dahl. “Discretionary accounting and the behavior of japanese banks under nancial duress.” *Journal of Banking and Finance.* , 2003.

Tendeloo, Brenda Van, e Ann Vanstraelen. “Earnings Management and Audit Quality in Europe: Evidence from the Private Client Segment Market.” *European Accounting Review*, 2008.

Tomy, Rimmy E. Threat of Entry and the Use of Discretion in Banks’ Financial Reporting. Journal of Accounting and Economics, 2018

Wahlen, James M. The Nature of Information in Commercial Bank Loan Loss Disclosures. The Accounting Review, 1994.

Watts, Ross L. “Accounting Choice Theory and Market-based Research in Accounting .” *Rochester*, 1992.

Watts, Ross L., e Jerold L. Zimmerman. “Positive Accounting Theory.” *Prentice-Hall*, 1986.

Watts, Ross L. “Towards a positive theory of the determination of accounting standards.” *The Accounting Review*, 1978.

Watts, R. L., & Zimmerman, J. L. (1978). Towards a positive theory of the determination of accounting standards. The Accounting Review

Wooldridge, J. M. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press, 2002.

Zendersky, Humberto Carlos. “Gerenciamento de Resultados em Instituiçoes Financeiras no Brasil - 2000 a 2004.” *Dissertação de mestrado: Programa Multiinstitucional e Inter-Regional de Pós Graduação em Ciências Contábeis, Brasília, Brasil.* 2005.

DECISION SCIENCES INSTITUTE

Do Information Technology Firms Influence Unemployment?

Gary F. Templeton
Mississippi State University
Box 9581
Starkville, MS 39759
Email: gtempleton@msstate.edu
1(662) 325-1956

Shan Xiao
Mississippi State University
Box 9581
Starkville, MS 39759
Email: sx45@msstate.edu
1(662) 325-3928

ABSTRACT

This paper investigates the relationship between the prominence of information technology (IT) firms and US unemployment. We operationalize IT firm prominence as four industry sectors, each of which is measured in two ways relative to the overall economy: firm counts and market capitalizations. Applying covariance-based structural equation modeling to a dataset spanning from 1970Q4 to 2018Q1, we found small overall significant relationships between the relative importance of IT firms and national unemployment using two measurement types. The preponderance of our findings suggested that contrary to popular belief, IT prominence increased in the US economy as the unemployment rate dropped.

KEYWORDS: IT firms, unemployment, market capitalization, structural equation modeling

INTRODUCTION

Although the current unemployment rate remains extremely low in the US, recent articles in the popular press continue to alarm readers that information technology (IT) adoption eliminates jobs (Feldstein, 2018; Whitehouse & Rojanasakul, 2017; Williams, 2017; Gardels, 2018). Further, writers fear the effect could exacerbate macro-economic unemployment in the next economic downturn, which experts view as inevitable (Miles, 2019).

Technological unemployment, referring to the loss of jobs caused by technological innovation, concerns a large population of the U.S. Many workers are worried they may lose their jobs due to robotics or artificial intelligence as modern society progresses. A 2018 report by McKinsey Global Institute forecasts that one-third of American workers will be replaced by automation by 2030. The institute predicts that among 60% of occupations, at least 30% of work activities will be automated. While organizations embrace and enjoy advanced technology, IT generating firms continue to proliferate. IT firms are making use of diverse technologies to improve the organizational efficiency, competitive advantage, and public image (Dery et al., 2017) of non-IT firms. As an example, Apple demonstrates innovation using IT to grow significantly in the market

(Templeton et al., 2018). However, economic anxiety persists in the US despite the deployment of world-class technology throughout the country. This macro-economic phenomenon has not yet been examined by Information Systems (IS) researchers. Based on this discussion, we believe it is better to illuminate how IT firms and unemployment have historically fluctuated through several important economic cycles in the US. To address this gap, our first research question is:

Research Question 1: How does relative IT firm prominence and unemployment change during known economic periods of recession and expansion?

Hardware, services, trades, and telecommunications are four representative sectors of IT industry recent years. We are interested in these specific industries and their influence on unemployment. Using each subsector will produce specific findings that offer greater insight into the relationship between IT firm prominence and national unemployment. Hence, our second research question is:

Research Question 2: How do specific IT industries (hardware, services, trades, and telecommunications) influence unemployment?

We start examining the first question by graphing the interrelationships among two measures of IT firm prominence (counts and market capitalizations) and the unemployment rate. To answer the second question, we classify IT firms into four sectors – hardware, services, trades, and telecommunications, then construct two models for measuring the casual relationships. Specifically, we explore how market capitalizations generated by the four IT industries influence the unemployment rate and how IT firm counts in the four IT industries affect the unemployment rate. We believe the findings may shed new light on the impact of the relative size of IT firms on national unemployment from an economic perspective.

This paper begins by reviewing the theoretical background of IT firms and the influence of the relative prominence of IT firms on the unemployment rate. The methodology is presented in the next section, including the applied analysis technique, data, and analysis process. Finally, we present and discuss our results, followed by concluding remarks in the last section.

BACKGROUND

The economic impact of unemployment has been widely studied by researchers. Okun's law (Okun, 1962) points out that an increasing unemployment rate negatively influences a country's GDP (Prachowny, 1993) and labor (e.g., weekly hours). The unemployment rate is also a good single indicator of economic health. When the unemployed lose wages, it results in less contribution to the economy in terms of less purchasing power and taxable income. Thus, unemployment plays a significant role in overall economic health.

Technological Innovation and National Unemployment

The link between technology and unemployment has been discussed often in the last century. Despite concerns in the popular press, empirical studies are mixed concerning the relationship between technological change and unemployment. This relationship is complex, since technological progress is known to simultaneously eliminate and create jobs (Aghion & Howitt, 1998). Some empirical studies show that "positive technology shocks" improve employment

(Trehan, 2003, p. 13). Others contradict that assessment (Gali, 1999) and it is believed that technology negatively affects less educated employees the most (Rotman, 2017).

Modern economic theorists reason that the cyclical nature of unemployment is driven by technological obsolescence and innovation (Manuelli, 2000). In this vein, periods of unemployment are believed to be initiated by employers adjusting to technological obsolescence. As older technologies become obsolete, firms will either 1) update technology and retrain workers or 2) eliminate relevant jobs. In this way, technological progress causes the knowledge embodied in a job to become obsolete (Mortensen and Pissarides (1998). The decision to update or eliminate workers is effected by the cost of technology adoption. The Internet is a form of technological advancement that is known to greatly lower the cost of employer and worker job searches (Gomme, 1998). This makes migration in the labor market more efficient and timely, theoretically improving the labor situation long term.

IT Firms and National Unemployment

The significant role of IT in the global economy has been widely recognized (Porter, 2000). The unique capabilities of IT-generating firms to create value through digital disruption strategies (Templeton et al., 2018) is shown in the extraordinary market valuations of Alphabet, Microsoft, Amazon, Facebook, and others. The deployment of the products and services of these companies throughout the global marketplace also suggest their profound economic influences (Picardo, 2018). The value creating strategies most associated with IT firms are known to be R&D investment (Jha & Bose, 2016), outsourcing (Wickramasinghe & Jayaweera, 2011), mergers and acquisitions (Chang & Cho, 2017), and data analytics (Chen et al., 2015). Despite the success of these strategies, IT firms have been scarcely studied as objects of scientific experimentation. Consequently, there is very little available literature on the relationship between IT firm phenomena and national unemployment.

Given IT firms base their wealth on digital innovation, they are useful representations of the extent of technological revolution occurring in the national economy (Greenwood & Jovanovic, 1999; Hobijn & Jovanvic, 2001). The performance of IT firms are known to affect the U.S economy (Deutsch, 2018; Jorgenson et al, 2016), including the labor market. The literature supports two basic ways to assess the presence of IT firms in the US economy. First, firm asset values drive labor demand (Phelps, 1994). However, the relationship between firm value and unemployment is unclear. For example, Manuelli (2000) suggested firm technology improvements reduce market values, which in turn cause firms to reduce investment and hiring. Second, IT firm counts may drive unemployment. Manuelli (2000) explains that when firm market valuations are low, new firms are less willing to enter the market because existing technology is becoming obsolete. Thus, we expect both measurement types to be reasonable paths for assessing the IT firm-unemployment relationship.

IT Industries and National Unemployment

Our study splits the IT economy into four industry sectors. Figure 1 shows that the IT subsectors (hardware, services, trades, and telecommunications) defined in Templeton et al. (2018) are a substantial component of the U.S. economy. Despite their prominence, virtually no existing literature addresses these IT industries.

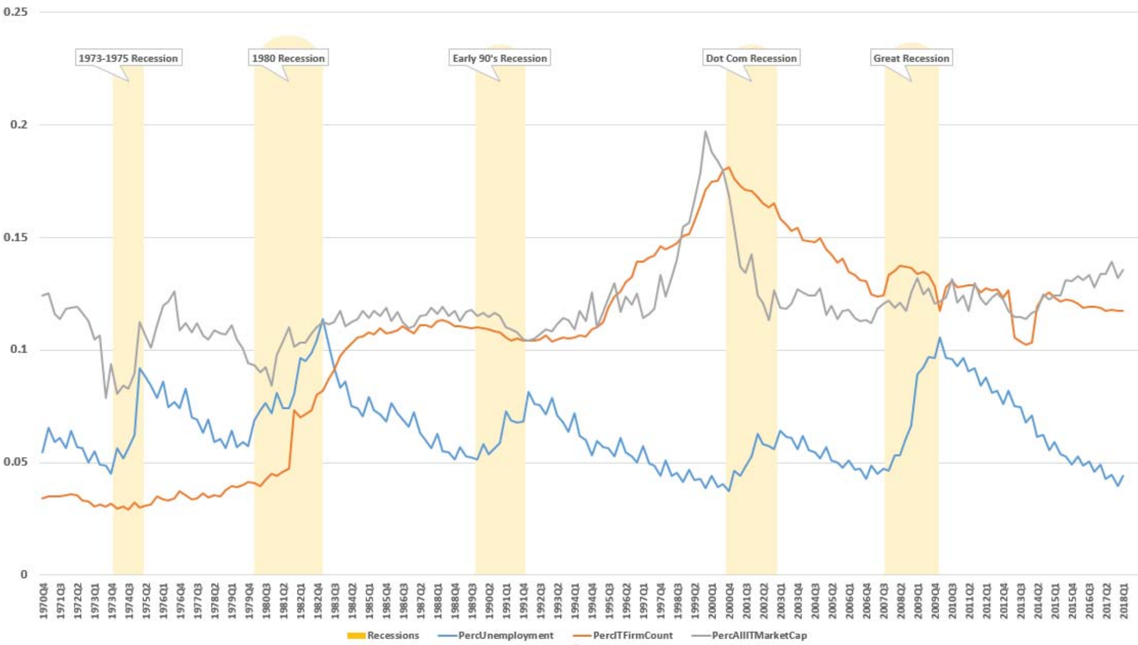
The hardware IT sector includes firms that manufacture computer, telephone, television, and other communications equipment. Discussions of automation and robotics are relevant to this subsector. The threat of process automation to employment has been a prominent social issue over the last half-century (Davenport & Kirby, 2015; Miller, 1964). According to studies, the linkn between IT hardware and unemployment is confounding and complex. Robotics are believed to increase organizational unemployment rates by replacing workers from a variety of occupations (Fortunati et al., 2015). Therefore, experts believe robotics contribute to labor uncertainty (Frick, 2014) and decrease unemployment (Sorells, 2018). Schwitter (1965) argued that computer hardware has had a minor effect on the unemployment rate. Aside from its labor consequences, automation also fulfills employees' needs in ways that may not be reflected in economic indicators (Diebold, 1953). Technological advancement is believed to cause a substitution of capital for labor (Falatoonzadeh & Safarzadeh, 2017) in ways that cannot readily be measured. IT services include programming services, computer processing, and developed systems. Writers and researchers are usually referring to this sector when discussing the effects of artificial intelligence (AI) on unemployment. For instance, Rotman (2017) warns that labor issues should be considered while enjoying the benefits brought by artificial intelligence. IT-enabled service innovations are known to save labor costs in organizations (Seifer, 1971). For example, technical advancement allows self-servicing such that customers make transactions online without assistance. Organizations are known to apply IT to reduce reliance on scarce and expensive skills (Miles, 1993).

IT trades include the wholesale and retail sales of digital devices and software. The relationship between IT trades and unemployment also remains obscure. While Borjas, Freeman, and Katz (1991) found a small negative relationship between trade and labor conditions at the national level, they ignored the role of computing. Results concerning computer equipment may be differentiated from other economic phenomena due to Moore's Law (Moore, 1965). While this uniqueness has been demonstrated in studies (Sachs & Shatz, 1994), nothing has been done to isolate country-specific computer trades and unemployment. Morin (2014) explains that as computers become cheaper than labor, they become more in demand. The computer displacement of routine work occurs rapidly as the economy enters recessions and these jobs are not replaced during expansions. However, the world has not experienced continuously increasing unemployment due to falling computing costs.

The telecommunications IT sector includes firms involved in radio, telephone, messaging, and other communications. The literature regarding this sector is mixed concerning its relationship with unemployment. Investment within the telecommunications sector presumably leads to economic growth (Zahra et al., 2008). However, Czernich (2014) could not confirm their results that the diffusion of broadband internet is associated with lower unemployment rates in Germany. Montgomery and Benbasat (1983) advocated using telecommunications to reduce white collar labor costs, although they did not test this association.

Practically all publicized concerns about the negative economic effects of IT on unemployment ignore surges in influence among IT firms. We could not find any article investigating the link between the relative prominence of IT firms in society and macro-economic unemployment. For these reasons, we empirically investigate this relationship in the US.

Figure 1: The Relative Prominence of IT firms and Unemployment (1970Q4 – 2018Q1)[1]



[1]Marked recessions in Figure 1 are based on prominently understood US economic recessions. We defined the timeframes for each, only for the purpose of charting, by estimated each recessionary period. The low points in unemployment marked the starts of recessions and the high points marked recessionary finishes.

METHODLOGY

Data

All IT firm data was collected from the *Compustat – Capital IQ* database component of the *Wharton Research Database Service (WRDS)*. To capture the largest and reliable sample frame, the study utilized a quarterly timeframe spanning quarter 4 of 1970 until quarter 1 of 2018. We used multiple measures of IT firms to gauge their prominence in the US economy: 1) aggregated market capitalization and 2) IT firm count. We calculated these values within each of four IT sectors: hardware, services, trades, and telecommunications. Within each sector-quarter, we divided each aggregated IT firm measure by the same for all firms to arrive at relative scores. This provided eight measurement options (relative count and capitalization measures for each of the four IT sectors) for assessing the relationship between IT economic prominence and national unemployment. We collected the US non-seasonally adjusted unemployment rate from the Federal Reserve Economic Database (FRED). We divided this value by 100 in order to standarize with other study measures. Table 1 displays the descriptive statistics for each study variable.

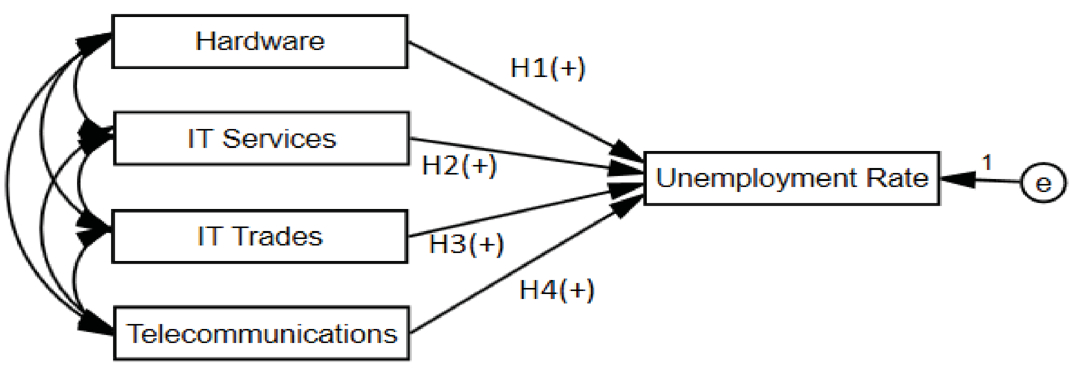
Table 1: Study Variables

Variable Name	**n**	**μ**	**Σ**
Unemployment Rate (%)	190	0.06	0.02
Relative Market Capitalization			
Hardware	190	0.03	0.00
Telecommunications	190	0.05	0.01
Trades	190	0.00	0.00
Services	190	0.05	0.01
Relative Firm Count			
Hardware	190	0.03	0.00
Telecommunications	190	0.02	0.01
Trades	190	0.00	0.00
Services	190	0.03	0.01

Procedures

We used Excel 2016, SPSS 25, and Amos 25 for all analyses. We used covariance-based structural equation modeling (SEM) to test hypotheses associated with RQ2 (concerned with the link between IT subsector prominence scores and unemployment). Figure 2 illustrates the path diagram used to test the relationships between each of the four industry sectors and unemployment. This model was tested using two bases for measurement: relative market capitalization and relative firm count.

Figure 2: Research Model



RESULTS

RQ1 is concerned with how economic recessions effected aggregated measures of IT firm prominence and unemployment. Figure 1 illustrates the relative stability of these IT firm economic measures, their behavior during and between five prominent recessions[1], and their relationship with national unemployment. For example, the graph indicates that the relative prominence of IT firms (using both measures) was at its peak in 2000, just prior to the Dot Com Recession. Further, it shows an unclear relationship between both IT firm prominence omnibus measures and unemployment. While unemployment rates are clearly cyclical over time, there does not appear to be a pattern for IT firm promenence in both boom and bust economic periods. This is true using both IT firm count and market capitalization data. This data justifies further empirical work on this evidence.

Pursuant to RQ2, we found that the relative aggregate market capitalizations of IT industry sectors explained 18.1% (Adjusted R^2) of the national unemployment rate. Further, we found support for three of the four hypotheses (see Table 2, Panel A). All tests produced negative critical ratios, indicating that unemployment decreases as the prominence of IT firms increases. This, and especially the consistent pattern in three significant sectors (hardware, trades, and services), suggests that IT firm health improves the labor situation at the national level. Further, these indicators are adequately independent influences on the overall economy. RQ2 results also indicate that IT firm counts explain 14.2% of the national unemployment rate. The analysis elicited support for two of four hypotheses (see Table 2, Panel B). Consistent with results pertaining to H4a, findings for H4b suggested a negative relationship between IT services prominence and unemployment. However, contrary to all other significant findings, the path for H1b showed a positive relationship between IT hardware firm prominence and unemployment. Thus, this isolated finding generally supported the idea that IT causes unemployment.

Table 2: Market Capitalization in Four IT Industries and Unemployment Rate

Panel A: Using IT Firm Market Capitalization

Hyp	Test	Estimate	S.E.	C.R.	P
H1a	Hardware Market Cap→ Unemployment	-.481	.155	-3.102	.002
H2a	Telecom Market Cap→ Unemployment	-.087	.098	-.889	.374
H3a	IT Trades Market Cap→ Unemployment	-16.922	4.972	-3.403	***
H4a	IT Services Market Cap→ Unemployment	-.232	.096	-2.413	.016

Panel B: Using IT Firm Counts

Hyp	Test	Estimate	S.E.	C.R.	P
H1b	Hardware Firm Count → Unemployment	.833	.182	4.575	***
H2b	Telecom Firm Count → Unemployment	-.173	.096	-1.795	.073
H3b	IT Trades Firm Count → Unemployment	1.820	1.498	1.214	.225
H4b	IT Services Firm Count → Unemployment	-.906	.186	-4.870	***

DISCUSSION

Results concerning RQ1, concerned with observing the interplay between study variables during economic recessions and expansions, reveal insights about how IT firms interact with the overall economy. Importantly, it reveals that IT firm prominence is a stable and sensible economic indicator. The phenomenon, introduce in this study, is reliable over a timeframe spanning several decades. Further, we show that unemployment is a demonstrably cyclical economic indicator that correlates well with known economic peaks and troughs.

Pursuant to RQ2, which involves four hypotheses concerning relationships between IT firm prominence and unemployment, we found mostly consistent results. One basic interpretation from this result is that IT firms influence the U.S economy in diverse ways that affect overall employment opportunities (Jorgenson et al, 2016). Our results provided evidence that illuminates how four IT sectors uniquely effected unemployment. As research using novel economic concepts and measures, the findings promote the inducement of new economic theory applicable to IT firms.

Proponents of the belief that IT causes unemployment may be bolstered by our finding that the prominence of the hardware sector (assessed using firm counts) positively influences national unemployment (H1b). That is, as the sector grows in firms, the labor situation is harmed and vice versa. This could very well be due to robotics, which are widely feared in the popular press for this reason. On the other hand, this test resulted in a negative significant relationship when using market capitalization data (H1a). However, popularly held notions about the overall effects of IT on unemployment were challenged by the preponderance of our findings. As we discussed earlier, robotics or automation may be one of the possible reasons. This finding suggests it may be beneficial to train employees to have updated skills (Davenport & Kirby, 2015).

Interestingly, the prominence of the telecommunications sector was insignificant as a predictor of unemployment. This is consistent with prior questions about whether efficiency in telecommunications technology affected unemployment (Runhaar & Lafferty, 2009). By contrast, the number of trades firms, unlike the market capitalization, does not influence the unemployment rate. Therefore, the economic aspects of IT trades deployment show the mixed results in this paper. Policymakers can take advantage of our findings to make better decisions to steer the economy and counter unemployment. Last, the count of IT services firms helps with the unemployment rate, aligning with the result of its market capitalization.

CONCLUSIONS

This paper presents a study assessing the relationship between IT and unemployment at the national level. Using structural equation modeling (SEM) path analysis applied to 47 years, we found support for five out of eight hypotheses tested. However, four of our significant hypotheses diverged from a long history of publicized fears claiming IT causes unemployment. Given the temporal and industrial spans of this study, it is useful to revisit one of the direst warnings about communication systems in the last century in *Harvard Business Review*:

> *"It is perfectly clear that this will produce an unemployment situation, in comparison with which the present recession and even the depression of the thirties will seem a pleasant joke. This depression will ruin many industries -- possibly even the factories that have taken advantage of the new potentialities" (Wiener, 1954).*

Broad speculations such as these have been debunked by the last half century of experience, in addition to our findings. Our study reveals that it is more realistic to say there is a relationship between the influence of IT in society and unemployment, but the relationship is not profound, and is complex. In fact, most of our significant findings indicate that when IT firms are most influential, the labor market improves. Both our data (see Figure 1) and our findings reveal that the public should largely absolve IT firms and their customers further blame for job losses.

A few limitations should be recognized when interpreting the research findings. First, we used two variables to assess the relative prominence of IT firms and a single variable to assess the national unemployment rate. More perspectives on both measures would be more convincing while testing our hypotheses. For Instance, the US economy is very complex and is monitored using hundreds of diverse metrics. Second, the prominence of IT firms may very well effect other economic factors, such as wages (Helpman et al., 2010) and the growth of gross domestic product. Control variables may also be used to achieve more rigorous results. For these reasons, future research may incorporate additional indicators to extend our tested models. The association between the economic prominence of telecommunications on productivity would make interesting future research. In light of our finding that hardware firm count may potentially increase national unemployment, we suggest researchers uncover and use more specific data about this relationship.

REFERENCES

Borjas, G. J., Freeman, R. B., & Katz, L. F. (1991). On the labor market effects of immigration and trade (No. w3761). *National Bureau of Economic Research*.

Chang, Y. B., & Cho, W. (2017). The Risk Implications of Mergers and Acquisitions with Information Technology Firms. *Journal of Management Information Systems*, 34(1), 232-267.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, 32(4), 4-39.

Czernich, N. (2014). Does Broadband Internet Reduce the Unemployment Rate? Evidence for Germany. *Information Economics and Policy*, 29, 32–45.

Davenport, T. H., & Kirby, J. (2015). Beyond Automation. *Harvard Business Review*, 93(6), 58-65.

Dery, K., Sebastian, I. M., & van der Meulen, N. (2017). The Digital Workplace is Key to Digital Innovation. *MIS Quarterly Executive*, 16(2), pp. 135-152.

Deutsch, A.L. (2018).The 5 Industries Driving the U.S Economy. *Investopedia*. Available at: https://www.investopedia.com/articles/investing/042915/5-industries-driving-us-economy.asp [accessed by 11/27/18] (IT drives US Economy)

Diebold, J. (1953). Automation-the New Technology. *Harvard Business Review*, 31(6), 63-71.

Falatoonzadeh, H., & Safarzadeh, M. (2017). Has Automation Changed the Response of Unemployment Rate to GDP Growth? *Journal of Emerging Issues in Economics, Finance and Banking*, 6(2), 2260–2273.

Feldstein M. (2018). The Fed Can't Save Jobs from AI and Robots. *The Wall Street Journal*. Available at: https://www.wsj.com/articles/the-fed-cant-save-jobs-from-aiand-robots-1528652059 [accessed by 11/27/18]

Fortunati, L., Esposito, A., & Lugano, G. (2015). Introduction to the Special Issue "Beyond Industrial Robotics: Social Robots Entering Public and Domestic Spheres," *The Information Society: An International Journal*, 31(3), 229-236.

Frick, W. (2014). Experts Have No Idea If Robots Will Steal Your Job. *Harvard Business Review*, p2-5. 4p.

Gali, J. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review* 89(1) (March) pp. 249-271.

Gardels, N. (2018). When Robots Take Our Jobs. The Washington Post. Available at: https://www.washingtonpost.com/news/theworldpost/wp/2018/02/02/automation-jobs/?utm_term=.f5965e673f7c [accessed by 11/27/18]

Gomme, P. 1998. "What Labor Market Theory Tells Us about the 'New Economy.'" *FRB Cleveland Economic Review QIII*, pp. 16-24.

Greenwood, J., & Jovanovic, B. (1999). The Information Technology Revolution and the Stock Market. *American Economic Review*, 89(2), 116-122.

Helpman, E., Itskhoki, O., & Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4), 1239-1283.

Hobijn, B., & Jovanovic, B. (2001). The Information Technology Revolution and the Stock Market: Evidence. *American Economic Review*, 91(5), 1203-1220.

Jha, A. K., & Bose, I. (2016). Innovation Research in Information Systems: A Commentary on Contemporary Trends and Issues. *Information and Management*, 53(3), 297-306.

Jorgenson, D., Gollop, F. M., & Fraumeni, B. (2016). *Productivity and US Economic Growth* (Vol. 169). Elsevier.

Manuelli, R.E. 2000. "Technological Change, the Labor Market, and the Stock Market." *NBER Working Paper 8022* (November).

Miles, I. (1993). Services in the new industrial economy. *Futures*, 25(6), 653-672.

Miles, T. (2019). Keep Humans in Good Jobs in the Robot Era, ILO Commissions Urges.

https://www.reuters.com/article/us-ilo-robots/keep-humans-in-good-jobs-in-the-robot-era-ilo-commission-urges-idUSKCN1PG1QC [accessed by 1/31/19]

Miller, J. J. (1964). Automation, Job Creation, and Unemployment. *Academy of Management Journal*, 7(4), 300-307.

Montgomery, I., & Benbasat, I. (1983). Cost/benefit Analysis of Computer Based Message Systems. *MIS Quarterly*, 7(1), 1-14.

Moore, G.E. (1965). Cramming More Components onto Integrated Circuits. Electronics, pp. 114-117, April 19, 1965.

Morin, M. (2014). Computer Adoption and the Changing Labor Market. Unpublished Manuscript (Job Market Paper), Columbia University.

Mortensen, D.T. & Pissarides, C.A. 1998. "Technological Progress, Job Creation, and Job Destruction." *Review of Economic Dynamics* 1(4) pp. 733-753.

Okun, A. (1962). Potential GNP: Its Measurement and Significance. American Statistical Association, *Proceedings of the Business and Economic Statistics Section*, 98-104.

Phelps, Edmund S. (1994). *Structural Slumps: The Modern-Equilibrium Theory of Unemployment, Interest, and Assets*, Cambridge, Mass.: Harvard University Press.

Picardo, E. (2018). Eight of the World's Top Companies Are American. *Investopedia*. Available at: https://www.investopedia.com/articles/active-trading/111115/why-all-worlds-top-10-companies-are-american.asp [accessed by 11/27/18] (Top IT firm's market capitalization)

Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly*, 14(1), 15-34.

Prachowny, M. F. (1993). Okun's Law: Theoretical Foundations and Revised Estimates. *The Review Of Economics and Statistics*, 331-336.

Rotman, D. (2017). The Relentless Pace of Automation. Retrieved September, 14, 2017.

Runhaar, H., & Lafferty, H. (2009). Governing Corporate Social Responsibility: An Assessment of the Contribution of The UN Global Compact to CSR Strategies in The Telecommunications Industry. *Journal of Business Ethics*, 84(4), 479-495.

Sachs, J.D., Shatz, H.J., Deardorff, A., & Hall, R.E. (1994). Trade and jobs in US manufacturing. *Brookings Papers on Economic Activity*, 1994(1), 1-84.

Schwitter, J.P. (1965). Computer Effect upon Managerial Jobs. *Academy of Management Journal*, 8(3), 233-236.

Seifer, D.M. (1971). The Services Industry: Automation, Minimum Wages, Unemployment. *Ohio State University Bulletin of Business Research*, 46(8), 4.

Sorells, B. (2018). Will Robotization Really Cause Technological Unemployment? The Rate and Extent of Potential Job Displacement Caused by Workplace Automation. *Psychosociological Issues in Human Resource Management*, 6(2), 68–73. https://doi.org/10.22381/PIHRM6220186

Templeton, G.F., Pace, B., French, A., Petter, S. & Larsen, K.R. 2018. "IT Firms: Creating Value through Digital Disruption," *The Americas Conference on Information Systems (AMCIS)*, August 16-18, New Orleans, LA.

Trehan, B. (2003). Productivity Shocks and the Unemployment Rate. *Federal Reserve Bank of San Francisco Economic Review*, 13–27.

Wickramasinghe, V., & Jayaweera, M. (2011). Career Management Strategies Among Information Technology Professionals in Offshore Outsourced IT Firms in Sri Lanka. *Journal of Management Development*, 30(9), 914-926.

Wiener, N. 1954. The Human Use of Human Beings (Garden City, New York: Doubleday), p. 162.

Williams, A. (2017). Will Robots Take Our Children's Jobs? *The New York Times*. Available at: https://www.nytimes.com/2017/12/11/style/robots-jobs-children.html [accessed by 11/27/18]

Whitehouse, M., & Rojanasakul, M. (2017). Find Out If Your Job Will Be Automated. *Bloomberg*. Available at: https://www.bloomberg.com/graphics/2017-job-risk [accessed by 11/27/18]

Zahra, K., Azim, P., & Mahmood, A. (2008). Telecommunication Infrastructure Development and Economic Growth: A Panel Data Approach. *The Pakistan Development Review*, 711-72.

DECISION SCIENCES INSTITUTE

Working capital variability and financial spillover effects in supply chains: An explorative network analysis

Philipp Wetzel
University of St. Gallen (HSG), Institute of Supply Chain Management
Email: philipp.wetzel@unisg.ch

Erik Hofmann
University of St. Gallen (HSG), Institute of Supply Chain Management
Email: erik.hofmann@unisg.ch

ABSTRACT

The purpose of this paper is to analyze how working capital is allocated in supply chain networks. It investigates working capital fluctuations along supply chain networks. Inter-organizational network data was collected by snowball sampling in Bloomberg Terminal's database. The results suggest the existence of a financial spillover effect in the supply chain network and a mechanism that transmits working capital burden from customers to suppliers. Interestingly, the financial spillover effect is "reset" in the midstream section of supply chain networks where cash-rich companies provide liquidity to its supply chain partners. Similarly, inter-firm collaboration through the adoption of supply chain finance (SCF) solutions may reduce the cascading spillover reactions – the financial bullwhip effect – and stabilize a supply chain. network.

KEYWORDS: Working capital management, Financial spillover effects, Network theory, Snowball sampling, Supply chain finance

INTRODUCTION

Working capital is a measure of liquidity as it involves, by definition, all short-term items on the balance sheet. Consequently, net working capital might be interpreted as the liquidity that ensures continuity of a firms operations. Optimizing working capital has imminent benefits: First, it lowers capital requirements and releases liquidity (Deloof, 2003). Second, it reduces cost of capital (Knauer & Wöhrmann, 2013; Wöhrmann, 2012). Third, good working capital management (WCM) reduces financial and operational risks as a high proportion of business failures are due to poor decisions concerning the working capital of the firm (Smith, 1980). Working capital therefore is not only important from a finance perspective, but also from a logistics management perspective (Farris & Hutchinson, 2003). With respect to the latter, working capital is crucial in measuring supply chain management improvements and process efficiency (Farris & Hutchison, 2002). Companies usually do not improve liquidity management before reaching crisis condition or the verge of bankruptcy (Cole, 1991; Enqvist et al., 2014; Wöhrmann, 2012). This may have devastating consequences for an entire supply chain, because according to Wagner, Bode and Koziol (2009) default dependencies among suppliers exist. The difficulty of managing working capital stems from the inherent risk-return trade-off. "[…] too much investments in this area ties up liquidity that could be used more efficiently elsewhere" (Ek & Guerin, 2011). Conversely, too little investments in working capital might cause stock-outs and thus may alienate customers and suppliers. Furthermore, the traditional firm-individual approach to working capital management falls short in recognizing supply-chain

related inter-firm dependencies (Hofmann & Kotzab, 2010). It is decisive for decision makers to consider the following two lines of argumentation. First, companies no longer compete as stand-alone entities, but rather as integrated supply chains and second, managing working capital detached from a firms supply chain partners might no longer be the most efficient solution (Daugherty et al., 2006; Christopher, 2005). Supporting these statements, Protopappa-Sieke and Seifert (2017) and Viskari and Kärri (2012) both found significant benefits of collaborative WCM within the supply chain. Grosse-Ruyken, Wagner and Jonke (2011) suggested that shortening CCC in the entire supply chain increases the performance of a focal firm. In addition, Hofmann and Kotzab (2010) and Randall and Farris (2009) show optimization approaches for working capital management in a network context.
However, existing research is insufficient due to various reasons. First, the potential of network thinking similar to the interpretation in lean material supply chain management or JIT has not yet been quantitatively evaluated from a working capital perspective. In particular, the knowledge about financial spillover effects between companies in supply chain networks is very limited. Second, past studies in general conventionalized supply chains as linear systems of sequential relationships. However, a company's financial working capital relationships take place in an extended supply chain network beyond immediate dyadic relationships, which a linear model fails to capture. Third, current literature about working capital investigates companies as stand-alone entities. Conceptual studies have pointed towards a more holistic approach in WCM. However, only few quantitative studies exist that take into consideration that working capital is not only pivotal for the efficient management of an individual company but also greatly impact its extended supply chain network. The present paper addresses aforementioned gaps by analyzing working capital from a supply-chain-oriented perspective. In particular, the following overall research questions are examined:

RQ1: How is working capital allocated in supply chain networks?
RQ2: How does working capital fluctuate along supply chain networks?

This paper addresses the shortcomings in previous research attempts as follows: First, the paper at hand chooses a novel approach by analyzing working capital in supply chain networks from a social network perspective. It therefore bridges several literature streams from the working capital, supply chain/logistics and social network theory literature. Second, it established an understanding of how the presence of liquid companies with access to outside finance may break the cascading effect of financial spillover effects in supply chain networks (Serrano et al., 2018; Gropp, 2013). Third, it contributes to financial literature that deals with the effects of supplier-customer interactions by exploring supply chains in terms of financial contagions. Fourth, it provides practitioners with valuable hints on how supply chain networks can be stabilized to the benefit of all involved players through better working capital coordination (e.g. with the help of supply chain finance solutions). It extends current research and pushes towards a more holistic perspective on working capital management
The remainder of the paper is organized as follows. First, the theoretical background and hypotheses are presented. Second, the methodology and data analysis approach of the explorative supply chain network analysis are outlined. This includes the method for approximating supply chain network tiers and the statistical methods used to derive conclusions from observations in the dataset. Third, the results of the different analyses are presented. Fourth, the results of the data analysis are discussed and both theoretical and practical implications are provided.

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

A multitude of studies have analyzed working capital levels across numerous settings and varying scopes of analysis. In particular, differences amongst industries and geographies have been investigated comprehensively. For example, Hawawini et al. (1986) investigated whether the working capital requirement to sales ratio (i.e., CCC) is different among industries. The results indicate a significant industry effect on firm's investment in working capital over the examined periods covered. They argue that this industry effect exists partly due to similar technologies, processes and managerial efficiency. Similarly, Filbeck and Krueger (2005) found significant differences between industries regarding the working capital measures. In line with these observations, Weinraub and Visscher (1998) found relative differences in working capital policies between industries. The latter further propose that these differences persist over an extended period. Similarly, Ng et al. (1999) found that while credit terms vary widely across industries, only little variation within industries exists.
These studies are of great value because they improve our understanding about differences in working capital levels across distinct settings and therefore varying explanatory variables, in particular industry and countries. However, from a modelling perspective they are limited as they are not conducted in the setting of multiple supply chain relationships. Supply management research have used the complex system perspective as a theoretical basis to describe supply networks (e.g., Choi, Dooley, & Rungtusanatham, 2001; Kim, 2009; Pathak, Day, Nair, Sawaya, & Kristal, 2007; Surana, Kumara, Greaves, & Raghavan, 2005). This supply chain-related studies used both simulation models (e.g., Kim, 2009; Pathak et al., 2007) and real-world supply networks (e.g., Choi & Hong, 2002; Jarillo & Stevenson, 1991). Choi and Hong (2002), for example, study a sample of three complete networks from the automotive industry to conceptually analyze the structure of supply networks in this industry. Losbichler and Rothböck (2008) conducted a working capital study on 6,925 European companies over the period 1995–2004. By applying a supply chain network perspective on working capital they found that some industries were able to reduce CCC by up to 31.6 days, whereas in others the CCC increased by up to 59.9 days. In particular, car manufacturers were able to improve CCC on the expense of other companies in industries such as logistic providers and metal goods manufacturers that were faced with an increase in CCC. Another example is the study of Lorentz et al. (2016), who examined the effect of the business cycle on net trade credit including its components in different supply chain network tiers. They found that net trade credit decreases in further upstream tiers. Additionally, they found that period-to-period changes in accounts receivable and accounts payable are greater further up the supply chain. A similar mechanism was found by Chen (2012) who analyzed payment terms of project contractors, project owners, specialist contractors and suppliers. Against the background of these results our first hypothesis, addressing the level of working capital in supply chain networks, reads as follows:

H1: *Average CCC levels increase in companies located further upstream a supply chain network.*

While working capital-related papers using network theory are scarce, there exist various studies analyzing working capital-related spillover effects in customer-supplier relations. Regarding the inventory management component of WCM the most important spillover effect is known as the bullwhip effect (Chen & Lee, 2017). The bullwhip effect describes how the distortion of demand information in a supply chain amplifies demand variance as it moves from consumption point up the supply chain to layers of suppliers. The phenomenon arises when supply chain members place orders with a large variance to actual sales (demand distortion). The demand variation amplifies as it propagates to further upstream companies (Lee, Padmanabhan, & Whang, 1997). Several studies analyzed these spillover effects, highlighting that a bullwhip effect exists in a supply chain (e.g., Chen, Liao, & Kuo, 2013; Lee et al., 1997;

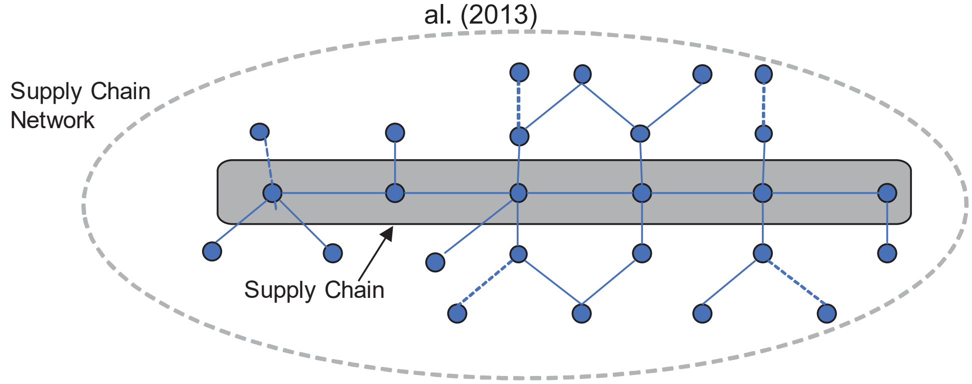
Lee, So, & Tang, 2000; Osadchiy, Gaur, & Seshadri, 2011). Only a few considered a supply network (e.g., Mizgier, Wagner, & Holyst, 2012; Ouyang & Li, 2010; Wangphanich, Kara, & Kayis, 2010).
Regarding the financial components of WCM a similar phenomenon is known as "financial contagion". It refers "to the increased likelihood of a firm defaulting to its suppliers as a result of its customers' defaults on trade credit" (Serrano, Olivio, & Kraiselburd, 2018). This suggests "not only that payments to suppliers are subject to variability, but also that variability is somehow transmitted upstream". Studies were conducted to analyze how this risk spreads from a single company to an industry. Serrano et al. (2018) showed that for strongly coupled relationships, the decision of a company not only affects the firm but also its suppliers. They confirm that large, liquid firms with access to outside finance ("deep pockets") have a reducing effect on the propagation of risk in a supply chain (Serrano et al., 2018). Similarly, Boissay and Gropp's (2013) results indicated that a credit-constrained firm facing a liquidity shock partly absorbs this shock and partly passes it on to suppliers. The liquidity shocks are transmitted through trade credit chains until a "deep pocket" fully absorbs the shock. Hence, firms with access to outside finance provide liquidity to the supply chain through maturity extensions on trade credit and by paying trade debits on time. Trade credit default chains might be seen as a mechanism that reallocates liquidity from deep pockets to small, credit-constrained firms. Boissay and Gropp (2013) interpret this mechanism as liquidity insurance through trade credit which is supplied by all types of companies to their customers. Also, small and credit-constrained firms can afford insuring their customer, because they are themselves insured by their suppliers. However, only deep pockets are able to inject fresh liquidity into the supply chain. Consequently, by extending maturities on trade credit supplier do not only "relax the financial constraints faced by their customers but also ultimately their customers' customers and other firms they may not have direct business relationships with". Chen et al. (2013) explore internal liquidity risk and financial bullwhip effects on corporate bond yield spreads along supply chain counterparties. Tangsucheeva and Prabhu (2013) introduce the financial (cash flow) bullwhip measured by the ratio of CCC variance over demand variance. Using simulation, they find "that increasing variability increases inventory and cash flow bullwhip along with lead time". Since the amplified order is higher than the actual demand, products remain in inventory and thereby cause higher inventory cost. Additionally, the company will also encounter higher opportunity- and financial cost as well as elevated levels of working capital. Itzkowitz (2013) showed that a supplier's cash holding increases as the importance of the relationship increases. This is argued to be a precaution against the additional operating risk by having an important customer. Kale and Shahrur (2007) show that lower debt level induces supplier and customer to make relationship-specific investments. Further, they indicate that companies have higher debt levels when customer or suppliers have higher bargaining power (bargaining role of debt). In line with these findings, Chu (2012) find decreasing firm leverage as the degree of competition between suppliers decreases. Oliveira, Kadapakkam, and Beyhaghi (2017) find that suppliers to distressed customers "increase their financing leverage to fortify their bargaining power". Against this background the second hypothesis, addressing the working capital variability and potential financial spillover effects in supply chain networks, reads as follows:

H2: *CCC variability increase in companies located further upstream a supply chain network.*

METHODOLOGY

Applying a supply chain network model, it is important to be clear about the distinction between a "supply network" and a "supply chain". A supply chain in general describes the extended network of dyadic exchange relationships that must exist for the creation of a product or service that is supplied to a final customer (Cox, Sanderson, & Watson, 2001). Hence, a supply chain is traditionally captured by a dyadic model that arranges suppliers and customers in a linear fashion. However, a dyadic perspective is limited in order to understand a company's working capital financing relationships as these depend on an extended network beyond immediate dyadic relationships. Kim et al. (2011) define a supply network as a firm's supply network that consists of ties to its immediate suppliers and customers, and ties between them and their immediate suppliers and customers, and so on. In this paper this definition is followed and illustrated in Figure 1.

Figure 1: Supply chain and supply network: a diagrammatical representation; Source: Rogers et al. (2013)



Data Source

All data used for this paper were downloaded from the Bloomberg Terminal database in November 2018. Bloomberg provides various information on supplier-customer relationships. Each relationship is categorized by cost type for the customer: cost of goods sold (COGS), capital expenditure (CAPEX), research and development (R&D) and selling, general and administrative expenses (SG&A). Bloomberg quantifies the relationship value if possible. For these "quantified" relationships, Bloomberg specifies percent revenue and percent cost. Revenue percent represents what the central company gets from is customer and cost percent represents what gets incurred by this customer.

Sampling Procedure and Data Collection

The supplier-customer relationship data were collected through snowball sampling. This is usually the preferred method for sampling network structures (Robins, 2015). Vogt and Johnson (2011) define snowball sampling as a technique for finding research subjects. One subject (ego) gives the researcher the name of another subject (alter), who in turn provides the name of a third, and so on. The ego initiating the snowball sample is called "seed". The snowball sampling approach is particularly useful for analyzing financial spillover effects as it can provide a better understanding about a topic for which little information is available (Griffiths et al., 1993). Further, large sample sizes and replication of results are likely to strengthen the level of generalization (Griffiths et al., 1993). According to Robins (2015) recent statistical models for social networks have demonstrated that estimated effects from a snowball sample can be consistent with the entire network structure. For the analysis in this paper the manufacturing company Hewlett Packard as part of the high-tech industry was chosen as the seed company. As Bloomberg's early data collection on supply chain information focused on the high-tech sector it is expected that this sector has a better supply chain coverage compared to other

sectors. Further, it deemed important to choose a company with (at least) 20 suppliers and 20 customers in the Bloomberg database. Fewer companies in the first wave would restrict the sample size of the next waves and render it impossible to come up with an extensive snowball sample.
After choosing HPE US Equity as the seed, four waves were sampled. In each wave a maximum of 20 suppliers and 20 customers were sampled per company of the previous wave (ego). This theoretically results in of maximum 2,560,000 (=1,280,000*2) alters in the last wave. However, due to lack of data (non-reporting, private company, government etc.), concentrated markets, loops or “beginning/end” of the supply chain the snowball sampling provided a total of 1,386,652 (=485,029+901,623) alters in the last wave. An overview of the sampled raw data is given in Table 1.

Table 1: Raw network sample; number of egos and alters per wave (including duplicates)

	Theoretically (per side)		**Downstream: effective**		**Upstream: effective**	
Wave	**Egos**	**Alters/no. dyads**	**Egos**	**Alters/no. dyads**	**Egos**	**Alters/no. dyads**
1	(HPE US Equity) 1	20	1	20	1	20
2	20	800	20	701	20	773
3	800	32,000	701	18,610	773	26,937
4	32,000	1,280,000	18,610	485,029	26,937	901,623

To clean the raw data, the following actions were conducted:

(1) removal of duplicate relationships;
(2) exclusion of all cost types other than COGS;
(3) removal of percent cost relation;
(4) removal of unquantified relationships;
(5) removal of isolated companies.

During the data cleansing all relationships with cost types other than COGS were excluded. This is motivated by the focus of the present paper on working capital relationships. In line with Itzkowitz (2013), this paper uses percent revenue as the measure for relationship importance. Revenue percentages are from a supplier perspective and represent material flow in a network (from node: supplier, to node: customer). Furthermore, only quantified relationships were considered in order to assign an importance to supplier-customer links. Lastly, companies that were no longer connected to the main cluster due to the data cleansing were removed. This process lead to a data set of 5,853 companies connected by 29,253 directed nodes. After collecting the network data, relevant financial working capital data for the represented companies were collected. In particular, data on accounts payable, accounts receivable, inventory, sales, sector (2-digit), industry group (4-digit), industry (6-digit), sub-industry (8-digit), Standard & Poor’s (S&P) long-term credit rating, Moody’s long-term credit rating, country origin and private indicator was collected. Data cleansing was performed again after collecting the financial data. All firms that had incomplete records in a metric necessary to calculate the working capital measures were removed from the sample (Baños-Caballero et al., 2014; Deloof, 2003). Further, all companies in the financial sector (2-digit GICS: Financials) and all real estate investment trusts (6-digit GICS: REITs) were removed because of the nature of their business (Deloof, 2003). Finally, the above data cleaning lead to companies isolated from the giant network, which were also eliminated. Table 2 gives a summary of the data cleaning. The final data set consists of 3,956 firms. Also note that DIH, DPO and DSO were winsorized at the 5% and 95% level to deal with outliers (Baños-Caballero et al., 2014; Deloof, 2003).

Table 2: Summary of data cleansing after retrieving company-specific data

Data cleansing	Number of companies
Number of initial companies	5,853
missing data	1,855
Financials (GICS 2-digit) and REITs (GICS 8-digit)	8
Isolated companies (due to above actions)	34
Final sample	**3,956**

Data Analysis Method

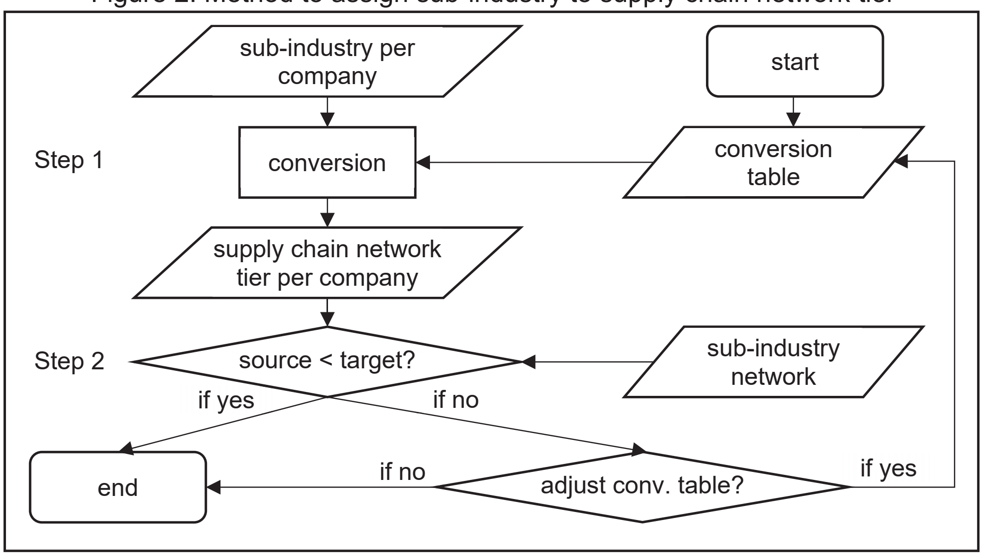
Table 3 provides an overview of the data analysis approach and sketches the structure of the following sections. The columns represent the supply chain approximation methods. The rows show the structure of the analysis by CCC levels and CCC variability. The analysis of the absolute working capital levels is applied to two different subsamples (complete sample and “subsample 10”). Subsample 10 is a more restricted sample and consist only of companies that are connected by an edge weight of 10 or more. The variability analysis is conducted on the complete sample with two measures. For each measure the analysis is divided into credit rating groups. The analyses were conducted with Gephi 0.9.2 and UCINET 6.645. Gephi was used to visualize the network and to calculate the network statistics. UCINET was used to perform the permutation tests.

Table 3: Overview of data analysis approach

<table>
<tr><td rowspan="2">Approx. approach</td><td colspan="6">Existent approach
by industry</td><td colspan="6">Network approach
by eccentricity</td></tr>
<tr><td colspan="6">Allocation of companies to supply chain network tiers by sub-industry</td><td colspan="6">Eccentricity as distance between a node v and the node farthest from node v</td></tr>
<tr><td>CCC levels</td><td colspan="3">Complete sample</td><td colspan="3">Weight ≥ 10 (Subsample 10)</td><td colspan="3">Complete sample</td><td colspan="3">Weight ≥ 10 (Subsample 10)</td></tr>
<tr><td rowspan="2">CCC variability</td><td colspan="3">ΔCCC</td><td colspan="3">$\frac{Var(CCC)}{Var(sales)}$</td><td colspan="3">ΔCCC</td><td colspan="3">$\frac{Var(CCC)}{Var(sales)}$</td></tr>
<tr><td>Invest ment</td><td>Specu lative</td><td>None (NR)</td><td>Invest ment</td><td>Specu lative</td><td>None (NR)</td><td>Invest ment</td><td>Specu lative</td><td>None (NR)</td><td>Invest ment</td><td>Specu lative</td><td>None (NR)</td></tr>
</table>

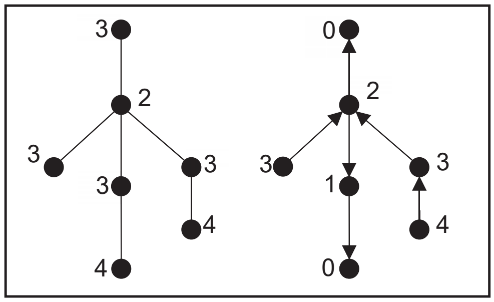
In order to determine how working capital is allocated in supply chain networks two different approaches are applied. Scholars such as Lind et al. (2012), Brandenburg (2016), Losbichler and Rothböck (2008) or Lorentz et al. (2016) used industries to divide a set of companies into different supply chain network tiers. The present paper applies this approach and labels it as the “existent approach” (see column 2 in table 3). The assignment of an industry to supply chain network tier is subjective as many different industries and products are present in the analyzed network. The process illustrated in Figure 2 was chosen to reduce subjectivity of the assignments. In the first step, a conversion table was created where the sub-industries (8-digit GICS) were assigned supply chain network tiers. The supply chain network tiers are retailer (Ret), distributor (Dist), product producers (Prod), component manufacturers (Comp), primary manufacturing (Prim) and raw material (Raw) in downstream to upstream order. The descriptions of the GICS classification were used to achieve the best possible allocation (Maitland & Gandhi, 2018). Integrated companies, consultants, service providers, energy providers, and transportation companies were assigned to the group “integrated companies” (Int) as these create value throughout the supply chain. Companies for which Bloomberg provided no values were assigned to the “not classified” group (NC). With this conversion table the corresponding supply chain network tier was assigned to each company. In the second step, the subjective assignments were checked and corrected if necessary. In particular, a sub-industry network based on the original network was created. For each relationship the sub-industry of the source node and the sub-industry of the target node were taken to form an edge list by sub-industries. Duplicate edges were merged by summing up their edge weights. Thereafter, the conversion was checked. If source and target tier did not match logical product flow it was further assessed. In the assessment connections to other industries of source and target were considered together with the edge weight. If deemed reasonable, the conversion table was adjusted to match major product flow.

Figure 2: Method to assign sub-industry to supply chain network tier



The second analysis approach incorporates the location of a company in the network instead of the industry. This approach is called the “network approach” (see column 2 in table 3). In particular, an eccentricity measure is used to group the companies, thereby, incorporating the network perspective into the analysis. The eccentricity of a graph is a distance measure. Distance is thereby defined as follows: If G has a u, v-path, then the distance from u to v, written [···] d(u, v), is the least length of a u, v-path. If G has no such path, then d(u, v) = ∞” (West, 2001). The eccentricity of a vertex u, written e(u), is maxv∈V(G)d(u, v) (West, 2001). Hence, the eccentricity e(v) of a vertex v is the distance between v and a vertex farthest from v. Figure 3 illustrates the eccentricity measure in both a directed and an undirected network. In a material flow network, vertices with an eccentricity of 0 can be interpreted as the furthest downstream as the material cannot flow any further. A high eccentricity on the other hand, means that the product has (potentially many nodes left before reaching the final customer. Therefore, a node with a high eccentricity is expected to be further upstream. In this interpretation, eccentricity represents the tier distance from the end-customer. The directed network in Figure 3 (right-hand side) illustrates this argument.

Figure 3: Graph eccentricities in an undirected (left) and directed network (right)



Statistical analysis

The analysis of working capital within the supply chain network is conducted with ANOVA (analysis of variance). However, as standard statistical tests make assumptions, which are violated by network data, a viable approach for statistical tests on a network sample is introduced. Models studying the distribution of ties or randomized tests are suitable for networks. In particular, this paper applies permutational tests for the analysis of CCC levels in the supply chain network. The general logic of permutation test in the network context is that one wants to compare the observed correlation against the distribution of correlations that one could obtain if the two variables were in fact independent of each other (Borgatti et al., 2018). The permutation test calculates all ways the experiment could have come out with actual independent scores. Then, the proportion of all assignments yielding a correlation as large as the one actually observed are counted. This value represents the “p-value” or significance of the test (Borgatti et al., 2018). The number of permutations of N objects grows very quickly with N, therefore enumerating all possible permutations is not feasible and a sample from all permutations is taken. Typically, a large sample, such as 20000 permutations is used (Borgatti et al., 2018). According to Good (2005) permutation tests rely on the assumption of exchangeability. Thereby, exchangeable means independent, identically distributed observations as well as dependent normally distributed random variables. Further, it is also a necessary condition for the application of any statistical test. Thus, permutational tests are a suitable method to identify potential differences in average working capital levels in supply chain networks. The null hypothesis for an ANOVA is presented in Equation (1), where µk represents average CCC in group k. Contrary, Equation (2) represents the alternative hypothesis.

H_0: $$\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k \quad (1)$$

H_1: Means are not all equal; at least one $\mu_i \neq \mu_j$ (2)

ANOVA is an omnibus test and as such does not indicate between which groups the statistically significant difference exists. Therefore, if the null hypothesis of the ANOVA is rejected, post-hoc permutation t-tests are necessary to identify which groups are significantly different. Equation (3) presents the null hypothesis of a t-test and Equation (4) is the alternative hypothesis of a t-test, where µk represents average CCC in group k.

H_0 $$\mu_1 = \mu_2 \quad (3)$$

H_1 $$\mu_1 \neq \mu_2 \quad (4)$$

These post-hoc tests raise the concern of alpha inflation or the multiple comparison problem. The problem arises when multiple tests are performed on the same data. Each of these tests has its own alpha. If multiple tests are performed, the cumulative error for all these tests is going to be higher than the initially set alpha (Shaffer, 1995). To mitigate this problem, alpha-adjustment procedures were developed. In particular, permutational t-tests with a subsequent manual Benjamini-Hochberg procedure and Bonferroni method were performed (Shaffer, 1995; Benjamini & Hochberg 1995).

The working capital variability in the network is analyzed based on two measures. First, analogous to Lorentz et al. (2016) the CCC changes between two consecutive periods are compared. Second, in style of the cash flow bullwhip theory by Tangsucheeva and Prabhu (2013) the working capital variability between the groups is measured by the ratio in Equation (5).

$$\text{Working capital variability (WCVar)} = \frac{\mathrm{Var(CCC)}}{\mathrm{Var(Sales)}} \tag{5}$$

The formula scales working capital fluctuations by the variance of actual sales. A relatively high ratio indicates that the variability in working capital is higher than in sales. Consequently, a firm may find it more difficult to manage working capital and may incur higher cost (Tangsucheeva & Prabhu, 2013). To control for outliers, values more than three standard deviations higher than the average were removed. The cash flow bullwhip effect represents fluctuations that occur due to seasonality and other short-term effects (Chen & Lee, 2017). The bullwhip effect may, therefore, not be adequately captured by yearly accounting data. Nevertheless, with reference to the “working capital wave” of Lorentz et al. (2016) and Chen’s (2012) theory the variability of CCC along the supply chain network is analyzed.

The analysis of CCC variability is performed on three company groups. Boissay and Gropp (2013) suggested that firms with access to outside finance (“deep pockets”) absorb the cascading effect of payment default chains. Therefore, the analysis is conducted on three credit rating groups: investment grade group, speculative-grade group and not rated group (NR) (Moody’s Investor Services, 2019; Standard & Poor’s Financial Services, 2018). The assignment to the three groups is according to the S&P long-term credit rating as of end of 2017. If a company was not rated by S&P but by Moody’s, the rating of Moody’s was used. In case Bloomberg did not provide a rating of either S&P or Moody’s, the company was assigned to the not rated group together with all companies that were listed as “not rated” in Bloomberg.

RESULTS

Descriptive Statistics

The complete supply chain network consists of 3,956 nodes that are interconnected through 19,328 edges (Table 4). Subsample 10 consists of about 15% of the nodes and 3.6% of all edges of the complete sample. The elimination of edges is also seen in the average degree, which represents the number of adjacent edges to the average node (Borgatti et al., 2018). The average degree in the complete sample is about 4.9 and 1.2 in Subsample 10. Further, average-weighted degree, which is based on the number of edges but scaled by the edge weight, is lowest in the complete sample (Newman, 2001).

The network diameter, describing the longest of all shortest paths in a network, decreases as lower weighted edges are eliminated. In the complete network the diameter has 20 edges and in the Subsample 10 only 3 edges. On the other hand, radius is defined as the shortest path between all pairs of nodes (= minimum eccentricity). The radius is zero for in the complete

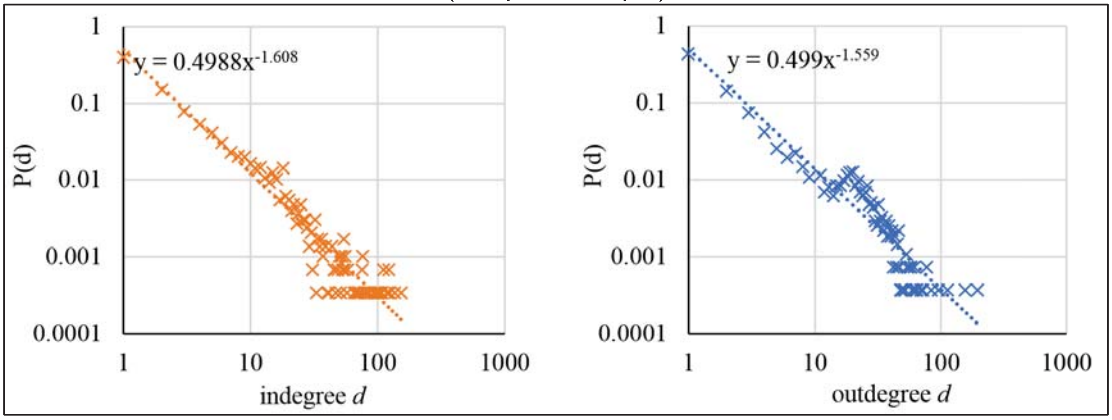
sample and in all subsamples. This indicates that at least one node has no outgoing edge. The average path length describes the average number of steps along the shortest paths for all possible pairs of network nodes. In the complete sample the measure is with about 5.3 steps much lower than the network diameter. For Subsample 10 it is 1.2 steps, respectively (Chartrand & Tian, 1997).
Graph density describes the number of ties in the network, expressed as a proportion of the number possible (Borgatti et al., 2018). The measure is slightly higher in Subsample 10 compared to the complete sample. Similarly, the average clustering coefficient is a measure for the average probability that "if A is related to B and B is related to C then there would be a relationship from A to C" (Borgatti et al., 2018). This value is highest in the complete sample.

Table 4: Network metrics

Metrics	**Complete**	**Weight 10**
Nodes	3,956	599
Edges	19,328	707
Directed	True	True
Average degree	4.886	1.18
Average weighted degree	12.676	28.064
Network diameter	20	3
Radius	0	0
Average path length	5.346	1.229
Graph density	0.001	0.002
Average clustering coefficient	0.050	0.009

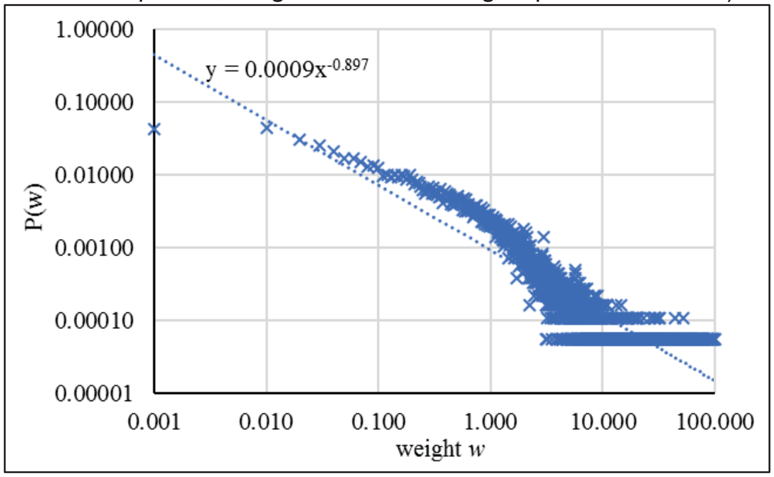
Figure 4 illustrates the connectivity distribution in the complete sample for indegree and outdegree edges. The top left point in the indegree plot indicates that a company has an indegree equal to 1 with a probability of 0.4 (i.e., 1,177 companies). The vertical lines on the bottom, right corners occur because only few companies have these relatively high degrees. If the connectivity distribution follows power law, the distribution would be following a linear line in the log-log plot. The dashed line represents the fitted power law curve. Thus, many nodes with a low degree and only a few nodes with a high degree are present in the network. Similarly, many nodes are connected by a relatively thin arrow indicating a relatively low edge weight.

Figure 4: Indegree distribution (left) and outdegree distribution (right), both on a log-log plot (complete sample)



Analog for edge weights, if the distribution follows power law, the distribution would be allocated in a linear fashion. Figure 5 depicts the edge weights in a log-log plot. The weights were rounded to two decimal places, weights of 0.005 or less were grouped in the 0.001 bin. Hence, the point on the top left indicates a probability of 0.04 that a company has a weight of 0.001 or less. The horizontal lines in the bottom right corner indicate large fluctuation in the tail for large degrees.

Figure 5: Edge weight distribution on a log-log plot (complete sample; weights rounded to two decimal places, weights below 0.005 grouped in 0.001 bin)



The descriptive statistics for the working capital measures are summarized in Table 5. Both, the mean CCC and the median CCC are highest in the complete sample at about 72 days and 64 days respectively. In the Subsample 10 the mean is at about 62 days and the median at 87 days. The difference between the means and medians as well as the values for skewness and kurtosis indicate not normally distributed samples.

Table 5: Working capital descriptive statistics, complete sample and subsamples

	mean	min	max	SD	25th %ile	median	75th %ile	kurtosis	skewness
Complete sample									
CCC	71.6	–99.1	374.6	56.7	33.7	63.5	100.3	1.829	0.982
DSO	66.8	1.2	251.5	45.4	37.1	58.4	85.7	3.262	1.508
DPO	42.4	9.3	101.1	24.9	23.3	37.1	56.9	–0.121	0.801
DIH	47.3	0.8	132.4	34.8	21.5	41.3	65.4	0.133	0.84
Subsample 10									
CCC	65.2	–84.8	282.7	43.3	37.0	56.8	87.3	1.881	1.019
DSO	64.6	1.2	251.5	32.9	43.7	58.1	78.5	4.846	1.633
DPO	41.8	9.3	101.1	21.5	25.4	39.0	54.5	0.232	0.731
DIH	42.4	0.8	132.4	28.1	24.6	37.6	54.4	1.820	1.204

CCC Level Results

The working capital allocation within the supply chain network was analyzed with two different grouping approaches: the “existent approach” and the “network approach”. Each analysis was performed on the complete sample and Subsample 10. For each combination of approach and subsample an ANOVA was calculated and, if necessary, post-hoc t-tests were performed. In the following, only the most important statistics of each analysis are presented. Regarding the post-hoc t-tests it is important to notice that the p-values presented in this paper cannot be blindly compared to an alpha of 5%, for example. The tests must be corrected for alpha inflation as explained previously. Each table showing the p-values for the t-tests indicates by a superscript which value is significant according to the Benjamini-Hochberg procedure (indicated by the superscript “a”) and the Bonferroni procedure (indicated by the superscript “b”). Thereby an alpha level of 0.05 is assumed.

<u>Existent Approach</u>

Figure 6 depicts the network graph of the complete sample with the nodes colored according to the respective supply chain tier. The node size is scaled by average CCC of each company over the entire observation period.

Figure 6: Network graph complete sample grouped by the existing approach (node size scaled by average CCC over all periods; edge size scaled by edge weight; node coloring by supply chain network tier, not classified and integrated companies colored the same)

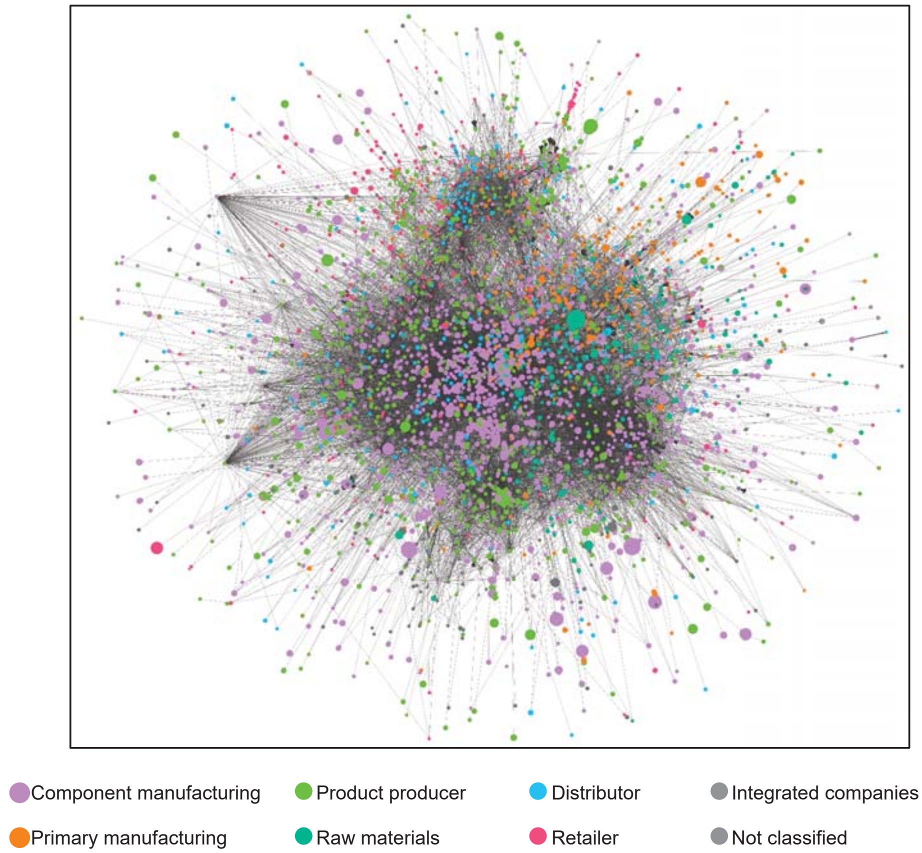
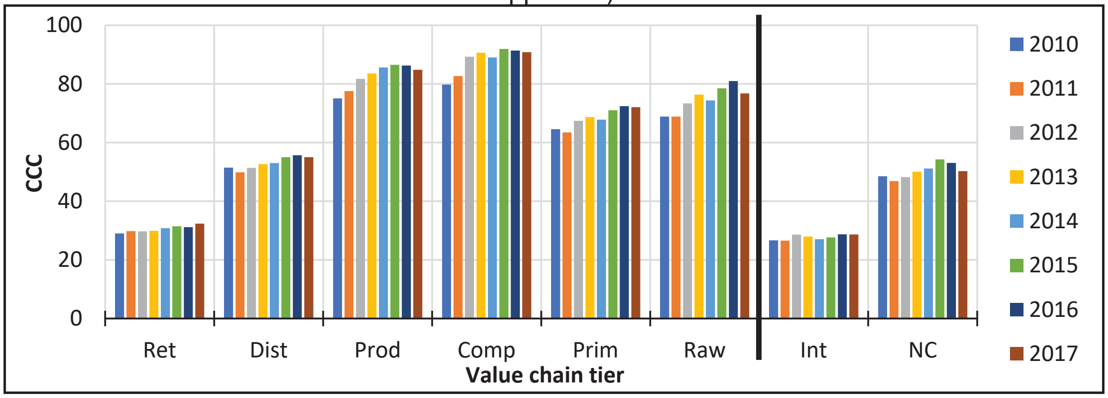


Figure 7 illustrates the average CCC per supply chain network tier and year of the complete sample. The average CCC increases between retailer (31 days average over all periods), distributor (53 days), product producer (83 days) and component manufacturers (88 days). Further up in the supply chain network, primary manufacturing (68 days) and raw materials (74 days) have both lower average CCC than product producers and component manufacturers. The black bar in Figure 7, between the raw-materials and integrated-companies groups, indicates the upstream end of the supply chain. Integrated companies and not classified companies provide services throughout the supply chain.

Figure 7: Average CCC per supply chain network tier and year (complete sample, existing approach)



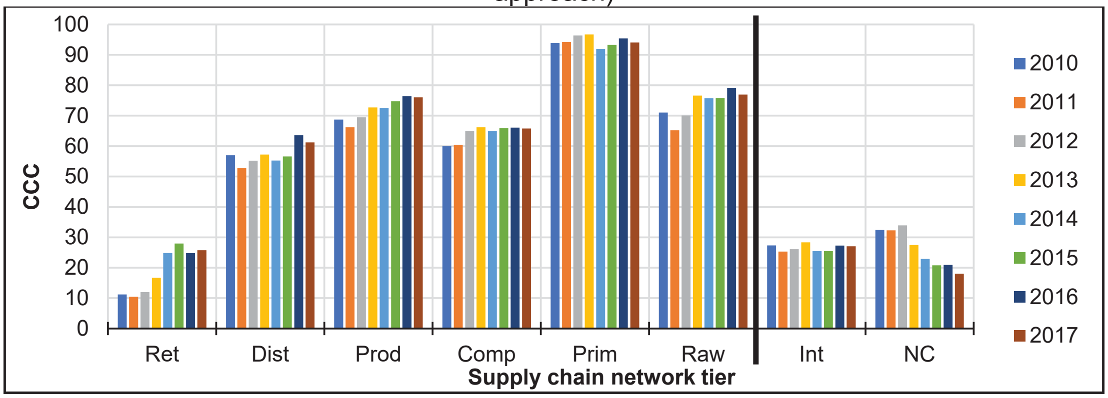
A permutational ANOVA was calculated to determine whether the average CCC is different between the groups. The results indicate that the average CCC in the supply chain network tiers are significantly different at the p<.001 level, [F(7, 31624)= 723.8016, p=0.000]. Therefore, the H0 must be rejected and post-hoc pairwise comparisons performed.
Table 6 shows the p-values of the pairwise t-tests. Only the not classified group and distributors are not significantly different at the alpha level of .05. The average CCC in the retail group is not significantly different to the integrated group after the correction by the Bonferroni method. All other possible combinations are significantly different at the .05 alpha level after Benjamini-Hochberg procedure and Bonferroni method. Qualitatively the same results hold true within the Subsample 10.

Table 6: P-values of pairwise t-test (complete sample, existing approach)

	Ret	Dist	Prod	Comp	Prim	Raw	Int	NC
Ret								
Dist	0.0000[a,b]							
Prod	0.0000[a,b]	0.0000[a,b]						
Comp	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]					
Prim	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]				
Raw	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0001[a,b]	0.0001[a,b]			
Int	0.0259[a]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]		
NC	0.0000[a,b]	0.2015	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	0.0000[a,b]	

Figure 8 shows the results from the Subsample 10. The average CCC over all periods increases from retailer to product producer. Retailers have an average CCC over all observation periods of 19 days. For distributors the average CCC is 57 days and for product producers the measure peaks at about 72 days. The CCC is at about 64 days on average for the component manufacturers. The average CCC of primary manufacturers is highest with 94 days. Raw material producers have an average CCC of 73 days. Integrated companies and the unclassified groups have an average CCC of 27 days and 20 days, respectively.

Figure 8: Average CCC per supply chain network tier and year (Subsample 10, existing approach)



The permutation ANOVA to determine whether the average CCC is different between the groups is significant at the p<.001 level [F(4, 8571)= 31.0974, p=0.000]. Therefore, H1 must be rejected and the groups further analyzed in post-hoc tests. Table 7 presents the p-values of the post-hoc t-test. The results indicate that the average CCC of retailers is not significantly different to both the unclassified group and integrated companies. Further, the average CCC of integrated companies and the not classified group is not found to be significantly different. H0 cannot be accepted for the combination raw material producers and product producers. The average CCC of all other possible combinations was found to be significantly different at the .05 alpha level after both corrections for alpha inflation. The results regarding the p-values of the post-hoc t-test indicate that the average CCC of retailers is not significantly different to both the unclassified group and integrated companies.

Table 7: P-values of pairwise t-test (complete sample, network approach)

	0	**1**	**2**	**3**	**4**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
0																
1	0.012[a]															
2	0.047	0.803														
3	0.000[a,b]	0.005[a]	0.025													
4	0.001[a]	0.016[a]	0.048	0.920												
8	0.050	0.294	0.429	0.291	0.181											
9	0.011[a]	0.231	0.427	0.087	0.066	0.756										
10	0.783	0.050	0.056	0.000[a,b]	0.000[a,b]	0.011[a]	0.001[a]									
11	0.000[a,b]	0.695	0.909	0.000[a,b]	0.002[a]	0.221	0.169	0.002[a]								

12	0.000[a,b]	0.675	0.914	0.001[a]	0.007[a]	0.286	0.235	0.006[a]	0.932							
13	0.118	0.492	0.422	0.001[a]	0.005[a]	0.162	0.087	0.237	0.193	0.213						
14	0.020[a]	0.461	0.700	0.040	0.053	0.544	0.618	0.011[a]	0.473	0.535	0.212					
15	0.382	0.944	0.930	0.089	0.061	0.418	0.456	0.291	0.951	0.919	0.730	0.713				
16	0.000[a,b]	0.002[a]	0.013[a]	0.670	0.779	0.171	0.040	0.000[a,b]	0.000[a,b]	0.001[a]	0.001[a]	0.019[a]	0.061			
17	0.052	0.016[a]	0.022[a]	0.001[a]	0.000[a,b]	0.001[a]	0.001[a]	0.005	0.001[a]	0.003[a]	0.021[a]	0.006[a]	0.045	0.001[a]		
18	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	0.000[a,b]	

Network Approach

Figure 9 visualizes the results of the network approach. Again, companies are colored by eccentricity group. The node size is scaled by average CCC of the particular company.

Figure 9: Graph complete sample grouped by the network approach (node size scaled by average CCC over all periods; edge size scaled by edge weight; node coloring by aggregated eccentricity groups)

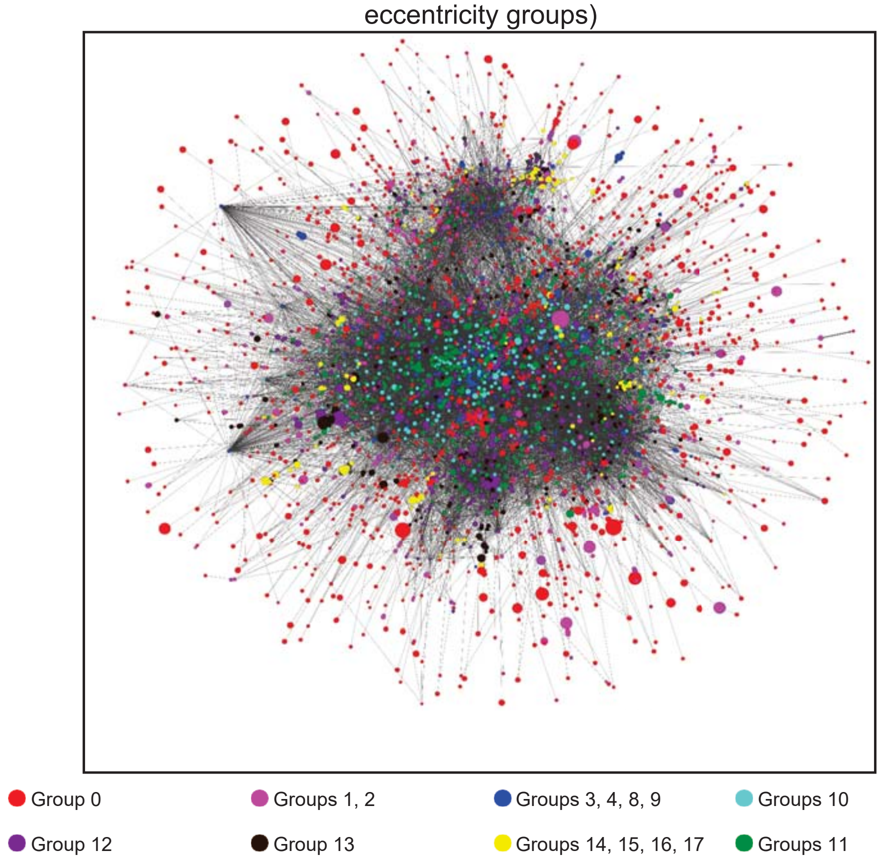
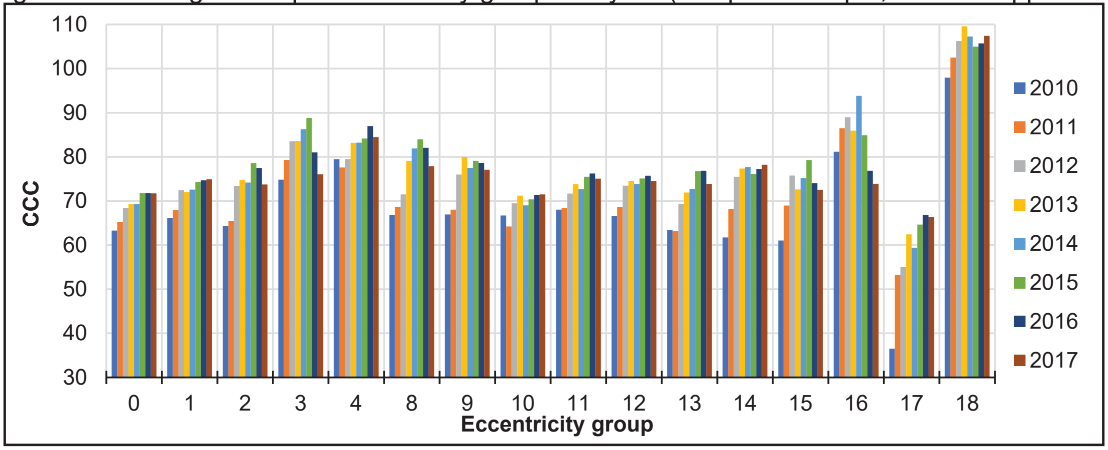


Figure 10 illustrates the development of average CCC of the 16 eccentricity groups in the complete sample. The average CCC over the entire period increases from about 69 days in group 0 to about 82 days in group 4. Thereafter the average CCC over all periods seems to decrease again to about 69 days in group 10. Further upstream, the measure stabilizes between 70 and 75 days until group 15. In group 16 the average CCC reaches a peak at about 84 days. In group 17 the measure is at about 58 days on the minimum for all groups. Group 18 has the highest average CCC of all groups with about 105 days.

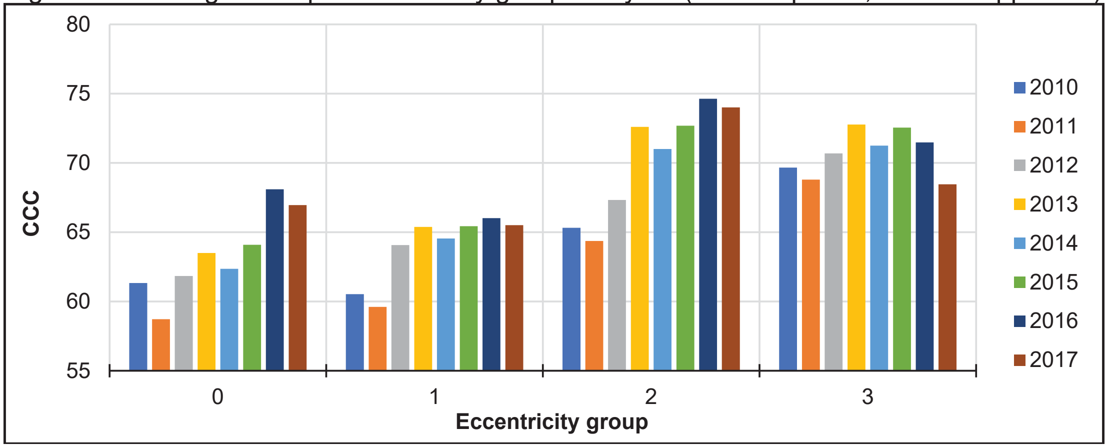
Figure 10: Average CCC per eccentricity group and year (complete sample, network approach)



The results of the permutation ANOVA revealed a significant difference of average CCC between the groups at the p<.001 level, [F(15, 31616) = 10.207, p=0.000]. Post-hoc permutation t-tests were facilitated to determine which eccentricity groups are significantly different. Table 7 summarizes the p-values for the post-hoc t-tests. H0 cannot be rejected for the tests of group 0 with Groups 2, 8, 10, 13, 15 and 17. Further, the average CCC in group 1 is found to be significantly different to the average CCC of Group 0, 3, 4, 16, 17 and 18. Group 2, however, is only significantly different to groups 16, 17 and 18. Groups 3 and 4 are both significantly different to the following groups: 0, 1, 10, 11, 12, 13, 17 and 18. Groups 8 and 9 are both significantly different to groups 10, 17 and 18. Additionally, group 9 is found to be different to group 0. These interpretations correspond to the Benjamini-Hochberg procedure or an alpha level of .05. According to the Bonferroni method fewer groups are significantly different.

Figure 11 shows the average CCC per eccentricity group in each period for Subsample 10. Group 0 has the lowest average CCC over all periods with about 63 days. In group 1 the measure is slightly higher at 64 days. Group 2 and group 3 have the highest CCC at an average of 70 days and 71 days, respectively. The permutation ANOVA to detect differences in the average CCC of the groups was significant at the p<.001 level, [F(3, 4788)= 6.0666, p=0.0007].

Figure 11: Average CCC per eccentricity group and year (Subsample 10, network approach)



The post-hoc t-tests suggest that the mean CCC of both group 0 and group 1 is significantly different to the average CCC of group 2 after the Benjamini-Hochberg procedure and Bonferroni method. For all other combinations H0 cannot be rejected at the .05 significance level (see Table 8).

Table 8: P-values of pairwise t-test (Subsample 10, network approach)

	0	**1**	**2**	**3**
0				
1	0.7476			
2	0.0011[a,b]	0.0004[a,b]		
3	0.1039	0.0766	0.9157	

CCC Variability Results

The variability of average CCC is analyzed by two measures. First, by the year-on-year CCC changes and, second, by the WCVar ratio. Further, each analysis is grouped by company rating. As in the analysis of working capital allocation the results are first presented for both the existing approach and network approach.

Existing Approach

Figure 12 shows the delta CCC for the companies by the example of the speculative-grade group. Figure 12 indicates that the year-on-year CCC changes of the speculative-grade group, the two most downstream tiers have relatively low delta CCC. This, however, increases in the further upstream groups. In the product producer group there was a more than 5-days increase between 2011 and 2012. Between the same periods the increase was even larger in the component manufacturer group. Raw material also experienced a pronounced increase 2011–2012. Similarly, for the period 2016–2017, a decrease in the CCC was observed in the product producer tier with decreasing values in further upstream tiers. The integrated-companies group shows relatively low delta CCC values.

Figure 12: Changes in CCC during the observation period (speculative grade-rated group, existing approach)

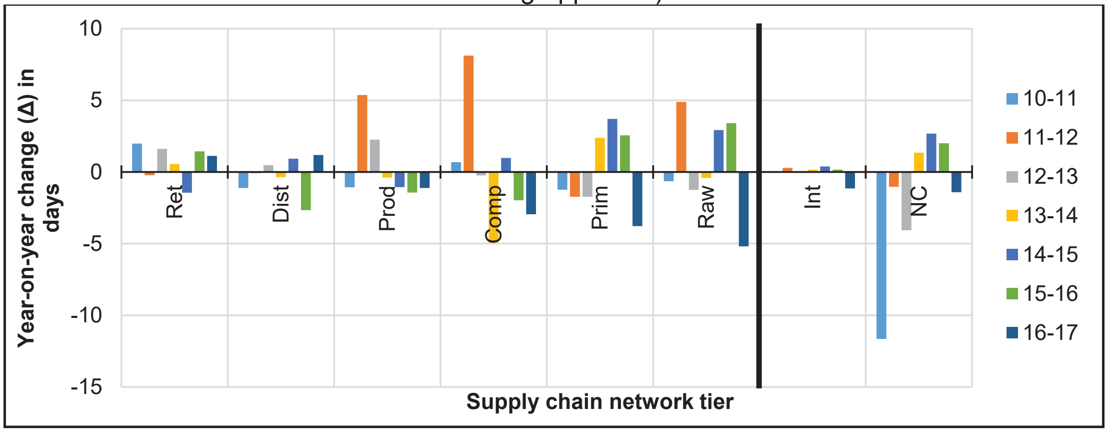
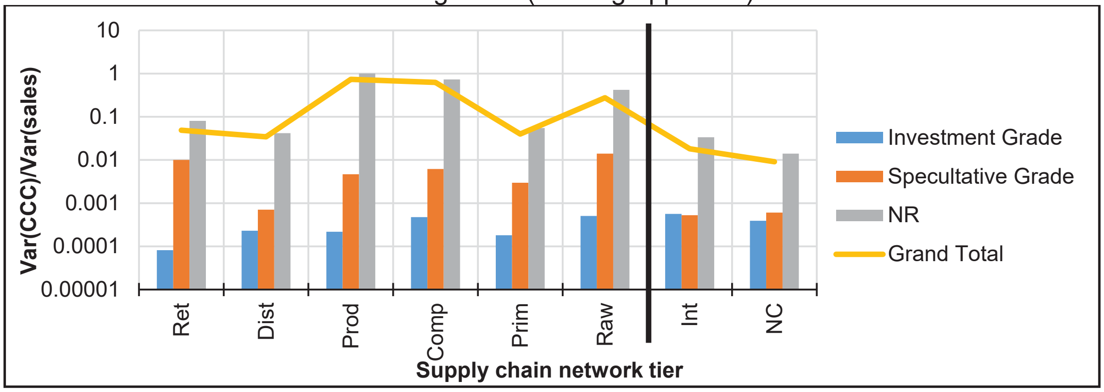


Figure 13 depicts the average WCVar ratio in each supply chain network tier on a log scale. The line illustrates the average ratio in all credit rating groups. The total measure is relatively low in the retail and distribution tiers before rising to a peak in the product producer tier. In the component manufacturer tier the measure is slightly lower than the peak. Thereafter, in the primary manufacturing tier the measure declines, only to increase again in the most upstream tier. Integrated companies and companies that were not classified have relatively low ratios. The bars in Figure 13 illustrate the level of the ratio for different rating groups. The variation seems to mainly come from the not rated companies. The values for the investment grade-rated companies are below the other two groups but show the same trend as the total measure.

Figure 13: Working capital variability ratio per supply chain network tier and credit rating group on a log scale (existing approach)



Network Approach

Figure 14 presents the changes for the companies in the speculative-grade group (similar results were obtained in the other groups). The year-on-year average CCC changes are smallest in the two most downstream groups. Further midstream, in eccentricities 2–9, the changes were more pronounced. In eccentricities 10–13 the changes were slightly lower, especially the changes between 2011 and 2012 were higher than the other year-on-year deltas in these groups. Further upstream the changes become relatively large. Figure 30 depicts the

CCC changes in the not rated group. These show a similar trend as the investment grade and speculative-grade groups: low values in the most downstream groups, followed by more pronounced changes in eccentricities 2–9, lower deltas in eccentricities 10–13 and another increasing trend until eccentricity group 17.

Figure 14: Changes in CCC during the observation period (speculative-grade-rated group, network approach)

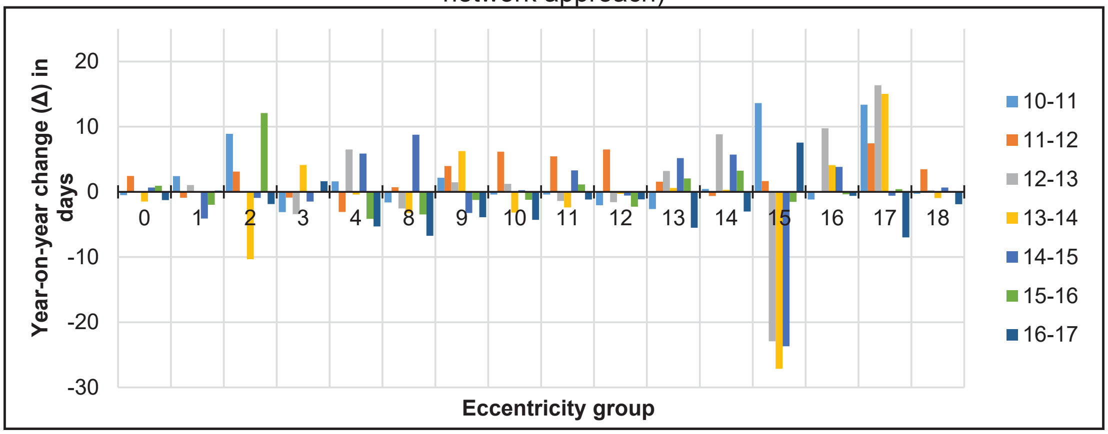
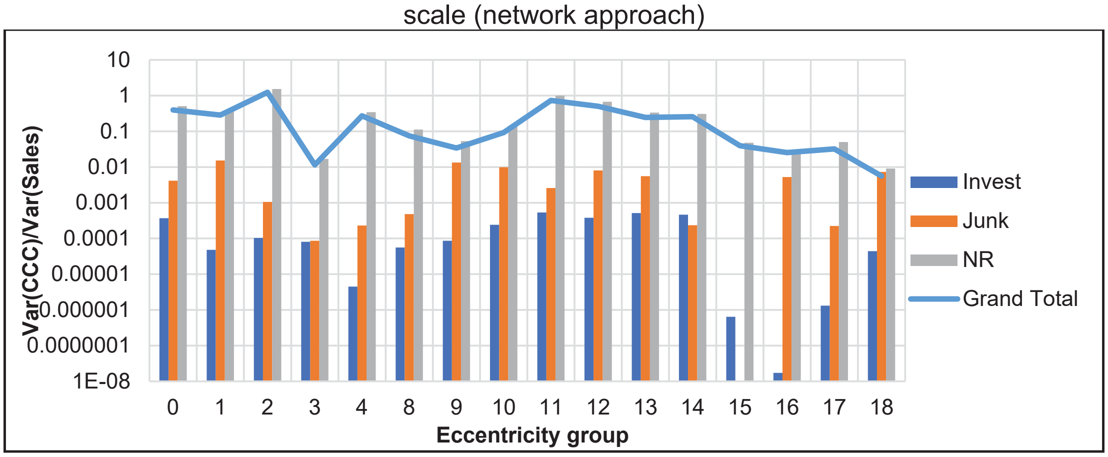


Figure 15 depicts the average WCVar ratio grouped by the network approach. In the two most downstream groups, the ratio is similar to the groups 13 and 14. It reaches a peak in group 2 and decreases again for group 3. In eccentricity 4 the measure reaches again a local peak but below the peaks in group 2 and group 11. Thereafter, the measure is relatively low until eccentricity 10. The bars in Figure 15 show the average WCVar ratio by credit rating group. The figure indicates that the main part of the variation stems from the not rated group. The values of this group closely follow the total values. The values for the speculative-grade group are generally between the not rated group and the investment grade group. The values of all groups seem to follow a similar trend as the total variation.

Figure 15: Working capital variability ratio per eccentricity group and credit rating group on a log scale (network approach)



DISCUSSION

Methodological

The network description showed that both connectivity distribution and edge weight distribution follow power law. The distributions are truncated for very high degrees and very high edge weights. According to Albert and Barabási (2002) this characteristic is found in many large real-world networks. The truncation indicates a limit to never ending growth (Amaral, Scala, Barthelemy, & Stanley, 2000). Despite having an exponential tail these distributions are significantly different to a random graph which follows a Poisson distribution. Barabási and Albert (1999) argue that a power law connectivity distribution is the consequence of network growth under preferential attachment. That is, new nodes preferentially attach to already well-connected nodes in a growing network. A consequence of preferential attachment is the formation of hub nodes and a “fit-gets-richer” phenomenon. This phenomenon occurs when in a supply chain system new firms attach to incumbent firms that are considered profitable, sizable and of use for the new entrant. The concept is similar to the first mover advantage but assumes heterogeneity of firms and includes the firm’s fitness. There is usually a limit to never ending growth, leading to the exponential cut-off and preventing a fit-gets-richer phenomenon from inexorably occurring. In a material flow network reasons for such a cut-off are, for example, the finite size of marketplaces, final consumer demand or production capacity (Hearnshaw & Wilson, 2013). Hearnshaw and Wilson (2013), therefore, reason that scale-free networks are typically found in competitive markets. Conversely, a “star network” where a large number of peripheral nodes group around a single hub could “represent an inefficient monopolistic market or a highly vertically integrated supply chain”. According to Hearnshaw and Wilson (2013) hub nodes are important in supply chain networks in that they are analogous to channel leaders. They control performance and provide supply chain wide coordination (Human & Provan, 2000; Jarillo, 1988). It might be argued that complex modern supply chains prevent authoritative, centralized structures. Still, Cowan and Jonard (2007) found that channel leaders can exert their influence, provide opportunities and motivation for other firms to align to the objectives of the leader. Achrol and Kotler (1999) state that supply chains are unlikely to coordinate, self-organize and lower transaction costs without channel leaders. These findings must be put into the perspective of the sampling method used. The degree bias in snowball sampling indicates that highly connected nodes (i.e., hub nodes) are more likely to be included (Illenberger et al., 2011). Hence, it is expected that the actual distribution shows a stronger truncation. Similarly to the connectivity distribution, the present paper finds a truncated power law distribution for edge weights. Many real-world networks have edge weights distribution that follow power law, that is, a few highly weighted connections and many lowly weighted relationships (Onnela et al., 2007). Hence, most firms have only a few critical suppliers. Hearnshaw and Wilson (2013) state that firms who have tightly coupled exchange relationships are more likely to improve their fitness compared to firms that continue with arms-length transactions.

To sum up, the supply chain network used for this paper follows the same characteristics that are found in efficient real-world networks from various disciplines: a small-world character like random graphs but unusually large clustering coefficient combined with a power law connectivity distribution as well as a power law relationship weights distribution (Albert & Barabási, 2002; Hearnshaw & Wilson, 2013). A characteristic of such a network is its resilience to disruptions and consequently its relative stability compared to random networks. Albert and Barabási (2002) argue that “hub nodes provide the network resilience against random disturbances. This is because a random disturbance event is more likely to impact a peripheral node that has few connections”. Conversely, if a hub node is impacted by disturbance the efficient network is more vulnerable than the random structure.

Content-related Implications

The analyses within the existing approach showed that the pattern of working capital allocation in the different tiers was similar across the subsamples. The retail tier and the integrated group have the lowest working capital levels in all samples. Further upstream from the retail tier, the average CCC increases and reaches a peak in the product producer or component manufacturer tier (midstream section) depending on the sample. After this peak a sharp drop in the measure is observed but it recovers again towards the upstream section. Comparing the average CCC levels between the different subsamples indicates that the CCC levels decrease as the edge weight increases. Hence, companies with stronger relationships seem to have lower average CCC. Within the network approach the difference of average CCC between the groups is not as pronounced compared to the existing approach. However, in the complete sample a similar trend in average CCC was found as in the existing approach. The average CCC in the eccentricity groups is increasing in the most downstream groups, hitting a peak in middle eccentricity groups. Overall the results suggest that the hypothesis of increasing average CCC, the further upstream a company is located, is only partly true. There are two “peaks” in the average CCC trend for the average supply chain underlying this paper with a “reset” in the midstream section.

In the existing approach the delta CCC was lowest in the most downstream group. However, between midstream tiers and upstream tiers the delta CCC did not seem to increase further. The WCVar ratio, however, increased again towards to most upstream tier showing two peaks. Further, in both variability measures the better rated groups showed lower values for both measures.

Theoretical Implications

The above presented findings are in line with existing research. Lorentz et al. (2016) found a lower net-trade-credit (i.e., subtracting DSO from DPO) for further upstream companies. This suggests an increasing CCC in further upstream companies when DIH is constant in all tiers. Further, Lorentz et al. (2016) indicated that period-on-period delta DSO and delta DPO appear larger further upstream explaining it by the relative power of the companies. While Lorentz et al. (2016) found a “working capital wave” for net trade credit and its components, the present paper suggests the existence of such a “working capital wave” for the CCC measure. In particular, the results suggest that there might be two consecutive working capital waves in a supply chain. The reason for these two waves can be traced back to highly liquid companies with access to outside finance (“deep pockets”) interrupt the wave from cascading further upstream. Boissay and Gropp (2013) and Serrano et al. (2018) have shown that the propagation of payment variability (i.e., financial contagion) is mitigated by supply chain members with solid financial statements. Conversely, “weak” financial statements amplified the effect (Serrano et al., 2018). Similarly, Shenoy and Williams (2017) found that companies with better access to bank liquidity offer more trade credit to customers. Serrano et al. (2018) argue that “deep pockets” decouple the supply chain in terms of financial contagion. A “double working capital wave” is further in line with the findings of Losbichler and Rothböck (2008). They suggested that in a supply chain the leading branch, i.e. a company with a high bargaining power, can reduce its working capital burden more than its suppliers and customers. Hence, the CCC level and possibly also the variability of CCC may not steadily increase between downstream and upstream tiers. Cash-rich companies might be able to break this pattern thanks to their bargaining power. Hence, the finding that the working capital of a firm is not only driven by firm specific factors but also the composition of the supply chain network. In this light, the paper at hand underlines the

importance of collaboration between companies to reduce negative material and financial spillover effects. Through the better coordination of working capital, companies can reduce cost, improve resilience to shocks and therewith increase the competitiveness of supply chain networks (Hofmann & Kotzab, 2010; Protopappa-Sieke & Seifert, 2017). For example, by adopting supply chain finance solutions such as reverse factoring as they have the potential to reduce the cascading effect in the supply chain networks like “deep pocket” companies.
One main reason for high working capitals in upstream supply chain networks can be traced back to the inventory bullwhip effect (Luo & Shang, 2012; Tangsucheeva & Prabhu, 2013). The spread of the amplified order and the actual demand results in products remaining in inventory for an extended period (Tangsucheeva & Prabhu, 2013). In the same vein, the amplified orders require more capital in the upstream stage to make the inventory payments (Luo & Shang, 2012). Accordingly, a reduction of the material bullwhip effect would reduce the working capital burden for upstream companies. Lee et al. (1997) give three ways to reduce the inventory bullwhip effect: information sharing, channel alignment and operational efficiency. These approaches all require inter-organizational coordination. Large inventories could also be financed through inventory financing where a financial institution provides a loan against collateralized inventory (Zhao & Huchzermeier, 2018).

Practical Implications

A supply-chain-wide coordination is necessary to tackle the working capital burdens of further upstream companies. Through coordination and an increased awareness for consequences of actions in a crisis, practitioners can stabilize to the benefit of all involved parties. This paper showed that similar patterns existed regardless of the relationship strength. Therefore, the key message for practitioners here is that collaboration with important supply chain partners should be approached first. This would improve the working capital situation of partners that contribute most to a product. Cash rich companies should take on the responsibility of pushing towards this goal. Because they can exert their influence on other firms to align to the objectives and reach a supply-chain-wide coordination (Hearnshaw & Wilson, 2013). Further, channel leaders provide resilience to an efficient network and, thus, their “fitness” should be improved (Albert, Jeong, & Barabási, 2000).
Practitioners can reduce CCC variability and financial contagion in supply chain networks with innovative financing solutions (Luo & Shang, 2012). Reverse factoring for example, aligns the cash flow and the product flow to some extent, thereby, reducing the impact of cash flow variability. This is line with Caniato et al. (2016) who found that companies use reverse factoring “with the main purpose of reducing bankruptcy risk in the upstream supply chain”. Another solution is dynamic discounting, improves the working capital situation for the companies within the sampled networks (Zhao & Huchzermeier, 2018).
The observed increasing average CCC along the supply chain poses a challenge to affected companies. They must be able to finance the higher working capital burden and withstand its variability. Attempts to reduce this cascading effect in CCC can be addressed on all of its components.

CONCLUSION

The aim of the present study was to analyze working capital from a network perspective. Previous research incorporated the network perspective into working capital analysis by approximating supply chains by industry. The present study extends previous research by applying social network analysis in the context of working capital management. The results of the empirical analysis revealed a “double working capital wave” in the network. Existing literature suggests increasing working capital levels and working capital variability in further upstream tiers. This study supports these findings but finds that the existence of liquid companies with access to outside finance may break the cascading effect. Therefore, the working capital increases in further upstream companies until reaching a "deep pocket". The "deep pocket" provides liquidity and, thus, the cascade restarts.

The originality of this research lies in its attempt to apply a social network perspective to the concept of supply chain-oriented working capital management. It therewith enables an integrated view on the topic by bridging several literature streams from the working capital, logistics/supply chain and social network theory. The paper at hand additionally establishes an understanding of how the presence of liquid companies with access to outside finance may interrupt the financial spillover effects in supply chain networks. Moreover, it complements existing financial management literature that focuses on the effects of supplier-customer interactions by including a thorough investigation of financial spillover effects in a supply chain environment. Finally, it provides valuable insights to both researchers as well as practitioners on how better working capital coordination in a supply chain network can benefit all players involved. This may be achieved, inter alia, through the usage of supply chain finance (SCF) solutions.

It can be argued that the main limitation of this paper lies in the snowball sampling approach, which is known to be prone for conceptual and methodological biases. Examples include the selection bias, degree bias and status homophily (Illenberger et al., 2012). Another limitation from this paper arises from the usage of yearly accounting data. This restricts the analysis of potential cash flow bullwhip effects, which are known to occur in short time intervals. Future research could address the methodical limitations arising from snowball sampling by using more seed companies or applying alternative methods (e.g., Kirchherr & Charles, 2018). In addition, the decomposition of the CCC into its components would be interesting and possibility yield a more differentiated analysis. Regarding the variability of CCC within supply chain networks it is further desirable to analyze the existence of a financial bullwhip effect on a real-world sample.

REFERENCES

Achrol, R. S., & Kotler, P. (1999). Marketing in the Network Economy. *Journal of Marketing*, *63*(special issue), 146–163.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, *74*(1), 47–97.

Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, *406*((6794)), 378–382.

Amaral, L. A., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11149–11152.

Baños-Caballero, S., García-Teruel, P. J., & Martínez-Solano, P. (2014). Working capital management, corporate performance, and financial constraints. *Journal of Business Research*, *67*(3), 332–338.

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Boissay, F., & Gropp, R. (2013). Payment Defaults and Interfirm Liquidity Provision. *Review of Finance*, *17*(6), 1853–1894.

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing Social Networks* (2nd ed.). London: SAGE.

Brandenburg, M. (2016). Supply chain efficiency, value creation and the economic crisis – An empirical assessment of the European automotive industry 2002–2010. *International Journal of Production Economics*, *171*(3), 321–335.

Caniato, F., Gelsomino, L. M., Perego, A., & Ronchi, S. (2016). Does finance solve the supply chain financing problem? *Supply Chain Management: An International Journal*, *21*(5), 534–549.

Chartrand, G., & Tian, S. (1997). Distance in Digraphs. *Computers & Mathematics with Applications*, *34*(11), 15–23.

Chen, H. L. (2012). Empirical behavioral analysis of project contractors' supply-chain payment terms. *Supply Chain Management: An International Journal*, *17*(3), 277–289.

Chen, L., & Lee, H. L. (2017). Modeling and Measuring the Bullwhip Effect. In A. Y. Ha & C. S. Tang (Eds.), *Handbook of Information Exchange in Supply Chain Management* (p. 3–25). Los Angeles: Springer.

Chen, T.-K., Liao, H.-H., & Kuo, H.-J. (2013). Internal liquidity risk, financial bullwhip effects, and corporate bond yield spreads: Supply chain perspectives. *Journal of Banking & Finance*, *37*(7), 2434–2456.

Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, *19*(3), 351–366.

Choi, T. Y., & Hong, Y. (2002). Unveiling the structure of supply networks: case studies in Honda, Acura, and DaimlerChrysler. *Journal of Operations Management*, *20*(5), 469–493.

Christopher, M. (2005). *Logistics and Supply Chain Management: Creating Value-Adding Networks* (3rd ed.). Harlow: FT Prentice Hall.

Chu, Y. (2012). Optimal capital structure, bargaining, and the supplier market structure. *Journal of Financial Economics*, *106*(2), 411–426.

Cole, N. (1991). When the Numbers Do Not Add Up. *Director*, *44*(6), 61–68.

Cowan, R., & Jonard, N. (2007). Structural holes, innovation and the distribution of ideas. *Journal of Economic Interaction and Coordination*, *2*(2), 93–110.

Cox, A., Sanderson, J., & Watson, G. (2001). Supply Chains and Power Regimes: Toward an Analytic Framework for Managing Extended Networks of Buyer and Supplier Relationships. *Journal of Supply Chain Management*, *37*(1), 28–35.

Daugherty, P. J., Richey, R. G., Roath, A. S., Min, S., Chen, H., Arndt, A. D., & Genchev, S. E. (2006). Is collaboration paying off for firms? *Business Horizons*, *49*(1), 61–70.

Deloof, M. (2003). Does Working Capital Management Affect Profitability of Belgian Firms? *Journal of Business Finance & Accounting*, *30*(3–4), 573–587.

Ek, R., & Guerin, S. (2011). Is there a right level of working capital? *Journal of Corporate Treasury Management*, *4*(2), 137–149.

Enqvist, J., Graham, M., & Nikkinen, J. (2014). The impact of working capital management on firm profitability in different business cycles: Evidence from Finland. *Research in International Business and Finance*, *32*(8), 36–49.

Farris, M. T., & Hutchison, P. D. (2002). Cash-to-cash: the new supply chain management metric. *International Journal of Physical Distribution & Logistics Management*, *32*(4), 288–298.

Farris, T. M., & Hutchinson, P. D. (2003). Measuring Cash-to-Cash Performance. *The International Journal of Logistics Management*, *14*(2), 83–92.

Filbeck, G., & Krueger, T. M. (2005). An Analysis of Working Capital Management Results Across Industries. *Mid-American Journal of Business*, *20*(2), 11–18.

Good, P. (2005). *Permutation, parametric, and bootstrap tests of hypotheses* (3rd ed.). New York: Springer.

Griffiths, P., Gossop, M., Powis, B., & Strang, J. (1993). Reaching hidden populations of drug users by privileged access interviewers: methodological and practical issues. *Addiction*, *88*(12), 1617–1626.

Grosse-Ruyken, P., Wagner, S., & Jonke, R. (2011). What is the Right Cash Conversion Cycle for Your Supply Chain? *Int. J. of Services and Operations Management*, *10*(1), 13–29.

Hawawini, G., Viallet, C., & Vora, A. (1986). Industry Influence on Corporate Working Capital Decisions. *Sloan Management Review*, *27*(4), 15–24.

Hearnshaw, E. J. S., & Wilson, M. M. J. (2013). A complex network approach to supply chain network theory. *International Journal of Operations & Production Management*, *33*(4), 442–469.

Hofmann, E., & Kotzab, H. (2010). A Supply Chain-Oriented Approach of Working Capital Management. *Journal of Business Logistics*, *31*(2), 305–330.

Human, S. E., & Provan, K. G. (2000). Legitimacy Building in the Evolution of Small-Firm Multilateral Networks: A Comparative Study of Success and Demise. *Administrative Science Quarterly*, *45*(2), 327–365.

Illenberger, J., & Flötteröd, G. (2012). Estimating network properties from snowball sampled data. *Social Networks*, *34*(4), 701–711.

Illenberger, J., Kowald, M., Axhausen, K. W., & Nagel, K. (2011). Insights into a spatially embedded social network from a large-scale snowball sample. *The European Physical Journal B*, *84*(4), 549–561.

Itzkowitz, J. (2013). Customers and cash: How relationships affect suppliers' cash holdings. *Journal of Corporate Finance*, *19*, 159–180.

Jarillo, C. J. (1988). On Strategic Networks. *Strategic Management Journal*, *9*(1), 31–41.

Jarillo, C. J., & Stevenson, H. H. (1991). Co-operative strategies—The payoffs and the pitfalls. *Long Range Planning*, *24*(1), 64–70.

Kale, J. R., & Shahrur, H. (2007). Corporate capital structure and the characteristics of suppliers and customers. *Journal of Financial Economics*, *83*(2), 321–365.

Kim, W.-S. (2009). Effects of a Trust Mechanism on Complex Adaptive Supply Networks: An Agent-Based Social Simulation Study. *Journal of Artificial Societies and Social Simulation*, *12*(3), 1–4.

Kim, Y., Choi, T. Y., Yan, T., & Dooley, K. (2011). Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management*, *29*(3), 194–211.

Kirchherr, J., & Charles, K. (2018). Enhancing the sample diversity of snowball samples: Recommendations from a research project on anti-dam movements in Southeast Asia. *PLoS ONE*, *13*(8), 1–17.

Knauer, T., & Wöhrmann, A. (2013). Working capital management and firm profitability. *Journal of Management Control*, *24*(1), 77–87.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). The Bullwhip Effect in Supply Chains. *Sloan Management Review 38*(3), 93–102.

Lee, H. L., So, K. C., & Tang, C. S. (2000). The Value of Information Sharing in Two-Level Supply Chain. *Management Science*, *46*(5), 626–643.

Lind, L., Pirttilä, M., Viskari, S., Schupp, F., & Kärri, T. (2012). Working capital management in the automotive industry: Financial value chain analysis. *Journal of Purchasing and Supply Management*, *18*(2), 92–100.

Lorentz, H., Solakivi, T., Töyli, J., & Ojala, L. (2016). Trade credit dynamics during the phases of the business cycle – a value chain perspective. *Supply Chain Management: An International Journal*, *21*(3), 363–380.

Losbichler, H., & Rothböck, M. (2008). Der Cash-to-cash Cycle als Werttreiber im SCM-Ergebnisse einer europäischen Studie. *Controlling & Management*, *52*(1), 47–57.

Luo, W., & Shang, K. (2012). *Integrating Inventory Replenishment and Cash Payment Decisions in Supply Chains*. Durhan.

Maitland, M. F., & Gandhi, S. (2018). GICS strucutre & sub-industry definition. Retrieved from https://www.msci.com/gics

Mizgier, K. J., Wagner, S. M., & Holyst, J. A. (2012). Modeling defaults of companies in multi-stage supply chain networks. *Internation Journal Production Economics*, *135*(1), 14–23.

Moody's Investor Services. (2019). Moody's Rating Scale and Definitions.

Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, *64*(1), 1–7.

Oliveira, M., Kadapakkam, P.-R., & Beyhaghi, M. (2017). Effects of customer financial distress on supplier capital structure. *Journal of Corporate Finance*, *42*, 131–149.

Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., & Barabási, A.-L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, *104*(18), 7332–7336.

Osadchiy, N., Gaur, V., & Seshadri, S. (2011). *Systematic Risk and the Bullwhip Effect in Supply Chains*.

Ouyang, Y., & Li, X. (2010). The bullwhip effect in supply chain networks. *European Journal of Operational Research*, *201*(3), 799–810.

Pathak, S., Day, J. M., Nair, A., Sawaya, W. J., & Kristal, M. (2007). Complexity and Adaptivity in Supply Networks: Building Supply Network Theory Using a Complex Adaptive Systems Perspective. *Decision Sciences*, *38*(4), 547–580.

Protopappa-Sieke, M., & Seifert, R. W. (2017). Benefits of working capital sharing in supply chains. *Journal of the Operational Research Society*, *68*(5), 521–532.

Randall, W. S., & Farris, M. T. (2009). Supply chain financing: Using cash-to-cash variables to strengthen the supply chain. *International Journal of Physical Distribution & Logistics Management*, *39*(8), 669–689.

Robins, G. (2015). *Doing social network research. network-based research design for social scientists*. (K. Metzler, Ed.). Thousand Oaks: Sage.

Rogers, H., Tannock, J., Bourlakis, M., & Braziotis, C. (2013). Supply chains and supply networks: distinctions and overlaps. *Supply Chain Management: An International Journal*, *18*(6), 644–652.

Serrano, A., Olivio, R., & Kraiselburd, S. (2018). Risk propagation through payment distortion in supply chains. *Journal of Operations Management*, *58–59*(1), 1–14.

Shaffer, J. P. (1995). Multiple Hypothesis Testing. *Annual Review of Psychology*, *46*(1), 561–584.

Shenoy, J., & Williams, R. (2017). Trade credit and the joint effects of supplier and customer financial characteristics R. *Journal of Financial Intermediation*, *29*, 68–80.

Smith, K. V. (1980). *Readings on the management of working capital* (2nd ed.). St.Paul: West Publishing.

Standard & Poor's Financial Services. (2018). Guide to Credit Rating Essentials: What are credit ratings and how do they work?

Surana, A., Kumara, S., Greaves, M., & Raghavan, N. (2005). Supply-chain networks: A complex adaptive systems perspective. *International Journal of Production Research*, *43*(20), 4235–4265.

Tangsucheeva, R., & Prabhu, V. (2013). Modeling and analysis of cash-flow bullwhip in supply chain. *International Journal of Production Economics*, *145*(1), 431–447.

Viskari, S., & Kärri, T. (2012). A model for working capital management in the inter-organisational context. *International Journal of Integrated Supply Management*, *7*(1/2/3), 61–79.

Vogt, W. P., & Johnson, B. (2011). *Dictionary of statistics & methodology: a nontechnical guide for the social sciences* (4th ed.). Thousand Oaks: Sage.

Wagner, S. M., Bode, C., & Koziol, P. (2009). Supplier default dependencies: Empirical evidence from the automotive industry. *European Journal of Operational Research*, *199*(1), 150–161.

Wangphanich, P., Kara, S., & Kayis, B. (2010). Analysis of the bullwhip effect in multi-product, multi-stage supply chain systems - a simulation approach. *International Journal of Production Research*, *48*(15), 4501–4517.

Weinraub, H. J., & Visscher, S. (1998). Industry Practice Relating to Aggressive Conservative Wokring Capital Policies. *Journal of Financial and Strategic Decisions*, *11*(2), 11–18.

West, D. B. (2001). *Introduction to Graph Theory* (2nd ed.). Dehli: Pearson Education.

Wöhrmann, A. (2012). Kostenmanagement in Krisenzeiten: Rentabilitätssteigerung durch Working Capital Management? *Controlling and Management*, *56*(3), 83–88.

Zhao, L., & Huchzermeier, A. (2018). *Supply Chain Finance - Integrating Operations and Finance in Global Supply Chains*. Cham: Springer.

DECISION SCIENCES INSTITUTE

Patient Hospital Satisfaction Ratings on Social Media versus HCAHPS Survey Scores

Subhajit Chakraborty
Coastal Carolina University
Email: schakrabo@coastal.edu

E. Mitchell Church
Coastal Carolina University
Email: mchurch@coastal.edu

ABSTRACT

We demonstrate that hospital satisfaction ratings expressed by patients freely on social media are positively related to the patient experience of care domain score that CMS calculates based on the HCAHPS surveys. We suggest that hospital leaders can know about their patient experiences by frequently tapping into social media.
Keywords: HCAHPS surveys; patient hospital satisfaction rating; patient experience of care score; social media.

INTRODUCTION

The last few years have seen a fundamental transformation of the modern U.S. hospital. A lot of consolidation efforts and changing margins have led to hospitals incorporating principles and practice from other business sectors, perhaps more than ever before. As more hospital executives run their hospitals like traditional businesses, one area of renewed focus is patient satisfaction in the quest to become more patient-centric. While always a point of emphasis, hospitals have recently begun to take note of the attention given to initiatives around customer satisfaction and retention in other service industries. Companies with high customer satisfaction scores enjoy a wide range of tangible economic benefits including increased sales, greater profitability and lower customer turnover (Heskett, 1987; Heskett, Jones, Loveman, Sasser, & Schlesinger, 1994; Heskett, Sasser, & Schlesinger, 1997).

The internet has made soliciting the opinions of customers easier than ever before as a result of advancements in Web 2.0 technologies and the development of websites built around online customer reviews. Many organizations, in particular manufacturing and retail service organizations, have been quick to embrace online customer reviews as a new way of truly understanding their customer grievances, thereby leading to their needs and wants. These organizations realize that in the age of social media, customers have a much louder voice and the ability to influence large segments of an organization's customer base (Brodalski, 2011; Holt, 2016; Killian & McManus, 2015). It is not uncommon to see many service organizations go to great lengths and provide significant incentives to solicit customer comments from organizations. However, despite these advantages and an espoused interest in improving patient satisfaction scores, many hospitals lag behind other service organizations in their use of new and emerging data resources and technology (Acohido, 2013; Hernandez, Machacz, & Robinson, 2015; Landi, 2018; Sullivan, 2018). As a result, investment in online customer reviews have been largely overlooked by U.S. hospitals.

The primary reason for this is that U.S. hospitals have for quite some time now relied on industry-specific customer feedback mechanisms—the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey and other tools (Cahill & Wang, 2012).

These standardized surveys are designed to capture key aspects of the patient experience relevant for regulatory requirements and key aspects of health outcomes. While HCAHPS surveys serve a valuable function within the hospital organizations, they are limited when compared with online customer reviews. First, an important consideration is the cost and requisite resources needed to administer HCAHPS. With online customer reviews, customers provide all of the information and are completely in charge of what they say with all the customer reviews. Thus, the onus of data creation and management is placed on the customer and the site collating the reviews such as Yelp or Healthgrades.

A second advantage of online customer reviews is their timeliness. While HCAHPS undeniably provide quality data, there is a significant lag time between patient interaction and HCAHPS reporting (Ranard et al., 2016). In today's hospital environment, many organizations should not be willing to wait so long to see the impact of their changing initiatives, especially when other data resources are available. This is especially true in the case of mergers and acquisitions, when hospitals move into new areas where patient feedback about the organizational transformation is crucial and may even adversely affect patient experience (V. Ho & Hamilton, 2000). Online customer reviews offer an almost instantaneous way for patients to share their feelings and experiences with the hospital organization.

A third difference between online customer reviews and hospital managed feedback mechanisms is the lack of constraints on customer interaction (Ranard et al., 2016). Authors of online customer reviews are free to deal with any topics that they choose, without the constraining influence inherent in all standardized survey instruments. Such freedom of response even extends to the online review authors themselves. Online customer reviews may even capture long-tail populations (Brynjolfsson, Hu, & Smith, 2006) totally untouched by HCAHPS, in the form of reviews authored by extended family members, caregivers or other individuals only tangentially related to patient interaction.

Our study attempts to highlight the value of online customer reviews by examining their usefulness as a metric of gauging hospital interactions. We use online review data from Yelp.com, one of the largest and most comprehensive online customer review websites today in an attempt to answer the research question: *do patient hospital satisfaction ratings on social media such as Yelp provide similar information as HCAHPS surveys?* Results from our study demonstrate that they do in several ways. We show that online customer reviews are useful for predicting some of the similar indicators of patient experience as calculated with the HCAHPS survey data. At the same time, we also control for the different hospital types thereby showing that while there may be challenges in soliciting customer data for certain hospitals, its importance lies undiminished for any type of hospital.

The remainder of our paper is organized as follows. First, we discuss the extant literature on patient satisfaction, the scant literature backing the use of social media in healthcare and introduce our hypothesis. Second, we describe the methods and procedures followed for data collection. Third, we elaborate the results of the study and discuss the findings. Finally, we note the limitations of the study and end with a discussion on practice implications and future directions.

LITERATURE REVIEW AND HYPOTHESIS

In healthcare and marketing literature, patient satisfaction is widely accepted as a multi-dimensional concept, based on a relationship between patient experiences and expectations (Sixma, Kerssens, Campen, & Peters, 1998). In one of the early works on patient satisfaction, Pascoe (1983) defined patient satisfaction as: "…the health care recipient's reaction to salient aspects of the context, process, and result of their service experiences… (pp. 189)." He noted that satisfaction consists of a "…cognitively based evaluation or grading of directly-received services including structure, process, and outcome of services… and an affectively based response to the structure, process, and outcome of services…(pp. 189)". Similar to customers of any service, most patients have some expectations of the health service in their mind when they visit a hospital. Patients evaluate the healthcare service actually received at the hospital against their expectations to form their overall impression about the service; and some of them choose to express their dissatisfaction or happiness later on social media such as Yelp.

In the U.S. healthcare system, typically, individual patient satisfaction has generally been obtained from a few random patients at every hospital by the HCAHPS surveys and aggregated to the hospital level to calculate the hospital's patient experience of care domain score. Once a hospital or its survey vendor submits the HCAHPS survey data to the data warehouse for the hospital inpatient quality reporting program, CMS cleans the submitted data by removing incomplete surveys, surveys from ineligible patients and calculates the "top-box" raw score for the nine HCAHPS measures that are used in the Hospital value-based purchasing (VBP)—Patient Experience of Care Domain. CMS applies the patient-mix adjustment for each of the nine HCAHPS measures used in Hospital VBP to obtain unrounded patient-mix adjusted scores. After performing more intermediate steps including survey mode adjustment for each of the nine HCAHPS measures used in Hospital VBP, CMS calculates the Unweighted Patient Experience of Care Domain Score (out of 100) for each hospital by summing the hospital's HCAHPS base score (0-80) and HCAHPS consistency score (0-20).

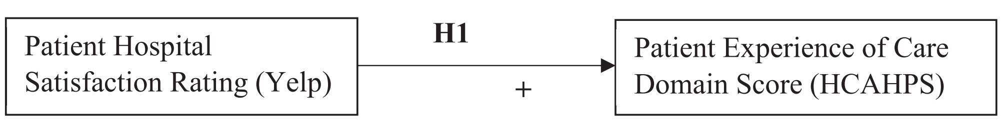
These established methods for capturing patient satisfaction have significant merits, yet online customer reviews also have a long history of successfully capturing these types of opinions across many different areas of the economy including retail and hospitality (Chevalier & Mayzlin, 2006; Liu & Park, 2015). Moreover, the few studies on use of social media use in healthcare (Church & Chakraborty, 2018; Ranard et al., 2016) have identified positive effects of social media use by patients. Social media can aid patients (Smailhodzic, Hooijsma, Boonstra, & Langley, 2016) by giving them the ability to read more about the disease/ailment and complement the information provided by healthcare professionals (Rupert et al., 2014) and also by providing psychosocial support (Y.-X. Ho, O'Connor, & Mulvaney, 2014). Use of social media by patients can also aid healthcare professionals by providing them a tool to become more responsive to patients' concerns thereby helping efforts for brand building and improved service delivery (Li, 2010; Williams, 2011) all of which ultimately strengthen the hospital's market position (McCaughey et al., 2014; Williams, 2011). However, as highlighted above, these social media sources are still not widely used in healthcare; at least not in a major way in U.S. hospitals.

Based on the conceptualization of patient satisfaction and the extant literature on the usefulness of social media ratings in healthcare (Church & Chakraborty, 2018; Greaves et al., 2014, 2012; Lagu, Goff, Hannon, Shatz, & Lindenauer, 2013; Ranard et al., 2016) we argue that information similar to HCAHPS ratings can be obtained from free social media websites such as Yelp. Therefore, we hypothesize:

H1: Patient hospital satisfaction rating obtained from Yelp is positively related to the patient experience of care score obtained from HCAHPS surveys.

Figure 1 depicts our research model. In the next sections, we discuss the methods used in the evaluation and testing of this hypothesis.

FIGURE 1: Research Model



METHODS

Procedure

We collected all available hospital reviews (24,313 reviews of 397 hospitals) from the commercial Yelp website (www.yelp.com) that were posted by patients/kin over a period spanning 11 years (9-May-2005 to 23-Jun-2016). Consistent with other studies, we avoided any hospitals that had reviews flagged as being suspicious or fake. Yelp ensures that review content is genuine and provides a notification if a business is accused of having reviews that have been faked or purchased by the organization. No reviews flagged as potentially fake were included in the sample. We also eliminated hospitals with less than 10 reviews so as to include only those hospitals which had a good presence among the Yelp reviewers. We matched up the hospital names and address information with CMS' 2016 Hospital Compare database. From CMS' database, two hospital attributes were also collected and dummy coded: (a) hospital type that had the following values: Acute Care Hospitals (1), Children's (2) and Critical Access Hospitals (3); and (b) hospital ownership that had the following values: Proprietary (1), Government-Federal (2), Government-Hospital District or Authority (3), Government-Local (4), Government-State (5), Physician (6), Voluntary non-profit-Church (7), Voluntary non-profit-Other (8), and Voluntary non-profit-Private (9).

We removed hospital reviews if they did not have either of the study variables; 26 hospitals did not have their average hospital star rating reported on Yelp while another 2 hospitals did not have the overall star rating reported in the CMS' Hospital Compare database. All 28 reviews were removed, leaving a final sample of 359 U.S. hospital level records which we used for our analyses.

Variables

Yelp calculates the average of all the star ratings (1 through 5) given by patients/kin for each hospital and displays it along with other hospital meta-data. We denoted the average hospital star rating given by patients/kin as the *patient hospital satisfaction*. Next, the two dummy coded variables taken from CMS—hospital type and the hospital ownership were used as controls in the hospital-level regression. We also used a Yelp meta-data—the actual number of review records available for each hospital—as a control in our regression because a greater number of positive or negative reviews for a hospital could have biased the overall Yelp-calculated patient hospital satisfaction ratings.

From the HCAHPS survey data, we used the unweighted patient experience of care dimension score for each hospital taken from the 2016 Hospital Compare database as our variable *patient experience of care*. We also used the recently introduced CMS' overall hospital star rating (CMS, 2016) to represent the *hospital quality performance*. CMS (2016) uses an elaborate methodology using 57 items from the HCAHPS survey grouped into seven subcategories, and is therefore appropriate for our use as a reflection of overall hospital quality.

RESULTS

All descriptive statistics including Pearson correlations among the variables in the study are indicated in Table 1. Our findings noted in Table 1, indicate a strong significant correlation between patient hospital satisfaction (collected from Yelp) and patient experience of care domain score (obtained from HCAHPS surveys), and also between patient hospital satisfaction and CMS overall hospital quality scores.

TABLE 1: Descriptive Statistics

		Mean	Std. Dev.	1	2
1	Patient Experience of Care (HCAHPS)	27.46	14.74		
2	Patient Hospital Satisfaction (Yelp)	2.82	0.60	0.51**	
3	Hospital Quality Performance (CMS)	2.90	0.95	0.55**	0.36**
4	Hospital Type	1	0.05		
5	Hospital Ownership	2.70	1.81		
6	Merged	1.93	0.26		
7	Number of Reviews	61.17	46.76		

Notes. n = 359 hospitals. **p < 0.01 (2-tailed). Model 1 (n = 359 reviews) with only controls; Model 2 with the predictor (n = 359 reviews). Pearson Correlations are reported. **p < 0.001.

Next, we performed an OLS regression with patient hospital satisfaction obtained from Yelp as the predictor and the patient experience of care domain scores obtained from HCAHPS as the criterion. Based on extant healthcare literature we introduced a few controls as these variables have been known to affect healthcare studies: the hospital type, the hospital ownership (Dobrzykowski, McFadden, & Vonderembse, 2016; Sharma, Chandrasekaran, Boyer, & McDermott, 2016; Wani & Malhotra, 2018) and number of online reviews received by the hospital (Ranard et al., 2016) and the dummy-coded merger indicator (1 = merged, 2 = not merged). None of the controls turned out to be significant. Our findings noted in Table 2, indicate a significant coefficient of patient hospital satisfaction predicting the patient experience of care thereby lending support to our hypothesis H1.

TABLE 2: OLS Regression Results

Variables / Model	**1**	**2**
Controls		
Hospital Type	-0.05	-0.05
Hospital Ownership	0.08	0.01
Merged	0.03	0.04
Number of Reviews	-0.02	-0.06
Predictors		
Patient Hospital Satisfaction	-	0.52**
R^2	0.01	0.27
Adjusted R^2	-0.00	0.26
ΔR^2		0.26

Notes*. Patient experience of care is the dependent variable. Model 1 (n = 359 reviews) with only controls; Model 2 with the predictor (n = 359 reviews). Standardized beta weights are reported. **p < 0.001.*

DISCUSSION

In this paper we investigated whether hospital satisfaction ratings obtained from social media sites such as Yelp can provide similar information to that obtained from HCAHPS ratings of hospitals. Support for our hypothesis demonstrates that freely-expressed patient hospital satisfaction ratings on social media sites such as Yelp are valuable. Our results show that the patient hospital satisfaction ratings from Yelp can predict the patient experience of care domain scores obtained through the HCAHPS surveys annually; and are also positively and significantly correlated to the overall hospital quality performance scores given by CMS. In so doing, we suggest that frequently checking hospital ratings from social media such as Yelp may provide hospital leaders an alternative to the long wait for the annually available HCAHPS survey results.

There are several disadvantages of the HCAHPS data collection process, and when the ratings are disclosed to each a participating hospital. First, each year on an average it takes quite a few months of actual survey on the part of the vendor to randomly collect the data from patients, collate and clean the data and get it ready for national level average calculations. Invariably, hospitals get to see their overall scores, and more specifically their patient experience of care scores almost a year from the time the data collection first began. Now, in the world of changing customer perceptions and even their feedback, a year may be too long a time. Yelp ratings, on the other hand, are freely available to anyone and they are constantly updated such that the scores are accurate at any time.

Second, hospitals enrolled in the value-based purchasing program (VBP), are now reimbursed their share of Medicare and Medicaid expenses or penalized based on their overall hospital score that is based on the HCAHPS ratings/data. Thus, hospitals leaders are generally apprehensive of the overall ratings published by CMS, which are all based on these HCAHPS data collection. Our study shows that Yelp ratings provide almost similar information, and hence could help hospital leaders know in advance, the overall CMS ratings that they are likely to receive. In that sense, Yelp ratings may help hospital leaders have a pre-emptive method to know their overall CMS ratings and could allow them to plan for their future steps better.

Like all studies, this study too has a few limitations. In this study we calculated the unweighted patient experience of care domain score for each hospital by summing the hospital’s HCAHPS base score (0-80) and HCAHPS consistency score (0-20). We demonstrate that patient hospital satisfaction ratings are positively and significantly correlated with overall hospital quality performance scores given by CMS. Nevertheless, we used the patient experience of care domain score as our criterion variable and not the overall CMS hospital quality performance score given by CMS each year because Yelp hospital ratings being satisfaction ratings of the patients/kin after all, we believe that Yelp ratings should be able to provide similar information as the patient experience score part of the data collected from HCAHPS. There are other dimensions too in the overall CMS hospital quality performance scores given each year but, in this paper, we argue that the patient experience domain scores being a very important part of the overall hospital score determined by CMS, hospital leaders would do well to keep a tab on Yelp scores and comments given by patients on free social media sites to gauge how their hospital is performing on the patient experience domain.

Another limitation of our study is that we have relied upon patient/kin comments expressed freely on only one social media website (Yelp). An avenue for future research involves collecting more patient comments from other popular social media websites such as *Healthgrades*, *Twitter* and *Facebook* as patients like other service customers are known to express their anguish on these social media sites too. We believe that patients freely express both their dissatisfaction and even happiness with on social media websites primarily because in the current US healthcare system, there are not too many avenues to express their comments and describe their experiences (Lagu et al., 2013). As we have demonstrated in this study, hospital leaders are able to obtain similar information that they routinely obtain from HCAHPS in a much easier manner and on a more frequent basis from social media websites.

Another future opportunity worth investigating is the difference in patient comments between merged hospitals and regular hospitals that have not consolidated. In our study, we had only a few hospitals that had merged during these 10 years; thus, our merged indicator as a control turned out to be insignificant. Future studies could specifically investigate if patient experiences improve after a merger or consolidation, as expressed by patients freely on social media for the hospital.

In closing, we would like to reiterate that patient hospital satisfaction ratings given freely on social media websites such as Yelp may provide a rich pool of up-to-date information about patient experiences in the hospital. As review sites continue to grow and proliferate in coming years, patient review data could provide an avenue for hospital leaders to know the areas for improvement at much greater frequency than waiting for the annual overall CMS scores obtained from CMS that are collected through the HCAHPS survey. In this manner, social media sites such as Yelp may help hospitals ultimately become more patient-centric.

REFERENCES

References available upon request

DECISION SCIENCES INSTITUTE

Reviewers Ranking as a Variable in Value Creation: Health Care Case

Ahmed Gomaa
University of Scranton
Email: ahmed.gomaa@scranton.edu

Satya Prattipati
University of Scranton
Email: Satya.Prattipati@scranton.edu

ABSTRACT

Users of healthcare services are providing online reviews of the services received from the healthcare providers. Positive patient reviews are likely to attract more patients and repeat visits. Health care providers are aware that a shift in the patient selection mechanism is taking place. One problem is the valuation of online reviews. This paper proposes a new algorithm that ranks each reviewer by weighting the type of feedback they provide for service providers to provide an economic valuation for the reviewers and their reviews. The algorithms calculate an Eigenvector and generate a weight for the reviews and the reviewers.

KEYWORDS: reviewers ranking, value creation, patient online reviews.

INTRODUCTION

Increasingly, more and more people are using Internet to select a service provider or a product. Online reviews have become a de-facto standard for all services providers. A majority of the consumers are expected to lookup a service provider or a product and find online reviews from prior users posted on sites like Facebook, google review, yelp, Amazon or even the providers web sites. Online reviews impact a purchase decision (Mudambi & Schuff, 2010); (Huang, Burtch, Hong, & Polma, 2016) and the absence of any online reviews will create a negative impression about the service provider or product.

This paper presents a new ranking algorithm for authors of health providers reviews to allow the economic valuation of the authors of the online reviews and their posted reviews. This research attempts to rank the reviewers of the healthcare providers, and their reviews, to assess the economic value of each reviewer. This paper models the value of the online reviewers of medical professionals by building on the **BiRank** ranking algorithm (He, Gao, Kan, & Wang, 2017)

The weighting is for the “Authority” or "contribution to," the online community, and “Hub” or "engagement with," the online community as well as the user’s interaction level with the reviews. The "Authority" variables are the summation of the A) number of reviews, and B) Stars rating left for the provider. The user’s interaction is the number of views and comments the reviews collected. Those variables help creating an "Authority Weight" for the reviewer. Another weight, the "Hub weight" is calculated based on the summation of A) the number of reactions and B) comments to other reviews. All variables are represented as a binary matrix to be able to calculate the corresponding Eigenvector and generate a weight for the reviews based on those variables. As shown in the demonstration results, the advantage of calculating the reviewer weight in this fashion is that it allows the reviewer weight to impact the reviews ranking even if that reviewer submitted review did not receive any positive feedback yet from the community. The algorithm weight the reviewer based on her prior interactions. The algorithm variables are selected based on two datasets. First, the data available from Medicare.gov which provide information about groups, individual physicians, and other clinicians currently enrolled in Medicare. Second, is from Yelp dataset, which provides a subset of business, along with the stored attributes.

LITERATURE REVIEW

In recent years, researchers started to show growing interest in the relationship between online reviews and the selection of healthcare providers. Some of this research is in Table 1:

Table 1: Research summary

Author	Year	Title	Finding
Luca and Vats (HBR)	2013	Digitizing Doctor Demand: The Impact of Online Reviews on Doctor Choice	Established a relationship between online and the healthcare provider selection
Patel and, Brombach (MISQ)	2016	Patient Engagement: Digital self-scheduling set to explode in healthcare over the next five years	By the end of 2019, 64% of patients will book appointments digitally, delivering $3.2 billion in value and a competitive boost for health systems
Goh, Gao, and Agarwal (MISQ)	2016	The Creation of Social Value: Can an Online Health Community Reduce Rural-Urban Health Disparities?	Rural areas generate value from consuming the content created by urban areas, thus establishing the fact that a value is transferable from one area to another due to the reviews.
Xu, Armony, and Ghose, (Management Science – under review)	2017	The Interplay between Online Reviews and Physician Demand: An Empirical Investigation"	Physician demand increase up to 7.24% from positive online reviews, the patient utility function increases up to 5.01% - thus establishing a monetary valuation of the online reviews
Lu and Rui (Management Science)	2017	Can We Trust Online Physician Ratings? Evidence from Cardiac Surgeons in Florida	Patients could trust the online ratings of physicians

Currently, there are several ranking algorithms that have proven to be effective. The common denominator across those algorithms is modeling the problem into a binary matrix of ones and zeros, and through an iterative process, the algorithm generates an eigenvector with the rank of the different components. PageRank developed by Google in 1998 (Page, Brin, Motwani, & Winograd, 1998), used it by modeling the web in a matrix of links, where one is an incoming link, and zero presents no link. The HITS algorithm (Kleinberg, 1999) presented both incoming links and outgoing links as matrixes and also used eigenvectors to rank the webpage. BiRank, developed in 2017, is an empirically superior algorithm (He, Gao, Kan, & Wang, 2017). It modeled as factors the time, current popularity of pages, and current influence of users. The algorithm is superior in predicting the level of popularity of a particular item with those three variables, time, item popularity, and user influence. In specific, BiRank algorithm has three hypotheses, 1) *Temporal Factor*. If an item has received many comments recently, it is more likely to be famous soon. More recent comments are a salient signal that more users focused on the item recently. 2) Item *Current Popularity*. If an item has already been widespread, it is likely to garner more views in the future. Popularity is partially affected by the existing visual interfaces of Web 2.0 systems: the more views an item has, the more likely it will be promoted to users.3) *User Social Influence*. If the users commenting on an item are more influential than uncommenting users, the item is more likely to receive more views in the future. It is important to note that BiRank is drastically different from PageRank and HITS. BiRank model users and their interactions and how that impact the popularity of the item, while the other two models model the webpage links. The fundamental difference lies within the information flow. As a result, PageRank and HITS will always rank popular websites like espn.com, CNN.com or FOXNews.com highly, while BiRank, given the influence of a celebrity like Michele Obama, her new page will be ranked high regardless of the website she decides to use too.

The BiRank algorithm includes the following variables/ equations.

$$w_{ij} = \delta^{a(t_0 - t_{ij}) + b} \tag{1}$$

The Weight w from user i to object *j* is a function of the time it was provided. In other words, the older the comment is, the less valuable it is.

$$u_i^0 = \frac{\log(1 + g_i)}{\sum_{k=1}^{|U|} \log(1 + g_k)} \tag{2}$$

$$p_j^0 = \frac{\log v_j}{\sum_{k=1}^{|P|} \log v_k} \tag{3}$$

u stands for user i, and the user influence is a function of the g, which is the number of a friend that user has. The *p* stands for a post, page or object, and its value is a function of *v* which is the number of comments that page or object received.

Figure1: BiRank Algorithm (He, Gao, Kan, & Wang, 2017)

Algorithm 1. The Iterative BiRank Algorithm

Input: Weight matrix W, query vector $\mathbf{p}^0, \mathbf{u}^0$, and hyper-parameters α, β;

Output: Ranking vectors $\mathbf{p}, \mathbf{u}$;

Symmetrically normalize W:$S = D_u^{-\frac{1}{2}} W D_p^{-\frac{1}{2}}$;

Randomly initialize $\mathbf{p}$ and $\mathbf{u}$;

while Stopping criteria is not met **do**

$\quad \mathbf{p} \leftarrow \alpha S^T \mathbf{u} + (1-\alpha)\mathbf{p}^0$;

$\quad \mathbf{u} \leftarrow \beta S \mathbf{p} + (1-\beta)\mathbf{u}^0$;

end

return $\mathbf{p}$ and $\mathbf{u}$

BiRank Algorithm in figure 1 presents how to calculate the rank of a particular page; the input is the weight matrix, which represents the comment and likes on a specific object, an initial random rank for the post or object *p* and an initial random rank for the user. The algorithm continues its iterations until *p* and *u* converges that is *p*, and *u* stop changing with further iterations. At this point, the algorithm finalizes the rank for the objects and the users.

ACCOUNTING FOR PATIENTS SPECIFIC CONDITIONS WITHIN THE BIRANK ALGORITHM

The paper applies “patients review” to the BiRank algorithm. First, the hypotheses are adjusted as follows:

H1. *Temporal Factor*: If a review received many comments recently, the author of the review is more likely to be popular in the short term. More recent comments are a salient signal that more Patients focused on the recent reviews.
H2. *User’s Social Influence*: If the users reviewing the physicians are more influential than others, the Physician is more likely to receive more views in the future. Interaction includes commenting, liking another comment, or commenting on another comment.
H3. *Item Current Popularity*: If a review is already viewed as helpful, it is likely to garner more patients to trust it in the future.

It is important to note that the variables and the equations remain the same, but the semantics of the variables are different. For instance, in H1, the algorithm accounts for the time of ALL interactions, not only the comments. Furthermore, in H2, the user influence is not only by friends but also by the level of interaction and engagement with others within the community. Also, in H3, the object represents the health provider review itself as a step to evaluate the reviewer value.

ALGORITHM

We need to consider the effect of user influence on ranking physicians. We define a user influence matrix U whose elements indicate how many times a user has contributed or engaged with a review. This should contribute to the reputation score of the review.

As mentioned before, the *p* represents the post, in this paper context, the physician current popularity, and *u* represent the user social influence. In step 5, the customized algorithm represents u^0 with the equation $u^o = c^0 x e^0$

$$\mathbf{p} \leftarrow \alpha S^T \mathbf{u} + (1-\alpha)\mathbf{p}^0; \tag{4}$$
$$\mathbf{u} \leftarrow \beta S \mathbf{p} + (1-\beta)\mathbf{u}^0 \tag{5}$$

Here the initial value the user influence is not her friends as in BiRank, but instead the importance of the user contribution and engagement to the system. The new algorithm starts assuming that all users are fully contributing and fully engaged with a perfect score. The new algorithm multiplies the transpose of both vectors to get the initial U value

$c^0 = [1.1.1.....1]^T$ (6)
vector indicates the score of contribution given to the community, number of check-ins, and stars

$e^0 = [1.1.1.....1]^T$ (7)
vector indicates the number of engagements: (likes) and comments back

To get the *U* value, the customized algorithm uses the generated *P* in step 4 and plug in the values in equation 5.

The *U* score in step 5 is calculated based on the number of contributions, and Engagement, both presented as matrixes

C_{ij}, number of contributions: reviews, number of check-ins, and number of stars
E_{ij} number of engagements: (likes) and comments back
$u^0 = C^T c + E^T e$ (8)

WEIGHTS ON CONTRIBUTIONS AND ENGAGEMENT

The "Authority weight" represents Authors of online reviews contributions and is based on
A) number of reviews,
B) Stars rating left for the provider.

The “Hub weight” represents Authors of online reviews engagement and is calculated based on
A) the number of reactions and
B) comments to other reviews.
Both variables are creating the reviewer influence level

So far, the paper modeled the authors rank by using the amount of engagement and contribution to the community instead of using the number of friends as a point of reference. The resulting ranked authors of reviews lead to ranking reviews. The algorithm output will be ranked reviews, based on the reviewers’ level of engagement and contribution. The reviews are first classified into eight separate aspects according to Medicare classification, namely,

1. Getting timely care, appointments, and information.
2. Between-visit communication.
3. Attention to patient medication cost.

4. How well clinicians communicate.
5. Patients' rating of clinicians.
6. Health promotion and education.
7. Courteous and helpful office staff.
8. Clinicians are working together for patient care.

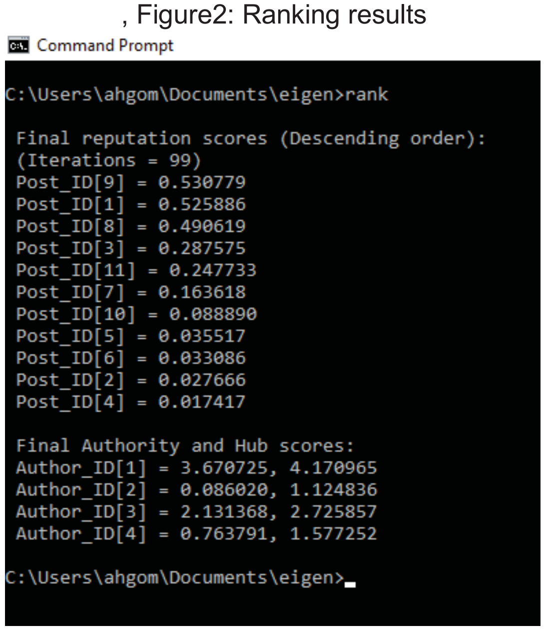
The paper ranks the reviewers based on the aspects they are reviewing a physician. Natural language processing allows for the classification of each review in one or more aspect, in addition to any other aspects presented in the Yelp reviews. In other words, for each identified aspect, related reviews are grouped and ranked based on the review popularity, reviewers influence and the temporal dimension.

IMPLEMENTATION DISCUSSION

For demonstration purposes, the implementation assumes the same time of review in one aspect. In the algorithm implementation, assume there are 4 reviewers addressing one identified aspect, which is attention to patient medication cost. All reviewers provided their reviews at the same time period. The four reviewers provided 11 reviews, and are interacting with each other's reviews. The Appendix presents the data used to generate the ranking of the reviews along with each reviewer rank as a hub and an authority to reflect the reviewer level of engagement and contribution to the community. The algorithm output shows that review 9, written by Author 3, is the most relevant review to the community, given the author of the review rank in terms of contribution and engagement. Author 1 is the most influential reviewer within the community, as she is the one the majority of the community interact with her posts. It is worth noting that post 1, written by Author 1 has a higher ranking than post 8 written by Author 3 due to the difference in the author authority and hub rank.

Accordingly, the algorithm allows for ranking reviews based on the reviewer influence, and the item popularity. The time of the review impacts the final ranking. Figure 2 shows the results of the algorithm. The results converge after 99 iterations. The levels of contributions, engagement and user interactions is presented in the Appendix.

, Figure2: Ranking results



ECONOMIC VALUE

Recent research suggests that there is an economic value from the online reviews of healthcare providers. The objective of exchanging the economic value with an online health community economy is to disseminate knowledge from the ecosystem point of view. In an attempt to formalize the ecosystem, it is essential to articulate the reviewers' economy variables in exchange. The reader of the reviews is trading their time and attention, in return, they are receiving the collective wisdom of the masses, and as a byproduct, the service providers are increasing their utility function, and the reviews readers book appointments with them more than the non-reviewed service providers. This paper uses the reviewer/patient engagement and contribution, and the health care provider reviews popularity to create a value function, where the patient generates value based on his effort (online reviews). According to the literature, the provider benefits from such engagement up to 7%, and the other system users benefit up to 5%. If a reservation system generates $250,000 per day, the benefits for the system to have a higher booking due to relevant reviews would be 7% daily or $17,500 per day. It might make sense to provide monetary compensation for the reviewers as a function of the value added they provide

$$ValueCreation_i^t = \frac{f(u_i^{(t)})}{\sum_{n \in all} f(u_n^{(t)})} \text{ x } income^t \tag{9}$$

The Value Creation is reviewer $\boldsymbol{i}$'s value creation in time $\boldsymbol{t}$. Income is the total income that may be distributed by the system to the community in time $\boldsymbol{t}$. $\boldsymbol{u}$ is the user influence score in time $\boldsymbol{t}$. $\boldsymbol{f}(.)$ is a reward function that controls the effect of change.

CONCLUSION

The paper presents a model to help incentivize online contributors in the form of reviews, where more compensation is delivered when more meaningful/sustainable contributions are provided. First, the paper identified an empirically tested superior ranking algorithm and modeled the patients' characteristics using the same methodology in the identified algorithm. The new algorithm defines the user influence as a result of the user contribution and engagement in addition to the importance of the reviews as viewed by the community. The algorithm implementation demonstrates the impact of the reviewer authority on the rank of the review. Second, the paper presents a value creation function based on the user influence, where each user is compensated based on her rank, which is derived from her engagement and contribution to the community, and the popularity of the physicians she is reviewing. To incentivize commentators, a financial model needs to pay the reviewer based on their impact. Future work includes automating the reviewers' compensation by presenting the knowledge created and disseminated in a blockchain ledger. Knowledge transactions may be recorded on a public ledger, similar to blockchain. The more usable the reviewer knowledge, the more transaction it generates. According to the paper model of value creation, it leads to more booking to the system in general, even if negative reviews exist. This is because the compensation is based on the quality and influence of the review, not the sentiment of the reviewers. An additional area of future research is to combine the temporal dimension with the spatial dimension of the patients using the build in mobile sensors to validate the physician popularity based on the number of signals detected in large medical buildings.

APPENDIX

Engagement			
Author_ID	Post_ID	Linked_Post	Linked_author
1	1	9	3
1	3	10	4
1	3	11	4
2	4	6	2
2	5	7	3
2	6	8	3
3	7	3	1
3	8	4	2
3	9	5	2
4	10	1	1

Contribution	
Author_ID (Reviewer)	Post_ID (Review)
1	1
1	2
1	3
2	4
2	5
2	6
3	7
3	8
3	9
4	10
4	11

User interaction			
Author_ID (Reviewer)	Post_ID (Review)	Views	Comments
1	1	4	1
1	3	2	1
1	1	1	0
2	4	1	1
2	5	1	1
2	6	1	0
3	7	1	1
3	8	10	1
3	9	11	1
4	10	1	1
1	11	1	1

REFERENCES

Goh, J. M., Gao, G. & Agarwal, R. (2016). The Creation of Social Value: Can an Online Health Community Reduce Rural-Urban Health Disparities? *MIS Quarterly, 40*(1), pp.247-263.

He, X., Gao, M., Kan, M.-Y., & Wang, D. (2017). BiRank: Towards Ranking on Bipartite Graphs. *IEEE Transactions on Knowledge and Data Engineering, 29*(1), pp 57 - 71.

Huang, N., Burtch, G., Hong, Y., & Polma, E. (2016). Effects of multiple psychological distances on consumer evaluation: A field study of online reviews. *Journal of Consumer Psychology, 26*(4), pp 474 – 482.

Kleinberg, J. M. (1999). Hubs, authorities, and communities. *ACM Computing Surveys, 31*(4), Article 5.

Lu, S. F., & Rui, a. H. (2017). Can We Trust Online Physician Ratings? Evidence from Cardiac Surgeons in Florida. *Management Science, 64*(6), 2557–2573.

Mudambi, S. M., & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly, 34*(1), pp: 185-200.

Page, L., Brin, S., Motwani, R., & Winograd, a. T. (1998). *The PageRank citation ranking: Bringing order to the web.* Stanford InfoLab.

Patel, D., & Brombach, M. (2016). *Insight Driven Health: Patient Engagement: Digital self-scheduling set to explode in healthcare over the next five years.* Accenture.

Xu, Y., Armony, M., & and Ghose, A. (2017). The Interplay between Online Reviews and Physician Demand: An Empirical Investigation. *Management Science - under review*.

DECISION SCIENCE INSTITUTE

Managing Product Complexity to Create More Effective Innovation Outcomes: A Comprehensive Examination of Coordination Mechanisms

Ryan Skiver
University of Wisconsin Oshkosh
Email: Skiverr@uwosh.edu

ABSTRACT

Literature suggests that product complexity has the potential to lead to greater profits if the complexity is managed effectively (Kekre & Srinivasan, 1990). Yet firms with higher levels of product complexity experience profit margins that are on average 3 percent lower than firms that do not have high product complexity (Hoole, 2006). Coordination Mechanisms should be put into place to manage the strain caused by product complexity. The results showed product complexity in two dimensions (multiplicity and relatedness) could create better innovative outcomes (Blue Ocean and Red Ocean) with coordination mechanisms in place.

KEYWORDS: Product complexity, coordination mechanisms, Blue ocean innovation, Red ocean innovation, survey research

INTRODUCTION

Product complexity has been increasing over the past couple decades due to customer demand for new technologies and customized products. Firms must adopt these differentiated products to continue thriving in today's customer-driven market. Literature suggests that added product complexity has the potential to lead to greater profits if the complexity is managed effectively (Kekre & Srinivasan, 1990; Meeker, Parikh, & Jhaveri, 2009; Meyer & Mugge, 2001). Yet firms with higher levels of product complexity experience profit margins that are on average three percent lower than firms that do not have high product complexity (Bozarth, Warsing, Flynn, & Flynn, 2009; Hoole, 2006). Firms with greater product complexity typically do have higher revenue growth, but have lower margins (profits). A key imperative is to find ways of more efficiently managing complexity. Literature has shown that managing product complexity effectively poses great difficulty in coordinating complex systems and aligning product complexity with strategic goals (Jacobs & Swink, 2011). This research was structured to find management techniques in product complex environments that would render gains in profits instead of profit losses.

LITERATURE REVIEW

Product complexity (and factors that cause complexity), strategy, and innovation have all been studied vastly in literature. This research put all of the areas together to create a model of successful innovation in a high complexity environment. There is a growing need to create multiple products that cater to individual needs (Jacobs, 2007; Lawton, 2007). However, there is a growing paradox associated with product complexity. Product complexity can create greater profits if the complexity is managed effectively (Meeker et al., 2009; Meyer & Mugge, 2001). At the same time, higher profits are not seen in highly complex industries (Bozarth et al., 2009; Hoole, 2006). This is due to a lack of strategic management of the complexity issues (Hoole, 2006). This research examined this paradox in detail and explored strategies using coordination mechanisms to alleviate the strain from complexity and to create more successful outcomes for firms in complex environments. Concurrently, innovation is necessary for a firm's survival (Steencamp, Hofstede, & Wedel, 1999). This research explores the link between complexity and innovation in the following sections and shows that innovation is an outcome of product complexity. This research also explores the literature on coordination mechanisms and combines this stream of literature with highly complex industry and innovation outcomes. The research then shows that certain coordination mechanisms can relieve the strain from complexity and at the same time create better outcomes in terms of both new and existing market innovation. Complexity, strategy, and innovation relationships are shown in Table 1 below.

Table 1: Complexity, Strategy, and Innovation: A Vital Relationship

Key Business Challenges	**Key Research Issues**	**Reference**
Complexity Challenges: Increasingly firms face complexity challenges that arise from diverse aspects including technological, environmental, and social factors. Firms need to understand these challenges and focus on how they can manage them successfully.	(1) Assess the source of complexity at firm and industry level; (2) Understand how firms manage the increasing level of complexity to add values to customers and to stay competitive.	Lawton, 2007; Meeker et al., 2009; Meyer & Mugge, 2001
Strategic Responses: A single firm cannot handle complexity challenges alone. Complexity challenges require organization-wide, purpose-driven efforts over a long-term horizon.	(1) Firms strategically engage employees, customers, and suppliers to create more effective outcomes in a complex environment; (2) Firms operationally involve employees, customers, and suppliers to create more effective innovation results.	Ashkenas et al., 2002; Bozarth et al., 2009
Innovation Outcomes: Complex problems and issues require innovative problem definition and solving and deploying processes that involve effective translation of challenges into valuable outcomes.	(1) Complexity challenges require new and different approaches (i.e., innovation); (2) Innovation is the lifeblood of the firm; (3) Outstanding innovation is often the result of adopting new perspectives, implementing different processes, and pursuing high value-reward outcomes.	Pellissier, 2012; Sivadas & Dwyer, 2000; Steencamp et al., 1999

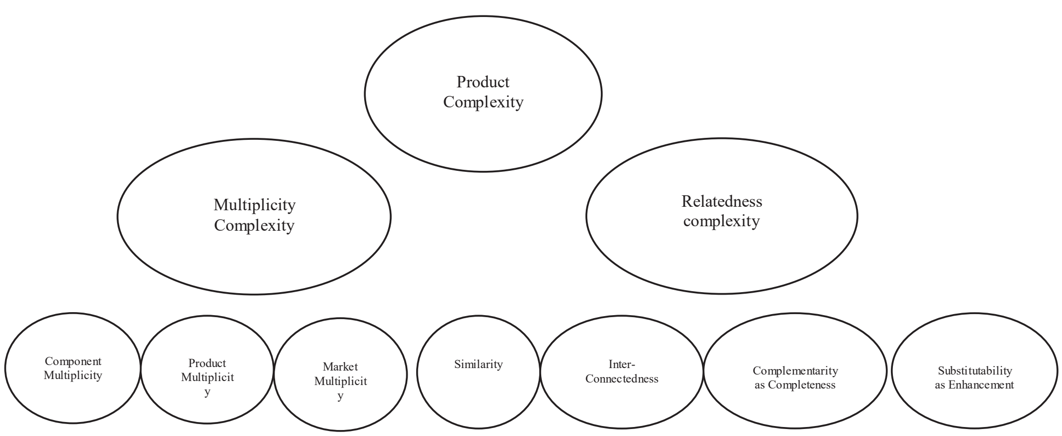
Table 1: Complexity challenges are significant and strategic responses are required. Firms that manage complexity well also show high levels of innovation outcome success.

This review will continue to explain the following three main areas of study: product complexity, the product complexity link to innovation outcomes, and coordination mechanisms to reduce strain caused by product complexity. This review provides a better understanding of the subject area and the reason for linking all the concepts together with theory. To understand what drives product complexity, we must first understand complexity challenges and what innovation is to discover how product complexity and innovation connect.

Complexity challenges. Increasingly, firms face complexity challenges that arise from diverse aspects including technological, environmental, and social factors (Lawton, 2007). Firms need to understand these challenges and to focus on how they can manage successfully (Meeker et al., 2009; Meyer & Mugge, 2001). This research aimed to accomplish two outcomes: (1) Explore the relationship between product complexity and innovation outcomes, and (2) Understand how firms manage the increasing level of complexity to add value to customers and to stay competitive with innovative products. To accomplish these three goals we must first understand the relationship between product complexity and innovation outcomes.

Product complexity dimensions. This research on product complexity built on the workings of Jacobs (2007) by expanding on the areas of multiplicity and relatedness. A greater understanding was obtained by dividing multiplicity into three dimensions: component, product, and market. This is because products can cause complexity for a firm due to many different products, many different components inside the products, and many different markets. By dividing multiplicity up into categories, this research attempted to gain a better understanding of the complexity of these three areas. Although the three dimensions of relatedness (similarity, interconnectedness, and complementarity) all fit the purpose of this research, a clearer understanding of complexity could be gained only if complementarity was divided into separate categories such as complementarity as completeness and substitutability as enhancement. In the previous literature complementarity complexity was all linked into one dimension. This research has found that complementarity is add-on product; however, add-on products can either be necessary for a product to function or unnecessary to enhance a product. For this reason to clarify this issues complementarity was divided into complementarity as completeness and substitutability as enhancement. Product complexity is quite complicated but by looking at product complexity in terms of the dimensions of multiplicity and relatedness, complexity can be captured in a specific, researchable way. The dimensions of product complexity are show below in Figure 1.

Figure 1: Dimensions of Product Complexity



Strategic coordination: a crucial role for coordination. Complexity challenges require responses from firms. Coordination is "the organization of the different elements of a complex body or activity so as to enable them to work together effectively" (Webster, 1964 p. 502). Therefore, coordination mechanisms are defined as the sustainable system-wide efforts that coordinate diverse business processes, functions, and people for the purpose of product development. The role of coordination mechanisms is to translate organizational resources into new ways of problem solving and value creation for customers in the form of products/services that increase value to the customers. Involvement and collaboration of employees, customers, and suppliers have been shown to have many advantages for the firm involved (Ashkenas et al., 2002; Cross, Yan, & Louis, 2000; Devanna & Tichy, 1990; Kerr & Ulrich, 1995). This collaboration has been studied extensively in the innovation literature (Powell, Koput, & Smith-Doerr, 1996; Shan, Walker, & Kogut, 1994). These authors have performed empirical tests and have shown that collaboration between: functional units, hierarchical levels, suppliers, and customers was positively related to new product development success both in financial performance and output of new products (Powell et al., 1996; Shan et al., 1994). The financial aspect is due to the close feedback from consumers so products that are necessary are being created. As the consumer demands greater complexity and newer products, coordination mechanisms will need to be in place to allow for more successful product creation. The success of new products is due to the collaboration between functional units, hierarchical levels, suppliers, and customers, allowing for products that are desired in the market at faster speeds and without technical problems.

Boundaryless organizations. **"**Boundaryless organizations" is the strategy of forming coordination mechanisms between firms and employees, customers, and suppliers. These coordination mechanisms allow for increased speed, clearer information, better decisions, etc. (Ashkenas et al., 2002). The idea of "boundaryless organizations" was created in 1988 inside General Electric by former CEO Jack Welch. The idea was to reduce bureaucracy, shorten cycle times, and increase capabilities for change (Ashkenas et al., 2002). A group of consultants that included Ashkenas, Jick, Ulrich, and Kerr created a blueprint for the term now known as boundaryless organizations (first coined in 1990). By 1992 global linkages were added and Kerr, Jick, and Ulrich began adding details to the complete the concept. They created a paradigm shift from describing successful businesses using old factors of size, role clarity, specialization, and control to using new success factors of speed, flexibility, integration, and innovation. The final concept consisted of four areas (included in Table 3 below) [modified from (Ashkenas et al., 2002).

Table 2: Types of Boundaries

Boundary Type	**Brief Summary**	**Description**	**Reference**
Vertical Boundaries	The various layers between the floor and ceiling of the organization.	Separate people by hierarchical levels, title, status, and rank.	Ashkenas et al., 2002; Devanna & Tichy, 1990
Horizontal Boundaries	The internal walls.	Separate people in organizations by function, business unit, product group, or division.	Ashkenas et al., 2002; Devanna & Tichy, 1990
External Boundaries	The external walls.	Divide companies from their suppliers,	Ashkenas et al., 2002; Cross et

		customers, communities, and other external constituencies.	al., 2000; Kerr & Ulrich, 1995
Geographic Boundaries	The cultural walls.	Include aspects of the other three but are applied over time and space, often across different cultures.	Ashkenas et al., 2002; Cross et al., 2000

Ashkenas et al. (2002) was amended to create Table 2 above. These changes were made to encompass the scope of product complexity and to focus on areas of product complexity and innovation to create working coordination mechanisms that could be used within a firm. For the purpose of this research, vertical, horizontal, upstream, and downstream boundaries covered the necessary scope of the coordination needed in product complex environments. Also because the area of focus for this research was primarily national, the instances of geographic boundaries from Ashkenas et al. (2002) that dealt with coordination between countries were excluded or changed to the appropriate coordination area vertical, horizontal, upstream, or downstream.

Innovation outcomes. Innovations and new product development are the lifeblood of any firm (Steencamp et al., 1999). To remain successful over long periods, firms must be ambidextrous in creating new market products (products that create new markets) through modular and radical innovation, and must upgrade existing market products (products that are new but do not form new markets) through architectural and incremental innovations. The management of radical and modular (new market innovation) poses a unique set of challenges for managers. It is a long and investment intensive process, marked by set-backs and unpleasant surprises, with no guarantee of success (Lynn, Morone, & Paulson, 1996). Success in the competitive environment involves exploitation of a firm's existing capabilities, whereas survival in a complex environment involves exploration for new capabilities at the same time (Levinthal & March, 1993; March, 1991).

With rapid increases in technology over the past thirty years, the environment has becoming more and more complex (Lawton, 2007). This brings an added importance to the exploration or new technology factor. In the literature there have been two main types of exploration of new market products—radical and modular innovation (Henderson & Clark, 1990). Yet, firms still experience significant failure rates when trying to implement new technologies (Sivadas & Dwyer, 2000). This research continues below with a review of the literature on innovation to put a better perspective on innovation and its relationship to product complexity.

This research viewed product complexity in terms of multiplicity and relatedness complexity. Product complexity and innovation go hand in hand (Pellissier, 2012). Multiplicity has three dimensions (component, product, and market) has four dimensions (simplicity, interconnectedness, complementarity as completeness, and substitutability as enhancement) discussed in the previous section that attribute to product complexity. These dimensions cause firms to introduce new products and thus create innovation in the firm. To fully understand product complexity we must understand innovation and how innovation is created from product complexity. These will be explained in the upcoming sections, but first a clear understanding of innovation must be developed from the literature to understand this relationship and to discover ways to manage complexity properly.

Contingency theory. Contingency theory was used to describe the different strategies needed for different levels of product complexity and why successful product complexity requires various coordination mechanisms. To understand the linkages of product complexity to coordination mechanisms, it was essential to investigate contingency theory further. Contingency theory links the characteristics of the organization (i.e., technology and strategy) to complexity, where different strategies are necessary to control the outcome of a changing system (Donaldson, 2001). This research explored strategies to deal with one type of system, a product complex system. Burns and Stalker (1961) stated that different kinds of management systems are appropriate to different kinds of technical environments. In other words, the same system cannot be used universally. Instead, technological complexity forces the use of different systems, depending on the conditions. Chandler (1962) stated that structure follows strategy in organizations. The firm creates a strategy based on what they hope to accomplish. The strategy is based on objectives, long-term goals, and allocation of resources. Different strategies create different needs; therefore, organizational structure will evolve to accommodate those needs.

A firm focuses on creating products that it controls and does not focus on products that are impossible to control (Lawrence & Lorsch, 1967). Contingency theory literature has shown that a fit between internal structure and the external environment must be obtained for maximized performance (Lawrence & Lorsch, 1967). If the strategy were changed due to external conditions, this would allow for better performance. Galbraith (1977) discussed contingency theory from the perspective of information processing in organizations by developing an information-processing model to explain the relationship between the information processing needs of a firm and the structural mechanisms that can address those needs. He argued that the ability of an organization to coordinate its activities successfully depends on how effectively and efficiently it can process the amount of information needed to be processed. This effective processing is done by reducing the need for information processing or by increasing the capacity to process the information. In other words, an information processing mechanism, or capability, should match the firm's information processing needs. Mintzberg (1981) stated that organizational success is matching or fitting the parts and characteristics of organizational structures to one another. This research focused on fitting strategy to practices based on varying levels of product complexity. A further explanation of contingency theory is provided below in Table 3. Table 3 demonstrates the key findings from contingency theory literature as applied in this research.

Table 3: Contingency Theory Time Frame

Key findings	**Literature Source**
Structures follow strategy in organizations	Burns & Stalker, 1961
Different kinds of management systems are appropriate to different kinds of technical environments	Chandler, 1962
Organization must establish a "fit" between its internal structure and its external environment	Lawrence & Lorsch, 1967
The ability of an organization to successfully coordinate its activities depends on how effectively and efficiently it can process the amount of information needed to be processed	Galbraith, 1977

The key to organizational success is matching or fitting the parts and characteristics of organizational structures to one another	Mintzberg, 1981
Appropriate fit between strategy and environment results in superior performance	Miller, 1988

Systems theory. In organizations systems consist of people, structures, and processes that work together to make an organization healthy or unhealthy (Ackoff, 1978; Bailey, 1994; Bánáthy, 1996; Bausch, 2001; Buckley, 1967; Capra, 1997). Systems thinking has been defined as an approach to problem solving, by viewing problems as parts of an overall system (Davidson, 1983; François, 1999; Gorelik, 1975). Firms are not judged solely by being in the market. Instead, they are judged by what they achieve in terms of outcomes (François, 1999). One firm alone cannot handle complexity; therefore, system resources (employees, customers, and suppliers) are needed to create the most successful outcomes (Davidson, 1983; Gorelik, 1975). Systems thinking is not one simple concept, but rather a set of habits or mechanisms within a framework that is based on the belief that the parts of a system can best be understood in the context of relationships with each other and with other systems, rather than in isolation (Checkland, 1981, 1997; Churchman, 1968, 1971). Product complexity challenges can be handled by an organizational system network to allow an organization to achieve improved outcomes.

What is the rationale for multiplicity and relatedness? Complexity theory details that single items are not complex; it is the number of different items or the interaction of the items that form complexity (Anderson, 1999; Brown & Eisenhardt, 1997). Because of this complexity, there is a great importance in understanding the multiplicity and the relatedness of products.

Social exchange theory. Social exchange theory is a social, psychological, and sociological perspective that explains social change and stability as a process of negotiated exchanges between parties (Befu, 1977; Cook & Emerson, 1978; Ekeh, 1974). Social exchange theory posits that all human relationships are formed by the use of a subjective cost-benefit analysis and the comparison of alternatives (Miller, 2005; Roloff, 1981; West & Turner, 2007; Zafirovski, 2005). Another part of social exchange theory is structuralism, which describes the benefits of coordination or why employees, customers, and suppliers would want to coordinate to form a network.

Why are coordination mechanisms the key strategic response? Systems theory and social exchange theory have demonstrated that one firm alone cannot handle complexity; therefore, resource coordination (between employees, customers, and suppliers) in the entire system is needed (Checkland, 1981, 1997; Churchman, 1968, 1971). Systems theory also explains why coordination mechanisms are needed due to the valuable outcomes attained by all parties involved in network coordination (Miller, 2005; West & Turner, 2007). Thus, how firms coordinate employees, customers, and suppliers is important strategically to success in a firm.

Product multiplicity complexity. Jacobs (2007) defined multiplicity as enumeration of items. This research extended that definition to ask the question, “the enumeration of what items?” The literature has shown that multiplicity exists not only in the sheer number of skus but also in the number of components within the product, the number of products themselves (skus), and the number of markets in which the firm participates (Baldwin & Clark, 2000; Daft, 1983; Gupta & Krishnan, 1999; Novak & Eppinger, 2001; Ramdas, 2003; Scott, 1992).

Component multiplicity. Component multiplicity is the enumeration of components within a portfolio (Gupta & Krishnan, 1999; Ramdas, 2003). Different products contain different numbers of components. Books are for the most part very low in component complexity, whereas cell phones and computers may contain thousands of component parts and thus are very component complex. Firms must deal with this type of complexity to create the necessary products for the consumer.

Product multiplicity. Product multiplicity is the enumeration of products within a portfolio (Baldwin & Clark, 2000; Novak & Eppinger, 2001). A small firm may specialize in one product, but this is usually not the case for long. Soon there are new generations of the product or product spinoffs to generate more revenue for the firm. Markets for most products are not infinite, so firms search for ways to gain valuable market share. This is done by generating new products.

Market multiplicity. Market multiplicity is the enumeration of markets within a portfolio (Daft, 1983; Scott, 1992). As a firm grows larger, it looks outside its current market to new markets to earn revenue. The more markets it enters, the more complex its environment becomes and the more difficult to control the need for new products becomes. This adds to the product complexity of a firm.

Product relatedness complexity. Relatedness of elements is formed by three subcomponents identified from the research on complexity. They are similarity, interconnectedness, and complementarity (Jacobs, 2007). Similarity is the sharing of product architectures in the terms of a platform or module (Fixson, 2005; Novak & Eppinger, 2001). Interconnectedness is the sharing or interaction of components to make the product function when one component needs another component to function (Griffin, 1997a, 1997b; Kaski & Heikkila, 2002). The final dimension is complementarity, which occurs when one product requires another aftermarket product (Jacobs, 2007; Tatikonda & Stock, 2003). Examples are a Blu-ray player and a Blu-ray disc, mp4 and digital audio/video, etc. The original product requires additional products for the original product to function.

Similarity. Similarity is the extent to which a product shares characteristics such as part geometries or components, fulfilling the same strategic role in the portfolio as a prior product but allowing for customization through replacement of components or modules. (Fixson, 2005; Novak & Eppinger, 2001). Customization includes modifying the appearance of presentation objects or editing their attribute values by selecting among a set of predefined configuration options (Morch, 1997). Customization can be accomplished by the company during production or after the initial purchase by the consumer. For example, the consumer may purchase a car with power windows or a premium sound system. The car itself has not changed, yet the modules of the power windows or a premium sound system give the consumers more options and differentiation than found in the original product (the car). In this case, product differentiation exists, but the differentiation is easier to produce because different products use similar platforms and modules.

Interconnectedness. Interconnectedness is a connection via an interface (Jacobs, 2007). When component parts require additional component parts to communicate or to fulfill their role in a product, there will be high interconnectedness (Griffin, 1997a, 1997b; Kaski & Heikkila, 2002). There are many products that have connected components, where if one fails, the entire unit will fail. This interconnectedness makes the product much more complex by relying on every component to function without failure.

Complementarity as completeness. Complexity as completeness is the extent to which other aftermarket products need to be purchased by the end user to obtain the intended use of the original purchased product (Jacobs, 2007; Tatikonda & Stock, 2003). One product adds value to another that otherwise would be less valuable. An example of this is computer software. It is dependent on a computer to function. The way to test for complementarity as completeness is to determine if the original product works out of the box or if it is dependent on a second product purchase to work as its designers intended. Complementarity as completeness requires a great deal of technology for a product to function.

Substitutability as enhancement. Substitutability as enhancement is the extent to which other aftermarket products can be purchased by the end user to create enhanced benefits from the original purchased product. One product adds value to another product by enhancing the original product, making the original product more valuable to the consumer. An example of this is a cell phone case. A consumer may purchase a case and screen protector for a smart phone through the retailer. This purchase adds to the complexity of the product by adding on more components to the original product. The cell phone could function fine on its own without the case; however, the case provides extra features such as waterproofing and damage protection, which create greater value for the original product. The way to test for substitutability as enhancement is to determine if the original product works out of the box and could later be enhanced by extra component purchases.

Coordination mechanisms. Coordination is, “the organization of the different elements of a complex body or activity so as to enable them to work together effectively” (Webster, 1964, p. 502). We therefore define coordination mechanisms as the sustainable system-wide efforts that coordinate diverse business processes, functions, and people for the purpose of product development. Involvement and collaboration of the employees, customers, and suppliers provide many advantages for the firm involved (Ashkenas et al., 2002; Cross et al., 2000; Devanna & Tichy, 1990; Kerr & Ulrich, 1995; Roh, Hong, & Park, 2008). This collaboration of networks has been studied extensively in the complexity and innovation literature (Powell et al., 1996; Shan et al., 1994). There have been empirical tests discussed in the previous literature that showed the important relationship between networks and new product development success in financial performance and output of new products (Powell et al., 1996; Shan et al., 1994). The financial aspect is due to the close feedback from consumers, which leads to the firm creating only products that are needed and demanded by the consumer. As the consumer demands greater complexity and newer products, coordination mechanisms will need to be in place to allow for more successful product creation. The output of new product creation is due to the collaboration between functional units, allowing for the product to be created without serious setbacks (Hong, Doll, Nahm, & Li, 2004; Hong et al., 2011).

Vertical coordination. Vertical coordination is bringing together employees and the firm as one unit (Ashkenas et al., 2002; Devanna & Tichy, 1990). This collaboration creates a free flow of information and tears down the bureaucracy. Boundaryless organizations focus more on who has the useful ideas than on who has rank. The boundaryless concept also allows decisions to be made quickly, increases information sharing, increases innovation research, and increases speed and flexibility (Ashkenas et al., 2002). The idea is that when employees are free to experiment and to think for themselves, they will be more creative and the company will create new, innovative products. Firms with fluid job descriptions, loose organization charts, high communication, and few rules may be conducive to innovation because they free developers from constraints, allowing them to change flexibly and to create novel ideas (March, 1981; Peters, 1994).

Horizontal coordination. Horizontal coordination is bringing together the firm's functional units as one unit (Ashkenas et al., 2002; Devanna & Tichy, 1990). This collaboration creates a free flow of information among departments. Boundaryless organizations focus more on who has the useful ideas than on who has authority. The boundaryless concept also allows decisions to be made quickly, increases information sharing, increases innovation research, and increases speed and flexibility (Ashkenas et al., 2002). Firms with fluid job descriptions, loose organization charts, high communication, and few rules may be conducive to innovation because they free developers from constraints, allowing them to change flexibly and to create novel ideas (March, 1981; Peters, 1994).

Upstream coordination. Upstream coordination is bringing together suppliers and the firm as one unit (Ashkenas et al., 2002; Kerr & Ulrich, 1995). This collaboration allows firms to reduce cost and to increase flexibility by working as a single unit rather than as several separate entities. Due to increasing complexity, one firm can no longer do everything; so instead it must concentrate on core competencies and focused missions (Ashkenas et al., 2002; Jin & Hong, 2007; Kerr & Ulrich, 1995). Both the firm and the supplier could gain by creating necessary products and working out a financial situation where both will profit (Ashkenas et al., 2002; Youn, Hong, & Nahm, 2008). This collaboration also could increase innovation if ideas are shared cross-organizationally. Communication allows other intelligent minds to see the problem and to find a solution to benefit all parties.

Downstream coordination. Downstream coordination is bringing together customers and the firm as one unit (Cross et al., 2000). Customer needs change at an increasing rate. By involving the customer, the firm can be more informed about the customer's likes, dislikes, needs, and wants (Ashkenas et al., 2002; Cross et al., 2000). This relationship with the customer also fosters innovative outcomes by leading the firm to create products that it otherwise might not have thought of offering in both existing and new markets (Ashkenas et al., 2002; Jeong & Hong, 2007).

Blue Ocean innovation offerings performance. Kim and Mauborgne (2005) created the concept of Blue Ocean and Red Ocean innovation. These terms were used to highlight the importance of innovation in both forms (new market as well as existing market). Blue Ocean innovation offerings are defined as product and service offerings with new innovation(s) that target new markets. Blue Ocean refers to the creation of products that create a new market. Based on the previous literature, we further distinguished new market innovation as either modular or radical innovation. Both new and existing innovations are necessary; yet many efforts to understand ways to improve the success rates of these innovations have failed. This research incorporated product complexity and innovation along with coordination mechanisms to study the success outcomes of firms that use these mechanisms in a complex environment.

Red Ocean innovation offerings performance. As mentioned previously, Kim and Mauborgne (2005) created the concept of Blue Ocean and Red Ocean innovation. These terms were used to highlight the importance of innovation in both forms (new market as well as existing market). Red Ocean innovation offerings are defined as multigenerations of products and services that include new innovation(s) in existing and growing markets. Using the previous literature, we divided existing markets further into whether they used architectural or incremental innovations. Both new and existing innovations are necessary, yet many efforts to understand ways to improve the success rates of these innovations have failed. This research incorporated product complexity and innovation along with coordination mechanisms to study the success outcomes of firms that use these mechanisms in a complex environment.

Table 4: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
Product multiplicity complexity A design state resulting from the multiplicity of product architectural elements.	Component multiplicity	The enumeration (i.e., sheer number) of components and parts received from suppliers, and parts and components to be added through transformation processes.	1) Our firm's products contain a large number of components? 2) Our firm's products contain a large number of component variants? 3) Our firm's product part list is large?	Gupta & Krishnan, 1999; Jacobs, 2007; Ramdas, 2003
	Product multiplicity	The enumeration (i.e., sheer number) of product lines that a firm offers to customers.	1) Our firm produces a large number of product variants? 2) Our firm produces a large number of products? 3) Our firm produces many SKUs?	Baldwin & Clark, 2000; Novak & Eppinger, 2001
	Market multiplicity	The enumeration (i.e., sheer number) of customer segments (i.e., local, regional, national, and global) that a firm targets and serves.	1) Our firm serves different markets for our products? 2) Our firm serves different geographic locations? 3) Our firm serves different customer segments?	Daft, 1983; Scott, 1992

Table 8 Continued: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
Product relatedness complexity A design state resulting from the relatedness among product architectural elements.	Similarity	The extent to which a product shares characteristics such as part geometries or components, fulfilling the same strategic role in the portfolio as a prior product but allowing for customization through replacement of components or modules.	1) Our products contain exchangeable parts? 2) Our products' parts are customized in-house? 3) Our products are mass-customized?	Fixson, 2005; Jacobs, 2007; Novak & Eppinger, 2001
	Inter-connectedness	The extent to which components are joined to each other within a product to form a completed product that adds value to the customer by allowing aftermarket customization with additional purchases.	1) Our components are highly networked *within* the product? 2) Our components are networked to allow for aftermarket customization of our products *without* additional purchases? 3) Our components are interconnected within the product?	Griffin, 1997a, 1997b; Kaski & Heikkila, 2002
	Complementarity as completeness	The extent to which other aftermarket product(s) need to be purchased by the end user to obtain the intended use of the original purchased product.	1) Our products require other product purchases for the full intended use of the product? 2) Our products are complementary to other product purchases?	Jacobs, 2007; Tatikonda & Stock, 2003

			3) Additional product purchases are necessary for our products? 4) Our products need additional aftermarket purchases?	

Table 8 Continued: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
Product relatedness complexity A design state resulting from the relatedness among product architectural elements	Substitutability as enhancement	The extent to which other aftermarket product(s) can be purchased by the end user to create enhanced benefits from the original purchased product.	1) Our products are enhanced through other product purchases? 2) Our products allow substitution of parts with additional purchases? 3) Our products complete the same basic function with enhancements? 4) Additional product purchases will enhance our products? 5) Our products become enhanced with additional aftermarket purchases?	Griffin, 1997a, 1997b; Kaski & Heikkila, 2002

Table 8 Continued: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
	Vertical coordination	The extent to which the firm internally separates people by hierarchical	1) New ideas for product development projects are accepted across different hierarchical levels? 2) Information is shared about product	Ashkenas et al., 2002; Devanna & Tichy, 1990

Coordination mechanisms Sustainable system-wide efforts that coordinate diverse business processes, functions, and people for the purpose of product development.		levels, title, status, or rank.	development projects across different hierarchical levels? 3) Responsibilities for product development projects are shared across different hierarchical levels? 4) Authority to make decisions about product development projects is distributed across different hierarchical levels?	
	Horizontal coordination	The extent to which the firm internally separates people by functional groups, functions, business unit, product group, or division.	1) New product design teams have frequent interaction with various functional departments? 2) Various functional departments are involved in the early stages of new product development? 3) Various functional departments are involved in the creation of new product concepts?	Ashkenas et al., 2002; Devanna & Tichy, 1990
	Upstream coordination	The extent to which the firm externally divides company plans and processes from their suppliers.	1) We consult customers early in the design efforts for products? 2) We partner with customers for the design of products? 3) Customers are frequently consulted about the design of products? 4) Customers become involved in projects only after the design is completed? 5) Customers are an integral part of the design effort for projects?	Ashkenas et al., 2002; Kerr & Ulrich, 1995
	Downstream coordination	The extent to which the firm externally separates customers, communities, and other	1) Suppliers are involved early in the design efforts for projects? 2) We partner with suppliers for the design of products?	Ashkenas et al., 2002; Cross et al., 2000

		external constituencies.	3) Suppliers are frequently consulted about the design of products? 4) Suppliers are an integral part of the design effort? 5) Suppliers are selected after the design for products is completed?	

Table 8 Continued: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
Blue Ocean innovation offerings performance New product creation that allows the capture of unexplored markets and creates successful outcomes in term of market performance, financial performance, and new product performance.	New market innovation market performance	New product creation that allows the capture of unexplored markets that allow the firm to achieve its market oriented goals in terms of market share, market growth, and new customer acquisition relative to its objectives and its competitors.	1) Profits 2) Sales 3) ROI	Kim & Mauborgne, 1991, 1993, 2005; Prajogo & Ahmed, 2006
	New market innovation financial performance	New product creation that allows the capture of unexplored markets that allow the firm to achieve its financial objectives in terms of profits, sales, and ROI relative to its objectives and its competitors.	1) Market share 2) Market growth 3) New customers	Calantone et al., 2003; Eisenberger et al., 1990
	New market innovation new product performance	New product creation that allows the capture of unexplored markets that allow the firm to achieve its new products outcomes in terms of customer satisfaction, speed to market, possibilities of future opportunities, and success of the project relative to its objectives and its competitors.	1) Leading to future opportunities 2) Speed to the market 3) On-time completion 4) Customer satisfaction	He & Wong, 2004; Kim & Mauborgne, 1991, 1993, 2005

Table 8 Continued: Constructs, Definition, Key Variables, and Literature Base

Constructs and Definition	Key Variables	Definition	Items	Literature Base
Red Ocean innovation offerings performance Changes to existing products or new generations of products that exploit the current market and create successful outcomes in term of market performance, financial performance, and new product performance	Existing market innovation market performance	Changes to existing products or new generations of products that exploit the current market that allow the firm to achieve its market-oriented goals in terms of market share, market growth, and new customer acquisition relative to its objectives and its competitors.	1) Profits 2) Sales 3) ROI	Kim & Mauborgne, 1991, 1993, 2005; Prajogo & Ahmed, 2006
	Existing market innovation financial performance	Changes to existing products or new generations of products that exploit the current market that allow the firm to achieve its financial objectives in terms of profits, sales, and ROI relative to its objectives and its competitors.	1) Market share 2) Market growth 3) New customers	Calantone et al., 2003; Eisenberger et al., 1990
	Existing market innovation new product performance	Changes to existing products or new generations of products that exploit the current market that allow the firm to achieve its new products outcomes in terms of customer satisfaction, speed to market, possibilities of future opportunities, and success of the project relative to its objectives and its competitors.	1) Leading to future opportunities 2) Speed to the market 3) On-time completion 4) Customer satisfaction	He & Wong, 2004; Kim & Mauborgne, 1991, 1993, 2005

HYPOTHESES/MODEL

Development of Hypothesis 1 and 1A. Because complexity is dynamic and changing, in complex environment more products will be created. These products will be new market products (i.e., Blue Ocean innovation) (Pellissier, 2012). Innovation is the lifeblood of a firm (Steencamp et al., 1999). Yet, half of the innovations fail to achieve successful revenue (Sivadas & Dwyer, 2000). Systems theory and social exchange theory have been used to explain the reason coordination mechanisms have led to more successful innovative outcomes given high multiplicity complexity. Product complexity and innovation challenges can be handled by organizational system networks (Ackoff, 1978; Ash, 1992; Bánáthy, 1991, 1992, 1996, 1997; Checkland, 1981, 1997). Several strategies can link employees, customers, and suppliers to create a more effective innovation system network (Ashkenas et al., 2002). In other words, strategies can create better innovation outcomes. Effective strategies are accomplished through building appropriate system networks. These system networks are the firm's resources of vertical/horizontal (employees), downstream (customers), and upstream (suppliers). These system networks provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in new markets by creating new knowledge and improving the information flow needed in the firm.

Vertical and horizontal coordination allow for employees to address problems quickly and to change in a complex fast-moving environment (Ashkenas et al., 2002; Devanna & Tichy, 1990). Vertical and horizontal coordination cut down the bureaucracy and speed up decisions to allow success in this type of environment. Vertical and horizontal coordination can empower employees to create a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in new markets by creating new knowledge and improving the information flow in the firm.

Downstream coordination allows for customers to communicate problems, needs, and changes in a complex fast-moving environment (Ashkenas et al., 2002; Kerr & Ulrich, 1995). Downstream coordination forms collaboration with customers to create a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in new markets through creating new knowledge.

Upstream coordination allows for suppliers to collaborate in creation, to avoid bottlenecks, and to speed up product development in a complex fast-moving environment (Ashkenas et al., 2002; Cross et al., 2000). Upstream coordination creates collaboration with suppliers to generate a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in new markets through creating new knowledge. Therefore:

Hypothesis 1:
Coordination mechanisms moderate the relationship between product relatedness complexity and Blue Ocean innovation offerings performance.

Hypothesis 1A:
Coordination mechanisms moderate the relationship between product multiplicity complexity and Blue Ocean innovation offerings performance.

Development of Hypothesis 2 and 2A. Because complexity is dynamic and changing, in complex environments, more products will be created. These products will be existing market products (Red Ocean innovation) (Pellissier, 2012). Innovation is the lifeblood of a firm (Steencamp et al., 1999). Yet, half of the innovations fail to achieve successful revenue (Sivadas & Dwyer, 2000). Systems theory and social exchange theory have explained the reason coordination mechanisms have led to more successful innovative outcomes during high relatedness complexity. Product complexity and innovation challenges can be handled by organizational system networks (Ackoff, 1978; Ash, 1992; Bánáthy, 1991, 1992, 1996, 1997; Checkland, 1981, 1997). Several strategies can link employees, customers, and suppliers to create a more effective innovation system network (Ashkenas et al., 2002). In other words, strategies can create better innovation outcomes. This is accomplished as specified through system networks. These system networks are the firm's resources of vertical (employees), downstream (customers), and upstream (suppliers). These system networks provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in existing markets by providing knowledge of ways to advance current products in the firm.

Vertical and horizontal coordination allow for employees to address problems quickly and to make changes in a complex fast-moving environment (Ashkenas et al., 2002; Devanna & Tichy, 1990). Vertical and horizontal coordination cut down the bureaucracy and speed up decisions to allow success in this type of environment. Vertical and horizontal coordination can empower employees to create a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in existing markets by creating knowledge of issues and advancement to current products in the firm.

Downstream coordination allows for customers to communicate problems, needs, and changes in a complex fast-moving environment (Ashkenas et al., 2002; Kerr & Ulrich, 1995). Downstream coordination creates collaboration with customers to generate a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in existing markets by creating knowledge of issues and advancement to current products in the firm.

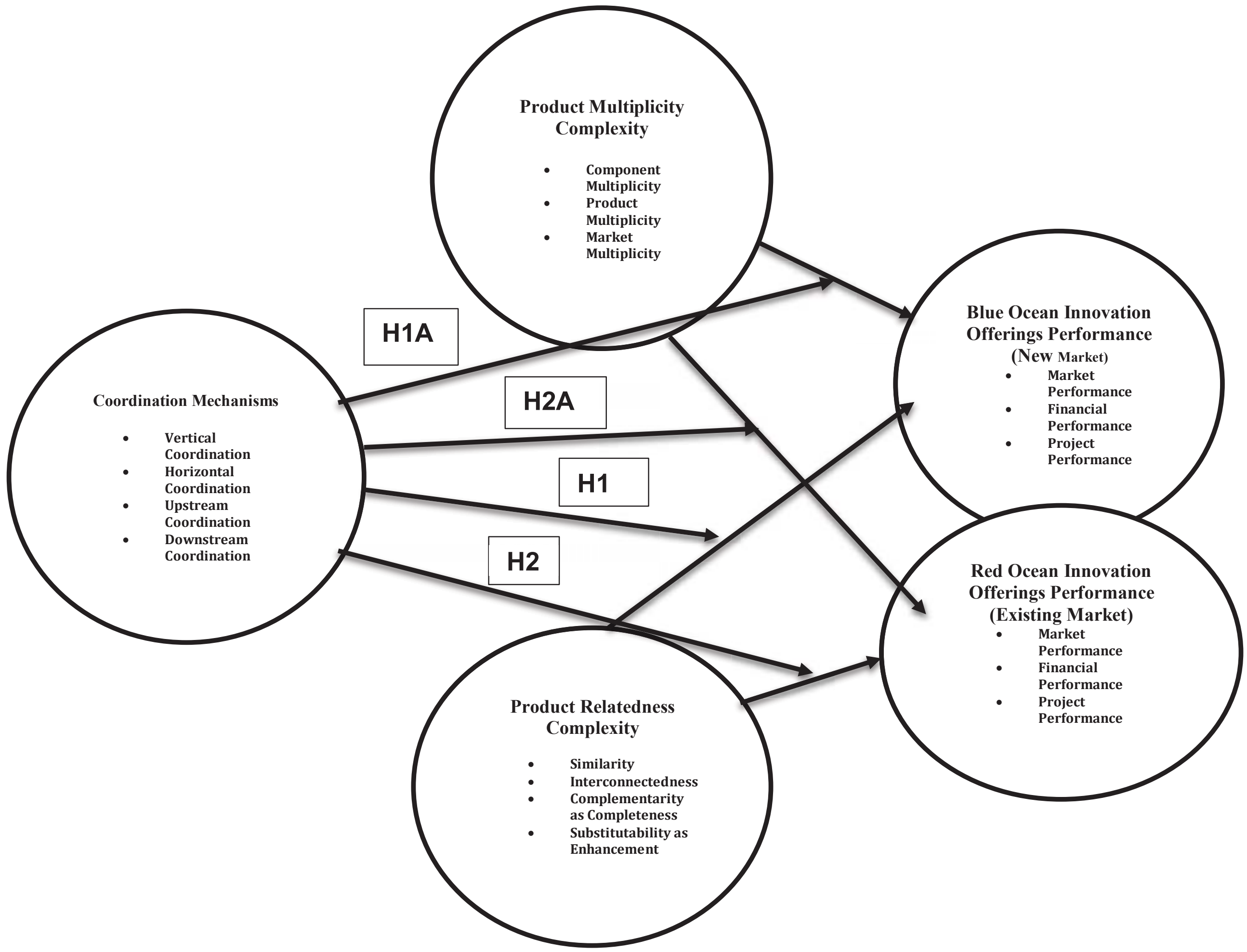
Upstream coordination allows for suppliers to collaborate in creation, to avoid bottlenecks, and to speed up product development in a complex fast-moving environment (Ashkenas et al., 2002; Cross et al., 2000). Upstream coordination creates collaboration with suppliers to generate a more effective innovation system network (Ashkenas et al., 2002). This system network can provide the information and collaboration to manage innovation effectively and to create successful innovative outcomes in existing markets by creating knowledge of issues and advancement to current products in the firm. Therefore:

Hypothesis 2:
Coordination mechanisms moderate the relationship between product relatedness complexity and Red Ocean innovation offerings performance.

Hypothesis 2A:
Coordination mechanisms moderate the relationship between product multiplicity complexity and Red Ocean innovation offerings performance.

Model

Figure 2: Model



METHODOLODGY

Testing the proposed hypotheses required a large representative sample. A cross-sectional field survey method was employed to meet the sample size requirement to provide an adequate level of statistical power (Gatignon & Xuereb, 1997; Iyer, 2001). The development and administration of surveys was done following Dillman's (1991, 2000) Tailored Design Method, which is often regarded as the standard for mail surveys. The Tailored Design Method (TDM) involves five elements for achieving high response rates: 1) a respondent-friendly questionnaire, 2) the use of four contacts by first-class mail, with an additional 'special' contact (e.g. certified mail or a telephone call), 3) the use of return envelopes with real first-class stamps, 4) personalized correspondence, and 5) a token financial incentive sent with the survey request. The questionnaires were emailed, faxed, or mailed as desired by the respondent (see next paragraph).

The survey was designed to be respondent friendly—not too long, with easily read terms and mostly Likert-type scales. The survey was prepared online, however, the respondents also were given an option to receive a mailed hard copy or a faxed copy of the survey. The questionnaires that were mailed were mailed with a postage-paid return envelope. Cover letters and questionnaires that were faxed were faxed with a return fax number listed. All letters were specifically addressed to the name of the firm's executive (i.e., CEO, VP of R&D, Director of Product Development, etc.) explaining the purpose of the survey. The survey cover page also included a nondisclosure agreement, indicating that the responses would be treated confidentially (no firm names would be used in this research) and that the data would be used for research purposes only. Appeals were used to highlight the importance of each response and of the research. Respondents who completed questionnaires were entered into a drawing as an incentive to participate and to increase the response rate.

Data collection involved a sequence of contacts that consisted of phone calls, faxes, and/or emails from late April to mid-August 2013. Phone calls were made to each of the 750 potential subjects to: 1) Verify that the information was correct (position of the key respondent), 2) Solicit participation and notify the key respondent about the survey, and 3) Ask the preferred method for survey delivery (i.e., email address, fax number, or mailing address) for the respondents.

Product multiplicity complexity model fit. Convergent validity and reliability for product multiplicity complexity: The nine items for product multiplicity complexity and their coded names are listed in Table 5. AVE and Cronbach's α values (for reliability testing) for the final measurement model are displayed following each dimension's name. Model fit statistical values of goodness of fit (GFI, AGFI, RMR, CFI, and NFI) are shown in the table under model fit. All of the model fit values and the AVE values shown in the table signified acceptable convergent validity for each dimension of the construct. The CFA of the product multiplicity complexity from AMOS is shown in Figure 3.

Table 5: Model Fit for Product Multiplicity Complexity

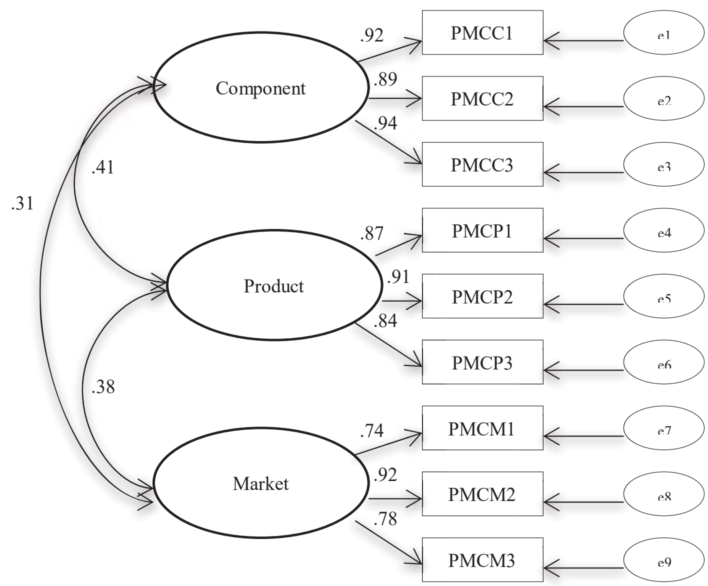
Coded Name	**Item**	**Model Fit**
	Component (Final AVE=.80, α=.92)	
PMCC1	Our firm's products contain a large number of components?	GFI= 0.973 AGFI= 0.915 RMR= 0.011 CFI= 0.980 NFI= 0.979
PMCC2	Our firm's products contain a large number of component variants?	
PMCC3	Our firm's product part list is large?	
	Product (Final AVE=.74, α=.88)	
PMCP1	Our firm produces a large number of product variants?	GFI= 0.964 AGFI= 0.928 RMR= 0.022 CFI= 0.985 NFI= 0.983
PMCP2	Our firm produces a large number of products?	
PMCP3	Our firm produces many Skus?	
	Market (Final AVE=.70, α=.83)	
PMCM1	Our firm serves different markets for our products?	GFI= 0.970 AGFI= 0.930 RMR= 0.018 CFI= 0.992 NFI= 0.991
PMCM2	Our firm serves different geographic locations?	
PMCM3	Our firm serves different customer segments?	

Discriminant validity for product multiplicity complexity: Table 6 displays the results of the single-factor/two-factor (pairwise) X^2 tests for discriminant validity. The differences in chi- squared values for each pair of dimensions were all statistically significant at $p < 0.001$ ($df = 1$, critical value = 12.485), offering appropriate evidence of discriminant validity.

Table 6: Discriminant Validity for Product Multiplicity Complexity

	PMCC			PMCP		
	COR.	**SIN.**	**Δ**	**COR.**	**SIN.**	**Δ**
PMCM	37.4	185.2	147.8	32.7	161.4	128.7
PMCC				22.4	134.5	112.1

Figure 3: Product Multiplicity Complexity Model Fit from AMOS



Product relatedness complexity model fit. Convergent validity and reliability for product relatedness complexity: The fifteen items for product relatedness complexity and their coded names are listed in Table 7. AVE and Cronbach‘s α values (for reliability testing) for the final measurement model are displayed following each dimension’s name. Model fit statistical values of goodness of fit (GFI, AGFI, RMR, CFI, and NFI) are shown in the table under model fit. All of the model fit values and the AVE values shown in the table signified acceptable convergent validity for each dimension of the construct. Three items were deleted: PRCC4, PRCSE3, and PRCSE5. After a re-evaluation of the definitions and a review of the literature, these items were found to be ambiguous and to not fit the scope of this study. Because they did not fit the scope of this study, and due to the higher values for model fit (GFI, AGFI, RMR, CFI, and NFI) when deleted, these three items were deleted. The CFA of the product relatedness complexity from AMOS is shown in Figure 4.

Table 7: Model Fit for Product Relatedness Complexity

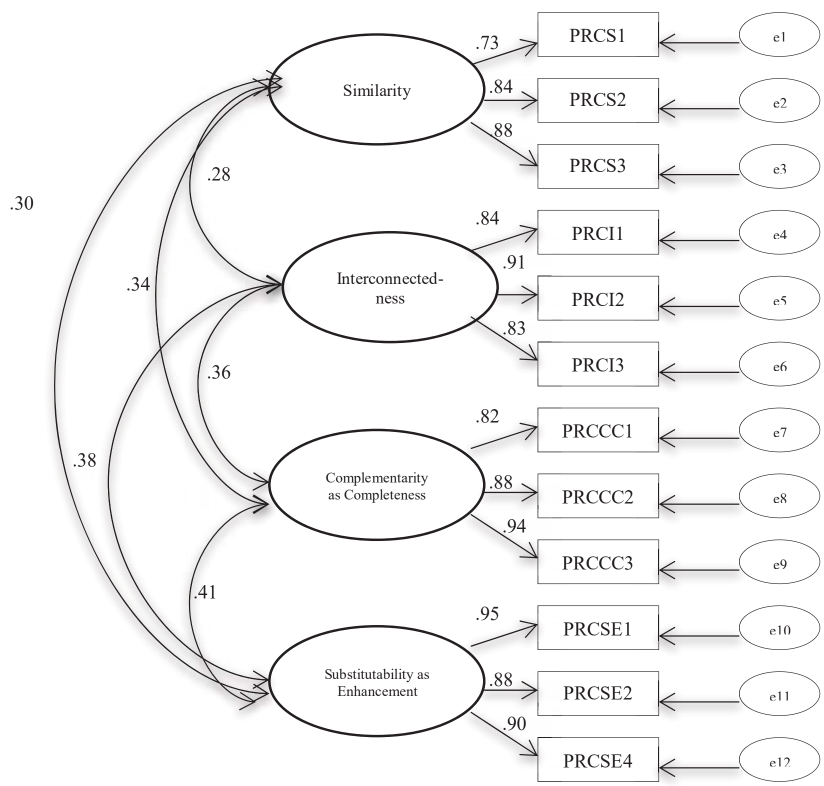
Coded Name	Item	Model Fit
Similarity (Final AVE=.69 α=.87)		
PRCS1	Our products contain exchangeable parts?	GFI= 0.944 AGFI= 0.901 RMR= 0.031 CFI= 0.934 NFI= 0.928
PRCS2	Our products parts are customized in-house?	
PRCS3	Our products are mass-customized?	
Interconnectedness (Final AVE=.63, α=.83)		
PRCI1	Our components are highly networked *within* the product?	GFI= 0.956 AGFI= 0.912 RMR= 0.021 CFI= 0.968 NFI= 0.966
PRCI2	Our components are networked to allow for aftermarket customization of our products *without* additional purchases?	
PRCI3	Our components are interconnected within the product?	
Complementarity as Completeness (Final AVE=.65, α=.85)		
PRCCC1	Our products require other product purchases for the full intended use of the product?	GFI= 0.980 AGFI= 0.947 RMR= 0.010 CFI= 0.991 NFI= 0.990
PRCCC2	Our products are complementary to other product purchases?	
PRCCC3	Additional product purchases are necessary for our products?	
PRCCC4	Our products need additional aftermarket purchases? (deleted)	
Substitutability as Enhancement (Final AVE=.70, α=.80)		
PRCSE1	Our products are enhanced through other product purchases?	GFI= 0.994 AGFI= 0.989 RMR= 0.008 CFI= 0.992 NFI= 0.990
PRCSE2	Our products allow substitution of parts with additional purchases?	
PRCSE3	Our products complete the same basic function with enhancements? (deleted)	
PRCSE4	Additional product purchases will enhance our products?	
PRCSE5	Our products become enhanced with additional aftermarket purchases? (deleted)	

Discriminant validity for product relatedness complexity: Table 8 displays the results of the single-factor/two-factor (pairwise) X^2 tests for discriminant validity. The differences in chi- squared values for each pair of dimensions were all statistically significant at $p < 0.001$ ($df = 1$, critical value = 13.675), offering appropriate evidence of discriminant validity.

Table 8: Discriminant Validity for Product Relatedness Complexity

	PRCS			PRCI			PRCCC		
	COR.	SIN.	Δ	COR.	SIN.	Δ	COR.	SIN.	Δ
PRCSE	41.2	402.6	361.4	19.9	297.2	277.3	25.8	368.5	342.7
PRCS				25.4	302.4	277.0	41.5	354.7	313.2
PRCI							40.2	333.7	293.5

Figure 4: Product Relatedness Complexity Model Fit from AMOS



Coordination mechanisms model fit. Convergent validity and reliability for coordination mechanisms: The seventeen items for coordination mechanisms and their coded names are listed in Table 9. AVE and Cronbach's α values (for reliability testing) for the final measurement model are displayed following each dimension's name. Model fit statistical values of goodness of fit (GFI, AGFI, RMR, CFI, and NFI) are shown in the table under model fit. All of the model fit values and the AVE values shown in the table signified acceptable convergent validity for each dimension of the construct. Five items were deleted: CMVC4, CMDC4, CMDC5, CMUC4, and CMUC5. After a re-evaluation of the definitions and a review of the literature, these items were found to be ambiguous and to not fit the scope of this study. Because they did not fit the scope of this study, and due to the higher values for model fit (GFI, AGFI, RMR, CFI, and NFI) when deleted, these five items were deleted. The CFA of the coordination mechanisms from AMOS is shown in Figure 5.

Table 9: Model Fit for Coordination Mechanisms

Coded Name	Item	Model Fit
Vertical Coordination (Final AVE=.73, α=.92)		
CMVC1	New ideas for product development projects are accepted across different hieratical levels	GFI= 0.984 AGFI= 0.931 RMR= 0.019 CFI= 0.990 NFI= 0.989
CMVC2	Information is shared about product development projects across different hieratical levels	
CMVC3	Responsibilities for product development projects are shared across different hieratical levels	
CMVC4	Authority to make decisions about product development projects is distributed across different hieratical levels (deleted)	
Horizontal Coordination (Final AVE=.75, α=.85)		
CMHC1	New product design teams have frequent interaction with various functional departments	GFI= 0.974 AGFI= 0.928 RMR= 0.014 CFI= 0.989 NFI= 0.987
CMHC2	Various functional departments are involved in the early stages of new product development.	
CMHC3	Various functional departments are involved in the creation of new product concepts.	
Downstream Coordination (Final AVE=.82, α=.94)		
CMDC1	We consult customers early in the design efforts for products?	GFI= 0.995 AGFI= 0.992 RMR= 0.003 CFI= 0.993 NFI= 0.990
CMDC2	We partner with customers for the design of products?	
CMDC3	Customers are frequently consulted about the design of products?	
CMDC4	Customers become involved in projects only after the design is completed? (deleted)	
CMDC5	Customers are an integral part of the design effort for projects? (deleted)	

Table 9 Continued: Model Fit for Coordination Mechanisms

Coded Name	Item	Model Fit
Upstream Coordination (Final AVE=.77, α=.91)		
CMUC1	Suppliers are involved early in the design efforts for projects?	GFI= 0.972 AGFI= 0.925 RMR=
CMUC2	We partner with suppliers for the design of products?	

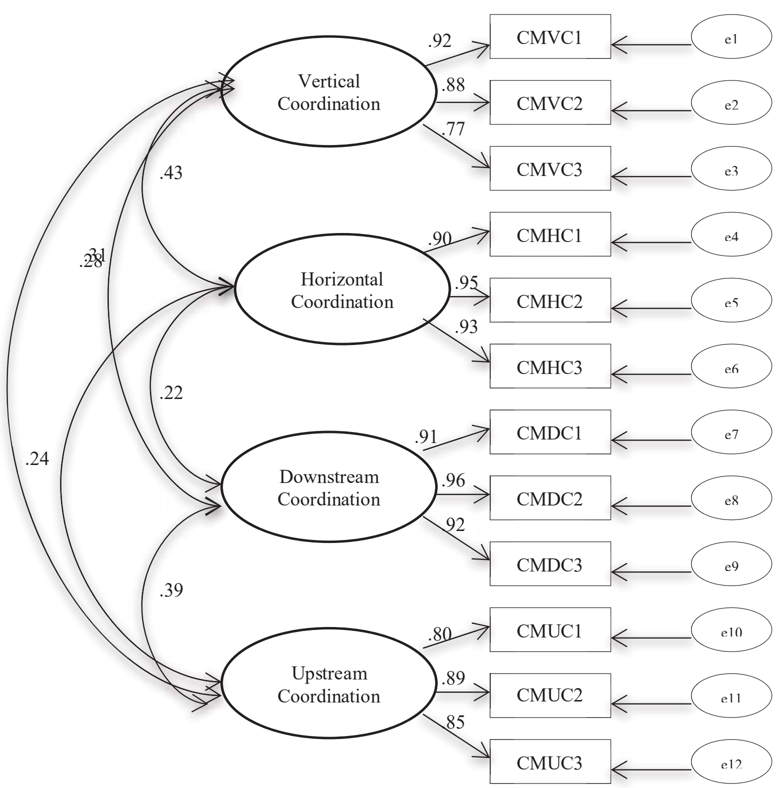
CMUC3	Suppliers are frequently consulted about the design of products?	0.017 CFI= 0.989 NFI= 0.982
CMUC4	Suppliers are an integral part of the design effort? (deleted)	
CMUC5	Suppliers are selected after the design for products is completed? (deleted)	

Discriminant validity for coordination mechanisms: Table 10 displays the results of the single-factor/two-factor (pairwise) X^2 tests for discriminant validity. The differences in chi- squared values for each pair of dimensions were all statistically significant at p < 0.001 (df = 1, critical value = 14.154), offering appropriate evidence of discriminant validity.

Table 10: Discriminant Validity for Coordination Mechanisms

	CMVC			CMHC			CMDC		
	COR.	**SIN.**	**Δ**	**COR.**	**SIN.**	**Δ**	**COR.**	**SIN.**	**Δ**
CMUC	44.7	484.8	440.1	22.4	297.5	275.1	30.4	401.7	371.3
CMVC				34.6	354.8	320.2	37.8	387.1	349.3
CMHC							40.1	367.9	327.8

Figure 5: Coordination Mechanisms Model Fit from AMOS



Blue Ocean innovation offerings performance model fit. Convergent validity and reliability for Blue Ocean innovation offerings performance: The ten items for Blue Ocean innovation offerings performance and their coded names are listed in Table 11. AVE and Cronbach's *α* values (for reliability testing) for the final measurement model are displayed following each dimension's name. Model fit statistical values of goodness of fit (GFI, AGFI, RMR, CFI, and NFI) are shown in the table under model fit. All of the model fit values and the AVE values shown in the table signified acceptable convergent validity for each dimension of the construct. One item was deleted: BOPP1. After a re-evaluation of the definitions and a review of the literature, this item was found to be ambiguous and to not fit the scope of this study. Because it did not fit the scope of this study, and due to the higher values for model fit (GFI, AGFI, RMR, CFI, and NFI) when deleted, this item was deleted. The CFA of the Blue Oceans innovation offerings from AMOS is shown in Figure 6.

Table 11: Model Fit for Blue Ocean Innovation Offerings Performance

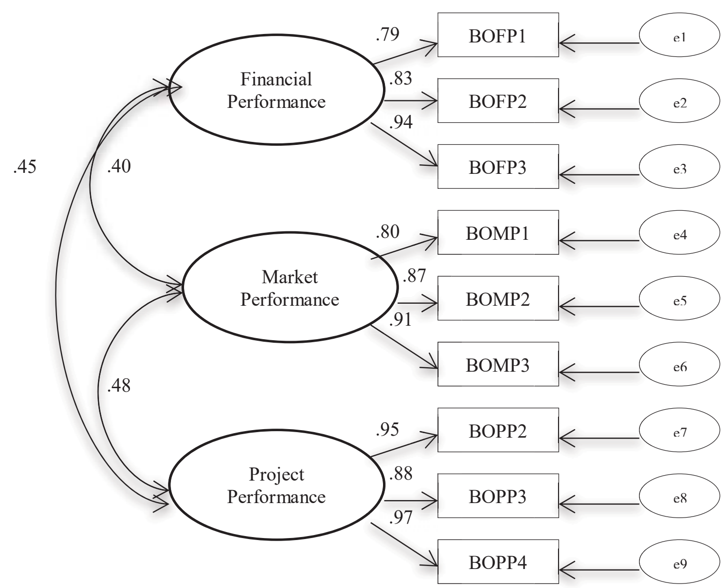
Coded Name	**Item**	**Model Fit**
Relative to your business unit's **objectives**, the **NEW MARKET project program** has been successful in terms of:		
Financial Performance (Final AVE=.78, α=.89)		
BOFP1	Profits?	GFI= 0.987 AGFI= 0.974 RMR= 0.014 CFI= 0.995 NFI= 0.992
BOFP2	Sales?	
BOFP3	Return on investment (ROI)?	
Market Performance (Final AVE=.76, α=.90)		
BOMP1	Market share?	GFI= 0.993 AGFI= 0.990 RMR= 0.004 CFI= 0.998 NFI= 0.994
BOMP2	Market growth?	
BOMP3	New customers?	
Project Performance (Final AVE=.79, α=.88)		
BOPP1	Leading to future opportunities? (deleted)	GFI= 0.988 AGFI= 0.976 RMR= 0.021 CFI= 0.964 NFI= 0.960
BOPP2	Speed to the market?	
BOPP3	On-time completion?	
BOPP4	Customer satisfaction?	

Discriminant validity for Blue Ocean innovation offerings performance: Table 12 displays the results of the single-factor/two-factor (pairwise) X^2 tests for discriminant validity. The differences in chi-squared values for each pair of dimensions were all statistically significant at $p < 0.001$ ($df = 1$, critical value = 12.687), offering appropriate evidence of discriminant validity.

Table 12: Discriminant Validity for Blue Ocean Innovation Offerings Performance

	BOFP			BOMP		
	COR.	**SIN.**	**Δ**	**COR.**	**SIN.**	**Δ**
BOPP	14.1	398.3	384.2	22.4	384.1	361.8
BOFP				25.1	342.9	317.8

Figure 6: Blue Ocean Innovation Offerings Performance Model Fit from AMOS



Red Ocean innovation offerings performance model fit. Convergent validity and reliability for Red Ocean innovation offerings performance: The ten items for Red Ocean innovation offerings performance and their coded names are listed in Table 13. AVE and Cronbach's *α* values (for reliability testing) for the final measurement model are displayed following each dimension's name. Model fit statistical values of goodness of fit (GFI, AGFI, RMR, CFI, and NFI) are shown in the table under model fit. All of the model fit values and the AVE values shown in the table signified acceptable convergent validity for each dimension of the construct. One item was deleted: ROPP1. After a re-evaluation of the definitions and a review of the literature, this item was found to be ambiguous and to not fit the scope of this study. Because it did not fit the scope of this study, and due to the higher values for model fit (GFI, AGFI, RMR, CFI, and NFI) when deleted, this item was deleted. The CFA of the Red Ocean innovation offerings performance from AMOS is shown in Figure 7.

Table 13: Model Fit for Red Ocean Innovation Offerings Performance

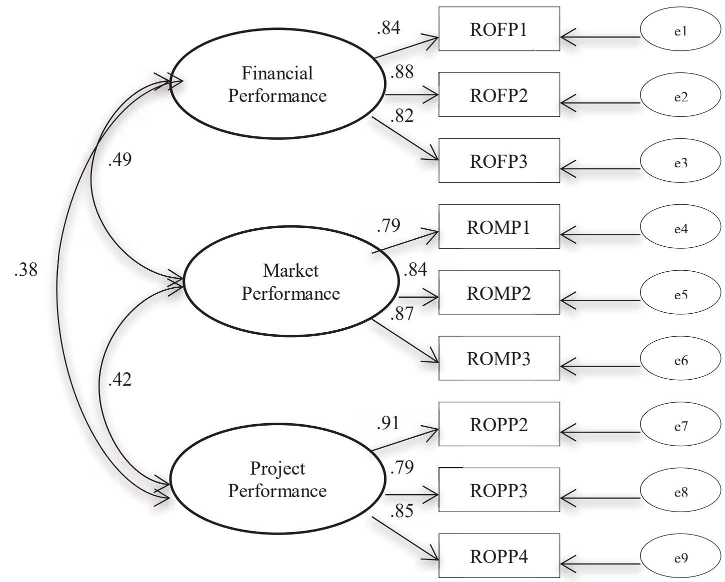
Coded Name	Item	Model Fit
Relative to your business unit's **objectives**, the **EXISTING MARKET new product project program** has been successful in terms of:		
Financial Performance (Final AVE=.70, α=.83)		
ROFP1	Profits?	GFI= 0.948 AGFI= 0.874 RMR= 0.035 CFI= 0.942 NFI= 0.938
ROFP2	Sales?	
ROFP3	Return on investment (ROI)?	
Market Performance (Final AVE=.65, α=.89)		
ROMP1	Market share?	GFI= 0.957 AGFI= 0.888 RMR= 0.031 CFI= 0.931 NFI= 0.928
ROMP2	Market growth?	
ROMP3	New customers?	
Project Performance (Final AVE=.68, α=.85)		
ROPP1	Leading to future opportunities? (deleted)	GFI= 0.936 AGFI= 0.875 RMR= 0.027 CFI= 0.947 NFI= 0.944
ROPP2	Speed to the market?	
ROPP3	On-time completion?	
ROPP4	Customer satisfaction?	

Discriminant validity for Red Ocean innovation offerings performance: Table 14 displays the results of the single-factor/two-factor (pairwise) X^2 tests for discriminant validity. The differences in chi-squared values for each pair of dimensions were all statistically significant at $p < 0.001$ ($df = 1$, critical value = 12.654), offering appropriate evidence of discriminant validity.

Table 14: Discriminant Validity for Red Ocean Innovation Offerings Performance

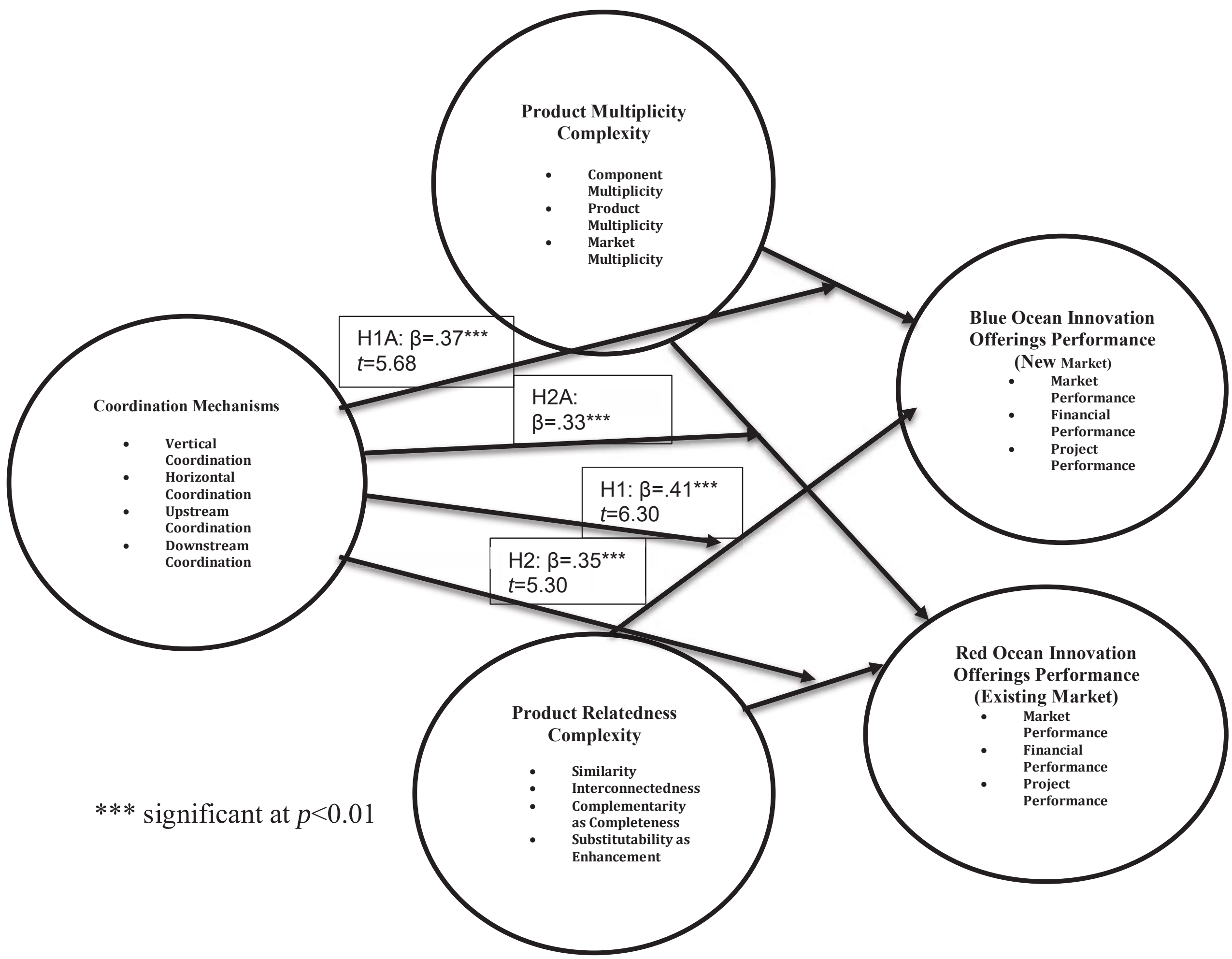
	ROFP			**ROMP**		
	COR.	**SIN.**	**Δ**	**COR.**	**SIN.**	**Δ**
ROPP	11.4	402.1	390.7	17.0	389.4	372.4
ROFP				16.5	365.4	348.9

Figure 7: Red Ocean Innovation Offerings Performance Model Fit from AMOS



RESULTS

Figure 8 Model Results



The model indicated that under conditions where coordination mechanisms were used to manage product complexity, outcomes were higher than in firms where coordination mechanisms were not used. Use of these coordination mechanisms led to better innovation performance outcomes (Blue Ocean and Red Ocean) under the conditions of high product complexity (multiplicity and relatedness) in Hypotheses 1, 1A, 2, and 2A. To sum up this model, firms need to use coordination mechanisms to manage product complexity to create successful innovation outcomes that create overall success for the firms.

Table 15: Structural Model Results

Hypothesis	**Relationship**	**Beta (β)**	**Hypothesis Supported**	***t*-Statistic (t)**	R^2
Hypothesis 1	**CM *PRC Moderates BOOP**	**.41***	**Yes**	**6.30**	**.39**
Hypothesis 2	**CM *PRC Moderates ROOP**	**.35***	**Yes**	**5.30**	**.37**
Hypothesis 1A	**CM *PMC Moderates BOOP**	**.37***	**Yes**	**5.68**	**.39**
Hypothesis 2A	**CM *PMC Moderates ROOP**	**.33***	**Yes**	**4.84**	**.37**

DISCUSSION AND CONCLUSIONS

Hypothesis 1 and 1A Supported The hypotheses were tested and found to be statistically significant at a $p < 0.01$ level, with β =.41, $t = 6.30$, and $R^2 = .39$ for hypothesis 1 and β =.37, $t = 5.68$, and $R^2 = .39$ for hypothesis 1A indicating a good amount of variance explained by the model. This indicates that coordination mechanisms moderate the relationship between product complexity and supply chain performance measured by Blue Ocean innovation offerings performance. This means that the coordination mechanisms support and enhance the long-term objectives and goals of firms in high product complexity environments. In particular, if a firm is dealing with multiple products from relatedness complexity, it should use these coordination mechanisms to create products that provide better Blue Ocean performance. Because complexity is dynamic and changing, in complex environment more products will be created. These products will be in terms of new market products (Blue Ocean innovation) (Pellissier, 2012). Innovation is the lifeblood of a firm (Steencamp et al., 1999). Yet, half of the innovations fail to achieve successful revenue (Sivadas & Dwyer, 2000). Systems theory and social exchange theory have been used to explain the reason that coordination mechanisms lead to more successful innovative outcomes given high multiplicity complexity. Product complexity and innovation challenges can be handled by organizational system networks (Ackoff, 1978; Ash, 1992; Bánáthy, 1991, 1992, 1996, 1997; Checkland, 1981, 1997). Several strategies can link employees, customers, and suppliers to create a more effective innovation system network (Ashkenas et al., 2002). In other words, strategies can create better innovation outcomes. Better innovation outcomes are accomplished as specified through system networks. These system networks are the firm's resources of vertical/horizontal (employees), downstream (customers), and upstream (suppliers). These system networks provide the information and collaboration to manage innovation effectively to create successful innovative outcomes in new markets by creating new knowledge and improving the information flow within the firm.

Hypothesis 2 and 2A: Supported

The hypotheses were tested and found to be statistically significant at a $p < 0.01$ level, with $\beta = .35$, $t = 5.30$, and $R^2 = .37$ for hypothesis 2 and $\beta = .33$, $t = 4.84$, and $R^2 = .37$ for yypothesis 2A indicating a good amount of variance explained by the model. This indicates that coordination mechanisms moderate the relationship between product complexity and supply chain performance measured by Red Ocean innovation offerings performance. This means that the coordination mechanisms support and enhance the long-term objectives and goals of firms in high product complexity environments. In particular, if a firm is dealing with multiple products from relatedness complexity, it should use these coordination mechanisms to create products that provide better Red Ocean performance. When relatedness complexity is very dynamic, many components and products are being created at a very rapid pace. Firms that do not manage complexity properly lose profits (Hoole, 2006). Yet, when firms properly manage complexity, they are expected to produce successful returns (Meeker et al., 2009; Meyer & Mugge, 2001). According to contingency theory, strategies need to change in different external contexts for success (Miller, 1988). Therefore, different strategies must be applied in different contexts. In this research, product complexity strategies have been presented as coordination mechanisms, which will be more effective at managing product complexity than current methods are. These coordination mechanisms can decrease the strain from product complexity to create successful outcomes (Burns & Stalker, 1961; Miller, 1988; Mintzberg, 1979). Complexity increases the strain between employees, customers, and suppliers due to rapid changing (Jacobs, 2007). Several coordination mechanisms can link employees, customers, and suppliers to create more effective outcomes in complex environments (Ashkenas et al., 2002; Bozarth et al., 2009). Similarity, interconnectedness, and complementarity complexity create a dynamic environment that can cause strain between employees, customers and suppliers. Collaboration allows a firm to be more productive by making better choices regarding the components needed (Ashkenas et al., 2002; Kerr and Ulrich, 1995; Cross et al., 2000; Devanna and Tichy, 1990). Coordination mechanisms help align who, what, why, and where in the organization to provide success through knowledge and information (Ashkenas et al., 2002; Cross et al., 2000; Devanna and Tichy, 1990; Kerr & Ulrich, 1995). In the same manner, the coordination of employees, customers, and suppliers can help firms manage similarity complexity, interconnectedness complexity, and complementarity complexity.

DISCUSSION/ CONCLUSION

This research provided several advances to the area of product complexity in both the academic and managerial fields. First, this research provided evidence that market conditions lead to a need for more complex products. This product complex environment will continue to limit a firm's profits if it does manage effectively. Thus, this research showed ways to relieve the strain from product complexity by providing coordination mechanisms to deal effectively with product complexity to create better outcomes. This research also explored the drivers of complexity, thereby creating a clearer understanding of the origin of complexity.

Second, a major contribution of this research has been the development of a reliable instrument and clear definitions that supports future research in the area of product complexity. Product complexity was identified and clear definitions and examples were produced, as well as an instrument to measure product complexity. Twenty one variables for product complexity (9 for Product multiplicity complexity and 12 for product relatedness complexity) were developed to measure product complexity and the relationship to innovation and coordination mechanisms. This research provided future researchers a better understanding of complexity and thus provided a means by which to research complexity deeper in the future. This enables researchers to further explore product complexity, which has been receiving increasing attention, but is still in need of more empirical research.

Second, a theoretical knowledge integration framework is provided that identifies product complexity, coordination mechanism, and Blue and Red Ocean innovations (market performance of projects). The key to success is coordination mechanisms which can effectively create Blue and Red Ocean innovations in highly product complex environment. Much more has to be learned about the linkages to types of coordination mechanisms on type of product complexities. This research provides a framework that will be beneficial for future study of product complexity. The use of the proposed constructs allows researchers to formulate and test numerous hypotheses. Other constructs may be added or modified in the future to further examine other connections or more detailed hypotheses.

Forth, this research also provided a valuable benchmarking tools for product development executives to assess the extent of their coordination mechanism in product complex environments. This means that in highly complex environments, firms could employ the benchmark of vertical, horizontal, upstream, and downstream coordination researched in this research to alleviate the strain and potential performance losses, possibly saving firms millions of dollars. Because complexity is growing, it must be managed appropriately; thus, mechanisms must be put into place to alleviate the strain from complexity. This research provided this possible benchmark tool on how to achieve this and how to be successful in product innovation under complex environments. Without using these coordination mechanisms, firms will not be as effective in creating successful innovation performance.

Finally, this research was an empirical research of coordination mechanism effect on product complexity in multiple different industries. The data was received from 208 firms from managers and executives that specialize in product development. This research adopted a sound methodology, which resulted in more precise measurement of underlying constructs of product complexity and examined the mechanisms that affect Blue Ocean and Red Ocean outcomes including overall performance.

REFERENCES

References available upon Request

DECISION SCIENCES INSTITUTE

Configuring Multiple-Load AGV Using Simulated Annealing Algorithm

Behnam Rahimikelarijani
Department of Industrial Engineering, Lamar University, Beaumont, TX 77710, USA
Email: brahimikelar@lamar.edu

Arash Abedi
Department of Industrial Engineering, Lamar University, Beaumont, TX 77710, USA
Email: aabedi@lamar.edu

Abstract

Automated Guided Vehicle (AGV) is one of the advanced equipment designed to move materials in manufacturing facilities. AGVs are mainly classified based on their layouts. This paper studies AVGs system in Tandem Layout by using a quadratic mathematical model to determine the optimal station layout and AGV load capacity in each cell. The optimization model presented in this study contains a quadratic function of several variables that makes the problem NP-hard. Therefore, a Simulated Annealing (SA) algorithm is proposed to solve the model in a timely manner. The numerical example shows that the exact solving time, an alternative solver, increases exponentially by increasing problem's size while the proposed SA algorithm slightly changes and results in better solutions.

KEYWORDS: Material handling system, AGV, multiple-load, Simulated annealing

INTRODUCTION

An efficient transportation system not only improves performance and increase the flexibility of a system but eventually results in reducing final product or service cost, production time and pollution [1]–[3]. The transportation could perform outside of a manufacturing plant which is related to supply chain logistic [4]–[6] or could be part of material flow in a manufacturing system. An efficient material flow network design can reduce material flow costs by 20% [7].

One of the most advanced material handle system in the manufacturing environment is Automated Guided Vehicle. AGVs are driverless vehicle that transfer parts between stations. The application of the AGVs was first developed during the late 20th when it was used for repeating transportation tasks in areas, such as warehouses, production lines, container terminals and external (underground) transportation systems[8].

AGVs transfer materials by following a wire guide-path. Designing an efficient guide path can improve the navigation process and performance of the system. The AGV guide path configurations were discussed in literature, which are Conventional/Traditional, Tandem, Single loop, Bi-directional shortest path and Segmented flow[9]. This paper focuses on tandem configuration which is proposed by [10]. In this configuration, stations are divided into non-overlapping zones where one AGV serves the demands in a cell. Transfer between zones is possible by defining the transfer point. This configuration was studied in the several researches because of the following advantages [11]:

- Simple control strategy by applying one AGV in each zone
- Removing collisions and traffic problems
- Finding optimal facilities' locations in each work station
- Developing a group technology (GT) system
- Increasing flexibility due to changes in the number of work stations and changes in the production planning
- Simplifying production processes in each loop

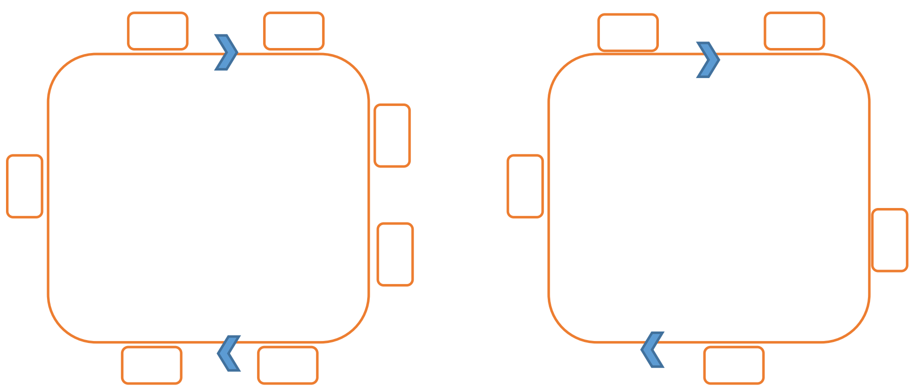


FIGURE 1: TANDEM CONFIGURATION AS PROPOSED BY [10].

There are many articles on dividing station into non-overlapping zones [9], [11]–[18]. These researches focus on balancing workload in different zones by minimizing the maximum work load or considering balancing strategy (see [11] for details). In addition, since the problem is binary quadratic assignment by nature, it is classified as NP-hard problem. Since, finding the optimal solution for large size problem is not possible using exact methods in timely manner[19], authors developed heuristic or metaheuristic algorithms to find optimal or near optimal answers in an efficient time.

After selecting AGV as the material handling system, AGV's capacity should be calculated [8]. In simulation studies conducted by [20] and [21], indicated that increasing load capacity will reduce the AGV fleet size and also will decrease the average throughput time. This paper significantly considers multiple load AGVs in modeling, this configuration was rarely considered in previous studies [8].

The first research that suggested a tandem AGV system with multi-load AGVs was [22]. They initially determined the subsets of workstations solving a travel salesman problem, then they applied a Markov chain model to calculate the AGV utilization in each partitioned subset, from which the final subsets of workstations were determined. Finally, they used simulation to validate their proposed methodology. In another study, [16], proposed a design methodology for multiple loads AGV in tandem layout. Their objective is to achieve the workload-balance between vehicles of different loops and the trade-off between the load-carrying capacity of a vehicle and its work capacity.

The present paper proposes a mathematical model to design a multiple-load AGV handling system in tandem configuration by considering balancing strategy. A metaheuristic algorithm is developed to solve the optimization model as the problem is NP-hard. Also a numerical example is provided to assess the efficiency of the algorithm and optimal results.

The rest of paper arranges as follow: in section 0 a detail problem statement will be presented. The parameters, variables and the mathematical model are included in this section. In section 0, a simulated annealing algorithm is developed to solve the proposed mathematical model. In section 0, set of illustrative examples are presented to solve the proposed model san Meta heuristic algorithm. Finally result analysis and validation concluded in this research.

MATHEMATICAL MODEL

Problem and notations

This paper presumes a system that AGVs are performing transportation task on bi-directional path in each cell, where at least two stations are in each cell. Each station is only allowed to assign to one cell, where one AGV serves as material handling facility. Empty transportations are not considered and the number of cells is the input variable.

The goal is to classify stations into non-overlapping loops in order to minimize both inter and intra loops transportation by considering multiple load AGV. A balancing strategy is used in this paper which was first proposed by [11]. The following notation represents indexes and variables that are utilized in the mathematical formulation.

- i(k): define i(k)th station;
- j(l): define j(l)th cell;
- p: define AGV capacity;

Please note that f_{ik} is number of transported loads between two stations.

The decision variables in the optimization problem are binary as follow:

x_{ij}: Equal to 1 if station i assigns to cell j else equal to zero.

y_{jp}: Equal to 1 if AGV with P capacity assigns to cell j else equal to zero.

The control parameters listed below are used to establish values, lower and upper bounds for constraints in the formulation:

w_p: Performance coefficient of AGV with P capacity;

f_{ik}: Flow between station k to i;

f_i: Flow from different station to station i ($\sum_k f_{ik}$);

T: The time that AGV is available;-

t_i: Mean loading, unloading and process time on part in ith station;

t_j': Bottleneck time in jth station;

N: Total number of predefined loop;

$\acute{\eta}, \eta$: Upper and lower bound of flow coefficient;

λ_p: Penalty of selecting AGV with P capacity;

n: Number of stations;

Proposed model

Based on the above notations and assumptions Model I can be formulated as follow:

Model I

(1) $$\text{Min } Z = \underbrace{\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\frac{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik}x_{ij}x_{kl}}{\sum_{p=1}^{m} y_{jp}w_p}\right)}_{i\neq k, j=l} + \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik}x_{ij}x_{kl}}_{i\neq k, j\neq l} + \sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp}w_p\lambda_p$$

Subject to:

(2) $$\sum_{j=1}^{N} x_{ij} = 1 \quad i = 1,2,3 \dots, n$$

(3) $$\sum_{p=1}^{m} y_{jp} = 1 \qquad j = 1,2,3 \dots, N$$

(4) $$\sum_{i=1}^{n} x_{ij} \geq 2 \qquad j = 1,2,3 \dots, N$$

(5) $$\eta\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik}x_{ij}x_{kl}}_{i\neq k}\times\sum_{p=1}^{m} y_{jp}w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp}w_p}\right) \leq \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik}x_{ij}x_{kj}}_{i\neq k} \qquad \varphi=1,2,3 \dots, N$$

(6) $$\acute{\eta}\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik}x_{ij}x_{kl}}_{i\neq k}\times\sum_{p=1}^{m} y_{jp}w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp}w_p}\right) \geq \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik}x_{ij}x_{kj}}_{i\neq k} \qquad \varphi=1,2,3 \dots, N$$

(7) $$t_j' = \max_i\{f_i t_i x_{ij}\} \qquad \varphi=1,2,3..,N$$

(8) $\sum_{i=1}^{n} x_{ij} \leq \frac{T\sum_{p=1}^{m} y_{jp} w_p}{t_j}$ φ=1,2,3,..,N

(9) $x_{ij} \in \{0,1\}$ $\forall$ i, j

(10) $y_{jp} \in \{0,1\}$ $\forall$ i, j

(11) $\boldsymbol{\eta} = \frac{1}{N} - \Delta$ $\boldsymbol{\acute{\eta}} = \frac{1}{N} + \Delta$

(12) Where , $0 \leq \Delta \leq \frac{1}{N}$

Equation (1) is the objective function that minimizes internal transportation (by considering load capacity), inter loop transportations and the penalty cost which is proportional to AGV capacity. Constraints (2) and (3) ensure that each station is assigned to one loop and in each loop there is only one AGV. Constraint (4) ensures that at least 2 stations are in each loop. Constraints (5) and (6) balance the work load based on the work capacity by defining η and ή. Equation (7) finds the bottleneck station in each loop. Constraint (8) assesses the feasibility of serving stations in loops by considering AGV capacity and bottleneck station, and constraints (9) and (10) enforce the binary constraint on decision variables. Equations (11) and (12) show that works are distributed based on the number of loops. Δ is defined to avoid impossibility of assigning exact number of work load to each loop (see [11] and [18] details).

Model I is a nonlinear binary mathematical programming model for multiple-load AGV system in Tandem configuration. In next the subsection, the objective function is modified to have more efficient solving procedure.

Modifying objective function

The first step is to simplify the mathematical model by linearizing the decision variables;

(1) $z_{ijkl} = x_{ij} \times x_{kl}$

The modification is done by adding decision variable z_{ijkl} to model and its necessary constraints.

Model II

(2) $$\mathrm{Min}\ Z = \underbrace{\sum_{j=1}^{N} \sum_{l=1}^{N} \left(\frac{\sum_{i=1}^{n} \sum_{k=1}^{n} f_{ik} z_{ijkl}}{\sum_{p=1}^{m} y_{jp} w_p} \right)}_{i \neq k, j=l} + \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{N} \sum_{k=1}^{n} \sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i \neq k, j \neq l} + \sum_{j=1}^{N} \sum_{p=1}^{m} y_{jp} w_p \lambda_p$$

(2) $\sum_{j=1}^{N} x_{ij} = 1$ $i = 1,2,3 \ldots, n$

(3) $\sum_{p=1}^{m} y_{jp} = 1$ $j = 1,2,3 \ldots, N$

(4) $\sum_{i=1}^{n} x_{ij} \geq 2 \quad j = 1,2,3 \dots, N$

(3) $\eta\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i\neq k} \times \sum_{p=1}^{m} y_{jp} w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp} w_p}\right) \leq \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik} z_{ijkj}}_{i\neq k} \quad j=1,2,3 \dots, N$

(4) $\acute{\eta}\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i\neq k} \times \sum_{p=1}^{m} y_{jp} w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp} w_p}\right) \geq \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik} z_{ijkj}}_{i\neq k} \quad j=1,2,3 \dots, N$

(7) $\acute{t}_j = \max_i \{f_i t_i x_{ij}\} \quad j=1,2,3..,N$

(8) $\sum_{i=1}^{n} x_{ij} \leq \frac{T\sum_{p=1}^{m} y_{jp} w_p}{\acute{t}_j} \quad j=1,2,3,..,N$

(5) $x_{ij} + x_{kl} - 2z_{ijkl} \geq 0 \quad \forall\, i, j, k, l$

(6) $\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} z_{ijkl} = \frac{n(n-1)}{2} \quad k > i$

(9) $x_{ij} \epsilon \{0,1\} \quad \forall\, i, j$

(10) $y_{jp} \epsilon \{0,1\} \quad \forall\, i, j$

(7) $z_{ijkl} \epsilon \{0,1\} \quad \forall\, i, j, k, l$

As it is mentioned, according to [23] constraints (5) and (6) are necessary for linearization.

The Lagrangian relaxation approach proposed by [24] is applied to relax some constraints, in this approach constraints which are hard to become feasible are rewritten. Consider Y as an objective function and $g_i(x)$ as constraints:

$\text{Min } Y = f(x)$

(8) S.t:

$g_i(x) \leq b_i, \forall i$

To remove inequality constraints and rewrite them in objective function the following change is made:

(9) $\min y = \{f(x) + rk(x)\}$

In above mentioned equation *r* is penalty coefficient and *k(x)* is penalty function;

(10) $k(x) = \sum_{i=1}^{m} \max\{(g_i(x) - b_i), 0\}^2$

Now problem (8), (9) are equivalent. The same modification will be made for constraints (4) and (8). Therefore the final objective function is formulated as Model III.

Model III

$$\begin{aligned}
\operatorname{Min} Z = &\underbrace{\sum_{j=1}^{N}\sum_{l=1}^{N}\left(\frac{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik} z_{ijkl}}{\sum_{p=1}^{m} y_{jp} w_p}\right)}_{i\neq k,\, j=l} + \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i\neq k,\, j\neq l} + \sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp} w_p \lambda_p \\
&+ r\sum_{j=1}^{N}\left(\operatorname{Max}\left[\left\{\eta\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i\neq k} \times \sum_{p=1}^{m} y_{jp} w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp} w_p}\right)\right.\right.\right. \\
&\left.\left. - \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik} z_{ijkj}}_{i\neq k}\right\}, 0\right]^2 \\
&+ \operatorname{Max}\left[\left\{\underbrace{\sum_{i=1}^{n}\sum_{k=1}^{n} f_{ik} z_{ijkj}}_{i\neq k} - \eta\left(\frac{\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{N} f_{ik} z_{ijkl}}_{i\neq k} \times \sum_{p=1}^{m} y_{jp} w_p}{\sum_{j=1}^{N}\sum_{p=1}^{m} y_{jp} w_p}\right)\right\}, 0\right]^2 \\
&\left. + \max\left[\left\{\sum_{i=1}^{n} x_{ij} \frac{T\sum_{p=1}^{m} y_{jp} w_p}{\acute{t}_j}\right\}, 0\right]^2\right)
\end{aligned}$$

Note that constraints (2), (3), (4), (7), (9), (5), (6) and (7) remain unchanged.

SIMULATED ANNEALING ALGORITHM

As can be seen from model formulation the proposed optimization model is a quadratic problem which is classified as NP-hard problem, thus simulated algorithm could be efficient to solve this problem in a timely manner. Simulated Annealing (SA) algorithm proposed by [25] is a metaheuristic algorithm to approximate global solutions. Algorithm 1 clarifies the algorithm steps clearly:

Algorithm 1:

Input: Cooling schedule.

s = s_0; /∗ Generation of the initial solution ∗/

T = β* f(s_0); /∗ Starting temperature ∗/

Repeat

Repeat /∗ At a fixed temperature ∗/

Generate a random neighbor s′ ;

ΔE = f (s′) − f (s) ;

If ΔE ⩽ 0 **Then** s = s′ /∗ Accept the neighbor solution ∗/

Else Accept s′ with a probability $e^{-\Delta E/T}$;

Until Equilibrium condition

/∗ e.g. a given number of iterations executed at each temperature T∗/

T = α*T; /∗ Temperature update ∗/

Until Stopping criteria satisfied /∗ e.g. T < T_{min}∗/

Output: Best solution found.

This is a probabilistic algorithm that avoids trapping in local optimum by accepting worse solutions. In this paper, by progressing in solution procedure, the rejection gets harder and the proposed SA reaches to one of the local optimum solutions.

The proposed algorithm starts with initial temperature T_0 which is calculated by the multiplying the initial solution (s_0) objective function to equalization coefficient (β). In the next step, the algorithm generate a random feasible solution, this process is explained in section generating neighborhood, if the generated solution answer gets better than the previous one, the answer will be accepted. Otherwise, the proposed algorithm generates a random probability λ and checks the following equation$\lambda \geq e^{-\frac{\Delta E}{T}}$, if the inequality satisfies then the worse answer is accepted otherwise it will be rejected. This algorithm iterates until termination criteria meets, which can be final temperature or optimal solution. The necessary parameters of the algorithms are tuned using [11] which are as follows:

TABLE I. SA's algorithm parameters

α	β	iteration
0.98	0.5	20

Initial solution

The solution structure is defined based on problem, which is a three rows matrix; the first row introduces the name of the station, the second row presents the loop name, where the above station is assigned, and the last row presents the type of AGV in that loop. TABLE II. depicts solution structure in which stations 1, 4 and 5 are in loop 1 that means they are served by single load AGV and the rest of stations are in loop 2 and served by three load AGV.

TABLE II. Solution structure

Station	1	2	3	4	5	6
Loop	1	2	2	1	1	2
AGV capacity	1	3	3	1	1	3

Generating neighborhood

The proposed algorithm first select a station and randomly assigns the station to a new loop, then revises the third row and select the best type AGV assignment. As an example in Figure II, stations 1, 4 and 5 are in loop 1 where an AGV with 1 load capacity is working, and station 2,3 and 6 are in loop 2 where a two load AGV doing transportation job. At first step station 4 will move to loop 2. The next happen in Figure III, where loop one AGV capacity will be the same but the second loop AGV capacity increase to three.

Station	1	2	3	4	5	6
Loop	1	2	2	1	1	2
AGV capacity	1	2	2	1	1	2

Station	1	2	3	4	5	6
Loop	1	2	2	2	1	2
AGV capacity	1	2	2	2	1	2

FIGURE II. First step of generating neighborhood

Station	1	2	3	4	5	6
Loop	1	2	2	2	1	2
AGV capacity	1	2	2	2	1	2

Station	1	2	3	4	5	6
Loop	1	2	2	2	1	2
AGV capacity	1	3	3	3	1	3

FIGURE III. Second step of generating neighborhood

Termination criteria

The proposed SA algorithm stops searching for solution when one of the termination criteria is satisfied;

1. Objective worsening (reaching optimal solution)
2. Iteration limit (final temperature)

In next section, we use some illustrative examples to check and validate the proposed simulated annealing algorithm.

NUMERICAL EXAMPLE

To validate the performance of the presented mathematical model and SA algorithm five numerical examples with different dimensions are tested (See TABLE III. For details). These examples are used in [11] as well.

TABLE III. Numerical examples parameters

instances	Parameters					
	n	N	η	η´	t	T
P1	6	2	0.25	0.75	{1,1,1,1,1,1}	2000
P2	7	2	0.2	0.8	{2,1,5,1,3,1,1}	1000
P3	8	2	0.25	0.75	{2,1,5,1,3,1,1,3}	4000
P4	9	2	0.2	0.8	{2,1,5,1,3,1,1,3,4}	4000
P5	10	2	0.15	0.85	1 ∀ i	5000
P6	12	3	0.09	0.57	1 ∀ i	5000

To conduct result analysis, three parameters should be defined which are w_p, r and λ. w_p is the performance coefficient of AGV with type P. This coefficient does not have a linear relationship with AGV's capacity so variation in loading and unloading policy can make difference. To determine this coefficient a uniform distribution is used for determining AGV capacity. The performance coefficient is 1 for AGV with one load capacity, AGV with 2 load capacity can have value in between [1.5, 2] and value for AGV with 3 load capacity could be a random number between [2.3, 3] [16]. After calculation, these values for AGV with 2 and 3 load capacity are 1.8 and 2.65 respectively. Parameter λ which is a penalty associated to increasing AGV capacity, and *r* which is Lagrangian relaxation are considered as 1.

According to above information and TABLE III. , the result analysis is as follows:

TABLE IV. Results of p1-p5 by λ=1

	GAMS	Time	Best SA	Max deviation	Mean Time	Enumeration
P1	901.82	3	901.82	0	24.1	901.82
P2	126.8	5	126.8	0	16.4	126.8
P3	181.77	25	178.39	0	19.7	178.39
P4	2337.65	50	2254.6	0	29	2254.6
P5	2948.77	67	2769.7	118.5	26	2769.7
P6	*	*	4725.9	38	29	4709.7

- STOP AFTER 2400 SECONDS

Enumeration presents the optimal solution of the problems. As TABLE IV. represents, by increasing the problem size, the complexity of the problem will increase and it may cause in trapping to local optimum. The GAMS solver trapped in local optimum after solving the first 2 problems. It was not able to solve P6 after limited time and stopped. While the proposed SA algorithm in the present paper solved first 4 problems in short time and reached to the global optimum. It also successfully solved *p5* in 26 seconds and found a local optimum for example *P6*. Even though increasing the size of problem increased the complexity, but the proposed algorithm could successfully solve the problems more effectively limited time.

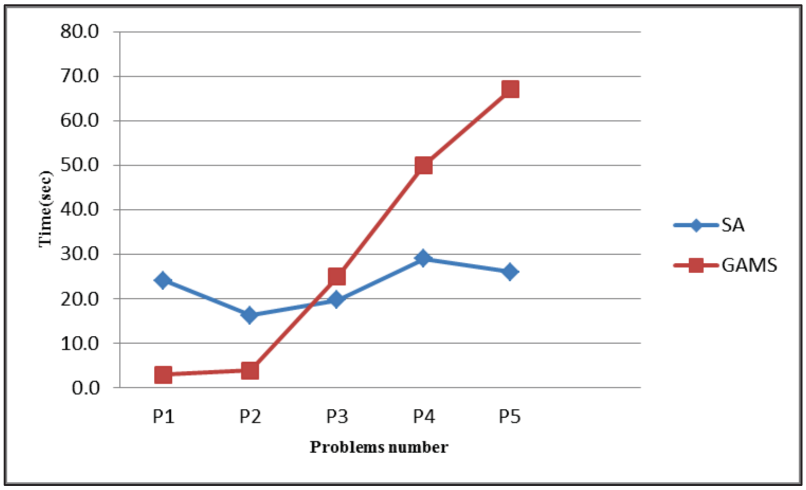


FIGURE IV. Analyze p1-p5 solving time

Figure IV indicates that increasing the size of problem will increase the complexity and GAMS solving time growing exponentially but the proposed SA algorithm solving time shows slight fluctuation. Figure V also compares the set of best results for both solving algorithms. As it is shown, in problem P4 and P5 the SA algorithm found a better solution comparing to GAMS.

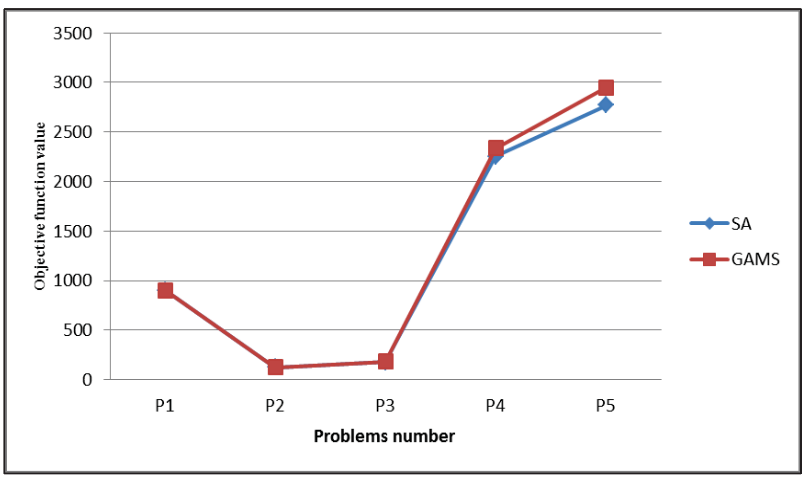


FIGURE V. Analyze p1-p5 solving time

CONCLUSION

This paper mainly focuses on modeling multiple load AGVs and conducts a review on the related literature which indicated that few number of researches have been done on this topic. Researches mentioned that multiple load AGVs cause decrease in average throughput time and fleet size, so this paper proposed a non-linear integer mathematical programming model for multiple-load AGV in tandem configuration. A simulated annealing algorithm was proposed to solve the problem in efficient time. Result analysis section reveals that increasing the problems dimension significantly increases the GAMS solving time, however the proposed SA algorithm solves the problem in short time and generates better solutions. In future research a new approach for modeling multiple loads AGV can be considered which minimizes the travel time and loading waiting time.

REFERENCES

[1] M. Alizadeh, J. Ma, M. Marufuzzaman, and F. Yu, “Sustainable olefin supply chain network design under seasonal feedstock supplies and uncertain carbon tax rate,” *Journal of Cleaner Production*, vol. 222, pp. 280–299, 2019.

[2] B. Rahimikelarijani, A. Abedi, M. Hamidi, and J. Cho, “Simulation modeling of Houston Ship Channel vessel traffic for optimal closure scheduling,” *Simulation Modelling Practice and Theory*, vol. 80, 2018.

[3] H. Behbahani, S. Nazari, M. J. Kang, and T. Litman, “A conceptual framework to formulate transportation network design problem considering social equity criteria,” *Transportation Research Part A: Policy and Practice*, vol. 125, pp. 171–183, 2019.

[4] M. Pirhooshyaran and L. V Snyder, “Optimization of inventory and distribution for hip and knee joint replacements via multistage stochastic programming,” in *Modeling and optimization: Theory and applications*, 2017, pp. 139–155.

[5] A. Abedi and W. Zhu, “An optimisation model for purchase, production and distribution in fish supply chain–a case study,” *International Journal of Production Research*, vol. 55, no. 12, pp. 3451–3464, 2017.

[6] A. Gharehyakheh and R. Tavakkoli-Moghaddam, “A fuzzy solution approach for a multi-objective integrated production-distribution model with multi products and multi periods under uncertainty,” *Management Science Letters*, vol. 2, no. 7, pp. 2425–2434, 2012.

[7] A. Asef-Vaziri, H. Jahandideh, and M. Modarres, “Loop-based facility layout design under flexible bay structures,” *International Journal of Production Economics*, vol. 193, pp. 713–725, 2017.

[8] I. F. A. Vis, “Survey of research in the design and control of automated guided vehicle systems,” *European Journal of Operational Research*, vol. 170, no. 3, pp. 677–709, 2006.

[9] R. Zanjirani Farahani and E. Miandoabchi, “Solving the tandem AGV network design problem using tabu search: Cases of maximum workload and workload balance with fixed and non-fixed number of loops,” *Journal of Industrial Engineering International*, vol. 5, no. 8, pp. 37–51, 2009.

[10] Y. A. Bozer and M. M. Srinivasan, “Tandem configurations for AGV systems offer simplicity and flexibility,” *Industrial Engineering*, vol. 21, no. 2, pp. 23–27, 1989.

[11] M. B. A. R.Tavakkoli-Moghaddam H.Kazemipoor, A.Salehipour, “Partitioning machines in tandem AGV systems based on ‘balanced flow strategy’ by simulated annealing,” *The International Journal of Advanced Manufacturing Technology*, vol. 38, no. 3–4, pp. 355–366, 2008.

[12] R. Zanjirani Farahani, G. Laporte, E. Miandoabchi, and S. Bina, “Designing efficient methods for the tandem AGV network design problem using tabu search and genetic algorithm,” *The International Journal of Advanced Manufacturing Technology*, vol. 36, no. 9, pp. 996–1009, 2008.
[13] T. Y. ElMekkawy and S. Liu, “A new memetic algorithm for optimizing the partitioning problem of tandem AGV systems,” *International Journal of Production Economics*, vol. 118, no. 2, pp. 508–520, 2009.
[14] S. Rezapour, R. Zanjirani-Farahani, and E. Miandoabchi, “A machine-to-loop assignment and layout design methodology for tandem AGV systems with single-load vehicles,” *International Journal of Production Research*, vol. 49, no. 12, pp. 3605–3633, 2010.
[15] Y. A. Bozer and M. M. Srinivasan, “Tandem configurations for automated guided vehicle systems and the analysis of single vehicle loops,” *IIE transactions*, vol. 23, no. 1, pp. 72–82, 1991.
[16] Y.-C. Ho * and P.-F. Hsieh, “A machine-to-loop assignment and layout design methodology for tandem AGV systems with multiple-load vehicles,” *International Journal of Production Research*, vol. 42, no. 4, pp. 801–832, 2004.
[17] A. Salehipour, H. Kazemipoor, and L. Moslemi Naeini, “Locating workstations in tandem automated guided vehicle systems,” *The International Journal of Advanced Manufacturing Technology*, vol. 52, no. 1, pp. 321–328, 2011.
[18] B. Rahimikelarijani, M. Saidi-Mehrabad, and F. Barzinpour, “A mathematical model for multiple-load AGVs in Tandem layout,” *Journal of Optimization in Industrial Engineering*, 2018.
[19] M. Alizadeh, I. Mahdavi, N. Mahdavi-Amiri, and S. Shiripour, “A capacitated location-allocation problem with stochastic demands using sub-sources: An empirical study,” *Applied Soft Computing*, vol. 34, pp. 551–571, 2015.
[20] M. Ozden, “A simulation study of multiple-load-carrying automated guided vehicles in a flexible manufacturing system,” *The International Journal Of Production Research*, vol. 26, no. 8, pp. 1353–1366, 1988.
[21] J. R. van der Meer, *Operational control of internal transport*. Erasmus University Rotterdam, 2000.
[22] K. Kim, B. Chung, and M. Jae, “A design for a tandem AGVS with multi-load AGVs,” *The International Journal of Advanced Manufacturing Technology*, vol. 22, no. 9, pp. 744–752, 2003.
[23] N. Christofides, A. Mingozzi, and P. Toth, “Contributions to the quadratic assignment problem,” *European Journal of Operational Research*, vol. 4, no. 4, pp. 243–247, Apr. 1980.
[24] H. S. MS.Bazaraa, *Nonlinear programming*, Second edn. Wiley, NY, 1993.
[25] S. Kirkpatrick, J. Gelatt, and M. Vecchi, “Optimization by simulated annealing,” *Science*, no. 220, pp. 671–680, 1983.

DECISION SCIENCES INSTITUTE

Project Risk Management: An Integrated Decision-making Support Approach

Fen Wang
Central Washington University
Email: fen.wang@cwu.edu

Efosa C. Idemudia
Arkansas Tech University
Email: eidemudia@atu.edu

Mahesh (Michael) S. Raisinghani
TWU College of Business
Email: mraisinghani@twu.edu

Natalie Lupton
Central Washington University
Email: natalie.lupton@cwu.edu

Xinbao Wang
City of Seattle
Email: xinbao.wang@seattle.gov

ABSTRACT

The overall process of Risk Management (RM) is to identify, analyze and monitor risks as well as to design and implement risk treatments. For this paper, we focus on and review the characteristics of a set of specialized computer-based systems, Decision-Making Support Systems (DMSS), which is a valuable tool for supporting RM processes within the context of Project Management. We report six illustrative application examples of DMSS in the Project Risk Management (PRM) domain and develop two extended and adapted DMSS models for PRM factors and phases. Along the process, a theoretical mapping framework between the generic decision-making processes and PRM processes is identified and depicted. We have also developed an integrated application framework for applying relevant DMSS tools and capabilities to the specific PRM sub-processes. Our aim is to provide a well-structured introduction to DMSS tools to efficiently and effectively facilitate PRM. We also aim to apply a linkage, creating awareness for project and risk managers of the usefulness and value of a variety of DMSS tools.

KEYWORDS: Project Risk Management, Decision-Making Support Systems, Expert Systems, AHP, Business Intelligence Tools, Business Analytics Tools

INTRODUCTION

The world has seen an increase of large complex projects that involve vast amounts of time, money and resources. The fast development of technology and global competition are among the significant reasons that project management (PM) is getting more attention (Schwalbe, 2016). The success parameters for any project named the triple constraints are on-time completion, completion within a specified budget and completion with satisfying performance. The main barriers for achievement are the changes in the project environment and when the problem multiplies with the size of the project because uncertainties in project outcomes increase (Dey, 2001; 2010).

To better understand, identify, and analyze such uncertainties and more effectively mitigate their potential negative consequences, the field of Risk Management (RM) has emerged in the last 40 years. RM refers to anticipating the occurrence of potential negative events, estimating effects and elaborating and implementing suitable actions against them. RM methods are used in several relevant fields such as project management (Raz and Michael, 2001), financial investments (Aebi et al., 2012), hazards management (Liu et al., 2016) and business continuity management (Torabi et al., 2014) among others.

In this paper, we focus on the field of Project Risk Management (PRM) within the context of RM and investigate how an integrated decision-making support approach could facilitate more efficient and effective PRM. Specifically, we introduce and review the characteristics of a set of specialized computer-based systems named Decision-Making Support Systems (DMSS) as valuable tools for supporting PRM. We aim to provide a concise but well-structured overview of DMSS tools and the PRM process, and its application linkage for creating awareness in project managers and risk managers alike of the usefulness and value of DMSS tools.

A REVIEW OF PRM PROCESS

PRM is the overall process to identify, analyze and monitor risks as well as to design and implement risk response strategies, including analyzing and monitoring risks throughout the life cycle of a project (NIST, 2002; Schwalbe, 2016). PRM is a field of study and practice established and matured in the last 40 years with the overall aim to conceptualize, assess and manage risks. The concept of risk has been defined in several qualitative and quantitative forms, but all of them share the notion of an undesirable uncertainty event which would cause damage on humans and/or assets valued by humans (Aven, 2016).

According to Raz and Michael (2001), RM is an essential process that accompanies the entire project life cycle throughout the planning, execution, control and completion phases. In particular, it was found that banks with adequate governance structures on RM had a better performance than those with weak governance structures (Idem, 2001). Highlighted, are the usefulness of the RM process for addressing natural hazards such as earthquakes, hurricanes, volcanic eruption, river flood, and landslide among others (Liu et al., 2016). According to our research, there are classifications of hazard interactions for complementing the current Multi-Hazard Risk Assessment method (Marzocchi et al., 2009). There is support for the implicit and required utilization of RM techniques as well as of the Business Impact Analysis (BIA) one, as part of the Business Continuity Management (BCM) process with the aim to increase the business organizational resilience after disruptions caused by the realization of risks (Torabi et al., 2014). Torabi et al. (2014) elaborated a comprehensive BIA framework supported with multi-attribute decision making (MADM) methods based on the notion of the concept risk appetite defined as “the amount of risk that an organization is willing to pursue or retain” (Idem, 2014; p. 314).

A RM process usually covers three sub-processes: Risk Assessment, Risk Mitigation, and Risk Evaluation (NIST, 2002). An updated concept of RM in the context of information technology standards (NIST, 2012) covers the following phases: Risk Framing, Risk Assessment, Risk Responses, and Risk Monitoring. Both traditional and updated concepts can be mapped as follows: the traditional Risk Assessment sub-process includes updated Risk Framing and Assessment sub-processes; the traditional Risk Mitigation sub-process corresponding to updated Risk Responses and the traditional Risk Evaluation sub-processes corresponding to Risk Monitoring.

The purpose of the Risk Framing sub-process is to establish the high-level organizational context for implementing the right level, no under or over one, of a RM process (NIST, 2012). A RM strategy is delineated, and decisional boundaries are established. A high-level analysis of the organizational context including priorities, assumptions, trade-offs, as well as selection of risk approaches, models and methods are established. In particular, a Risk Model is a core operational concept because it defines the specific elements to be included in the analysis of risks and their interrelationships. In NIST (2002; 2012) the expected elements are: threats; sources of threats; threats scenarios; threats shifting; assets vulnerabilities; severities of vulnerabilities; likelihood of threat exploitation on vulnerabilities; harm impact on assets; likelihood of occurrence of harm impact; final organizational risk as the product of the harm impact and its likelihood of occurrence and risk uncertainty level. This first sub-process is critical to establish a formal RM process in the organization and is executed as a preparation sub-process for the next one of Risk Assessment.

The purpose of the Risk Assessment sub-process (NIST, 2002; 2012) is to determine the specific organizational risks through the detailed analysis of threats, vulnerabilities, impacts, likelihoods and uncertainties (i.e., by using the agreed Risk Model). Qualitative, quantitative or hybrid models are used for specific and agreed RM approaches along with method and type analysis. This sub-process can be considered the core essential sub-process because a wrong analysis of organizational risks will lead to failed actions and countermeasures for risk treatments with potentially severe economic and organizational impacts. In some cases, the harm can also have an impact on human beings; thus, this sub-process becomes essentially critical.

The next sub-process is Risk Mitigation/Risk Responses (NIST, 2002; 2012). Its purpose can be stated as "*prioritizing, evaluating, and implementing the appropriate risk-reducing controls recommended from the risk assessment process*" (NIST, 2002; p. 27) for each organizational risk must be developed, evaluated and implemented for the selected alternative courses of action for coping with the risk with the organizational risk strategy (NIST, 2012). This third sub-process is also very important because it is usually economically impractical or impossible to eliminate organizational risks; thus, adequate and economically viable countermeasures must be proposed and implemented. (NIST, 2002; 2012). The risk mitigations countermeasures can be one of the following ones according to NIST (2002): 1) risk acceptance; 2) risk avoidance; 3) risk limitation; 4) risk planning; 5) risk acknowledgement and research treatment and 6) risk transference. Naturally, every organization must be treated with different risk mitigation actions based on economic and impact analysis.

The final sub-process is Risk Evaluation/Risk Monitoring (NIST, 2002; 2012). The purpose of this sub-process is to conduct ongoing evaluation and monitoring of the implemented risk mitigation countermeasures, as well as to watch over organizational and external changes that may reduce the effectiveness of the current mitigation actions. In this sub-process, the compliance or adherence analysis to best RM practices and external regulations is conducted.

Table 1: PRM Process and Sub-Processes Profile

SUB-PROCESS	INPUTS	TASKS	OUTPUTS
Risk Assessment / Risk Framing	• Risk Context (assumptions, constraints, priorities and tradeoffs, tolerance and uncertainty levels)	• Identify the purpose of the assessment • Identify the scope of the assessment • Identify the assumptions and constraints associated with the assessment • Identify the sources of information to be used as inputs to the assessment • Identify the risk model and analytic approaches to be employed during the assessment	• Risk management strategy • Risk assessment methodology (risk steps, risk model, assessment approach, and analysis approach).
Risk Assessment / Risk Assessment	• Inventory of assets • History of attacks • Data from intelligence on attacks • Reports from prior Risk Assessments • Current Controls • Security Requirements • Threats-Sources Profiles • Criticality of assets	• Identify threat sources that are relevant to organizations • Identify threat events that could be produced by those sources • Identify vulnerabilities within organizations that could be exploited by threat sources • Determine the likelihood of identified threat sources • Determine the adverse impacts to organizational operations and assets • Determine information security risks including any uncertainties associated with the risk determinations	• System boundaries • List of Threats-Sources • List of Potential Vulnerabilities • Current and Potential Controls • Likelihood Rating • Impact Rating • Risks and Risks Levels • Recommended Controls
Risk Mitigation / Risk Responses	• Risk Assessment Report	• Prioritize actions • Evaluate recommended control actions • Conduct Cost-Benefit analysis • Select control • Assign responsibility • Develop a safeguard implementation plan • Implement selected controls	• Actions ranking from High to Low • List of possible Controls • Cost-Benefit Analysis • Selected Controls • List of Responsible Persons • Safeguard Implementation Plan • Residual Risks
Risk Evaluation / Risk Monitoring	• Risk Assessment Report • Risk Mitigation / Responses Report • Risk Management Regulations • Best Practices on Risk Management Reports	• Determine the ongoing effectiveness of risk responses • Identify risk-impacting changes to organizational assets and the environments • Conduct audit evaluations • Trace Risk Management best practices and new regulations	• Risk Evaluation / Monitoring Report

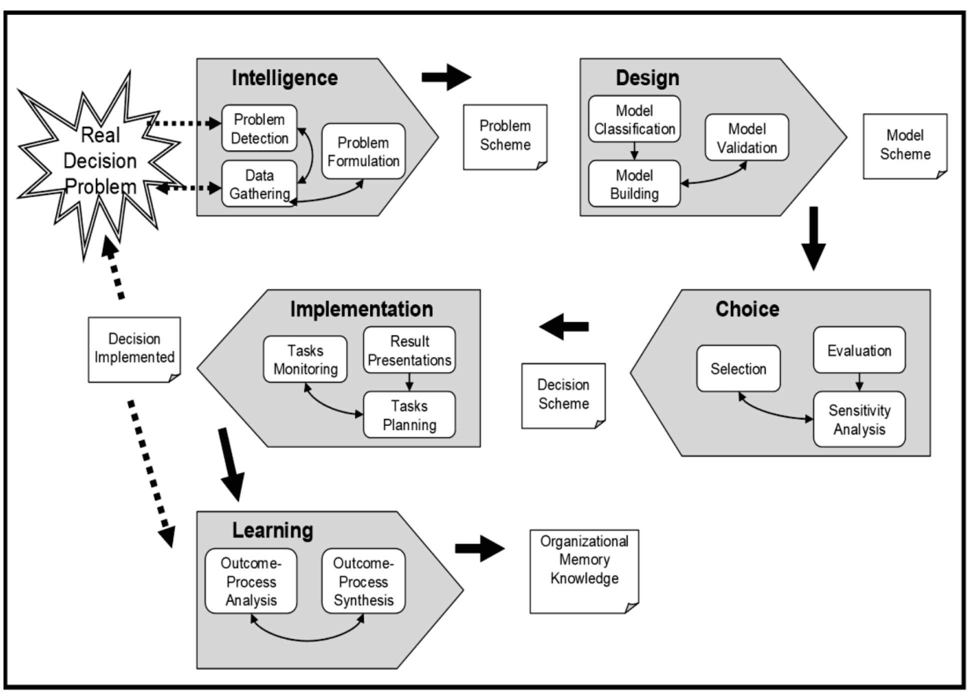
Table 1: PRM Process and Sub-Processes Profile, summarizes the essential characteristics of the RM sub-processes as reported in the consulted sources (NIST, 2002; 2012).

A REVIEW OF DECISION-MAKING PROCESS AND DECISION-MAKING SUPPORT SYSTEMS (DMSS) IN THE PROJECT MANAGEMENT DOMAIN

Decision-Making Support Systems (DMSS) are computer-based systems designed to support some or all phases of a decision-making process (Forgionne et al., 2000). A high-level generic decision-making process (Forgionne et al., 2000; Mora et al., 2014) can be defined as a managerial process of five phases: *intelligence, design, choice, implementation* and *learning*.

In the *intelligence* phase, the decision maker gains a fundamental understanding and acquisition of the general information needed to address the organization's problems or opportunities. In the *design* phase, the decision maker develops a specific and precise model that can be used to systematically examine the discovered problem or opportunity. Using the explicit model to logically evaluate the specified alternatives and to generate recommended actions constitute the ensuing *choice* phase. During the subsequent *implementation* phase, the decision maker ponders the analyses and recommendations, weighs the consequences, gains sufficient confidence in the decision, and implements a final decision. In the final *learning* phase, the decision maker or decision team reviews the process and outcomes with learning purposes (e.g., for reinforcing good practices and discouraging bad ones). Figure 1: A high-level generic decision-making process, illustrates this high-level generic decision-making process with their phases, inputs, and outputs (Mora et al., 2014).

Figure 1: A High-Level Generic Decision-Making Process



DMSS tools have evolved during these last 45 years (Shim et al., 2002). We report the existence of classic and modern DMSS (Forgionne et al., 2005). Classic DMSS are Decision Support Systems (DSS), Executive Information Systems (EIS), and Expert Systems/Knowledge-based Systems (ES/KBS). Modern DMSS are Data Warehouse-based Business Intelligence tools (DW/BI) and Data Mining-based and Business Analytics DMSS (DM/BA), and intelligent DMSS (i-DMSS).

Decision Support Systems (DSS)
According to Sprague (1980) "*… a DSS is an interactive computer-based system which help decision makers utilizing data and quantitative models to solve semi-structured problems*". A DSS offers the following capabilities: what-if scenario analysis, goal-seeking analysis, and sensitivity variables analysis. A DSS is usually used by staff personnel in the organizations, top managers and executive decisions makers. A survey of DSS applications pointed out that the main areas of use are production/operation, marketing, finance and strategic management (Eom & Kim, 2006). An interesting fact is that the underlying model which was quantitative in nature, either deterministic or stochastic have shifted to a hybrid nature (i.e., quantitative and qualitative). Qualitative models are now based on artificial intelligence techniques. Literature reports successful cases of the use of DSS (Sprague & Carlson, 1982; Eom & Kim, 2006).

Executive Information Systems (EIS)
An Executive Information Systems (EIS) according to Rockart and Tracy (1982) *"… is a computer-based system which let to access a common core of data covering key internal and external business variables by time and by business unit*". They discovered in the early '80s that some top executives (i.e., CEO's, CFO's, etc.) were using special information systems, which allowed them to monitor and track key performance indicators of the company.

Afterward, they coined the term EIS. The capabilities typical of EIS are access to summarized information as internal and external key performance indicators usually presented as graph and text-tables displays; analysis through drill-down, roll-up, slice and dice and pivoting operations, and networking communications to bulletin boards. The literature also reports successful cases of EIS usage (Watson et al., 1991; Elam & Leidner, 1995; Rainer & Watson, 1995).

Expert Systems/Knowledge-based Systems (ES/KBS)
According to several researchers in the field, ES/KBS is a computer-based system that exhibits, in a specific domain, a high degree of expertise in problem-solving that is comparable to that of a human expert (Feigenbaum et al., 1988; Liebowitz, 1990). An ES/KBS contains three main components: a knowledge base specialized in a specific domain, an inference algorithm, and a user interface. Two types of problems are usually addressed: classification (e.g., diagnosing, repairing, prediction, interpretation, monitoring and control) or construction (e.g., scheduling, configuration, planning or designing). An ES/KBS offers the following capabilities: intelligent advice, qualitative reasoning, problem-solving assistance, and explanation of advice. Successful cases of ES/KBS are reported from several sources (Feingenbaum et al., 1988; Eom, 1993). The difference between an ES and a KBS is that the former contains the knowledge of a recognized expert in the field and the second contains only knowledge valuable for the organization.

Data Warehouse-based and Business Intelligence Tools (DW/BI)
A data warehouse is a specialized organizational macro database composed usually by a large number of dimensions and usually containing a vast quantity of records (Wixom and Watson,

2001). It is defined as "*subject-oriented, integrated, time-invariant, non-updatable collection of data used to support management decision-making processes and business intelligence*" (March & Hevner, 2007; p. 1031). A data warehouse is the organizational repository of relevant data sets which can be exploited through computer-based systems named Business Intelligence (BI) tools. BI tools aim to support BI activities in an organization refers to the environment and organization relevant data. These tools scan the data for interpretation applying organizational knowledge and producing intelligence (interpreted information useful for a specific organizational context) (March & Hevner, 2007). BI tools emerged as the evolution of EIS (Adam & Pomerol, 2002) by the progress of information technologies to reach all organizational levels in addition to the top management level addressed for EIS; thus, provide the same capabilities. However, they are released to end-users through the concept of dashboards. Dashboards are systems that present a control display of the main indicators of interest for a decision-maker, and the name comes as an analogy of the car dashboards for drivers.

Similar to classic DMSS, DW/BI tools have provided several benefits to users such as: a better understanding of the decision context, increase decision-making productivity, positive change on how people perform tasks, effective decision support, improved on-line analytical processing; availability of high quality/accurate/secure data; high payback and low-cost access (Wixom & Watson, 2001; March & Hevner, 2007; Ramamurthy et al., 2008).

Data Mining and Business Analytics Tools (DM/BA)
These DMSS are based on the application of classic and modern statistical algorithms and heuristic intelligent - both numeric and logical- techniques which discover useful internal relationships among usually a vast quantity of heterogenous sources of data mainly for classification and prediction purposes (Davenport, 2006; Delen & Demirkan, 2013; Watson, 2014). Data Mining (also called Knowledge Discovery in Databases) emerged from the artificial intelligence research community (Fayyad et al., 1996) with the aim to cope with the massive quantity of heterogeneous sources of data available for the availability and convergence of new information and communication technologies (i.e., extensive use of internet, cloud computing, social web tools, RFID devices, mobile phones and tablets, and so on). This technological availability also enabled the convergence of advanced visual-oriented tools including data mining techniques and other data visualization ones in the concept of analytics. Analytics covers a set of statistical and intelligent heuristics algorithms for exploring, analyzing and visualizing huge quantities of heterogeneous sources of data.

According to Delen and Demirkan (2013) and Watson (2013), analytics can be classified in Descriptive, Predictive, and Prescriptive types. Descriptive Analytics refers to the already well-known reporting, OLAP, dashboards/scorecards, and data visualization techniques (i.e., DW/BI DMSS tools). Predictive Analytics covers techniques such as regression analysis, machine learning, and neural networks to elaborate forecasts. Finally, Prescriptive Analytics includes techniques including optimization modeling, simulation modeling, multi-criteria decision modeling, expert systems, and group support systems to obtain optimal or satisfactory solutions (i.e., decisional courses of action) on complex problems.

Intelligent and Integrated DMSS (i-DMSS)
These types of DMSS emerged from the need to combine the capabilities of a single DMSS (i.e., a DSS, an EIS or an ES/KBS) and add human-like intelligent capabilities to normal DMSS. Several research efforts were conducted for this aim.

Turban and Watkins (1986) proposed to integrate DSS and ES/KBS for adding the qualitative modeling capabilities of an ES/KBS to enhance a DSS. Delong and Rockart (1986) proposed the integration of a DSS with an EIS to add analytical capabilities to EIS and to extend the time frame of analysis from past and present to future forecasted data. Elam and Konsynski (1987) developed the prototypes to improve the management model modules of DSS by using intelligent techniques. Turban and Watson (1989), Forgionne (1991) and Forgionne and Kohli (1996) posed the integration of DSS, EIS and ES/KBS. The main aim was to support all the phases of the decision-making processes that stand-alone systems are unable to do. Turban and Watson (1989), and Forgionne (1991) developed a conceptual model proposing how this integration could be developed. Forgionne and Kohli (1996), based on this model, developed an intelligent integrated DMSS and tested its effectiveness and efficiency versus stand-alone systems in an experimental setting. King (1992) suggested the addition of why-analysis to traditional DSS. Why-analysis consists of explanation facilities in quantitative models to track how a numerical variable in the model reached a particular value. Klein and Methie (1990) suggested this integration to access data and knowledge bases in order to effectively support decision makers in complex and ill-structured tasks. Finally, Turban and Aronson (1998) surveyed the potential usage of other intelligent-based technologies including Genetic Algorithms, Fuzzy Logic Systems, Neural Networks, Case-Based Reasoning, and Knowledge-Management Systems. Most modern efforts are oriented to ontology-based systems at present (Chandrasekaran et al., 1999) as instances of these integrations, foundations and successful cases of intelligent DMSS were reported (Gupta et al., 2005)

Classic and modern DMSS have been widely used in organizations because when they are successfully implemented, they provide benefits such as: improved organizational performance, better decision quality, improved communication, enhanced mental models, amplified analytical skills of decision-makers, and reduced decision times among others (Turban & Aronson, 1998; Forgionne & Kohli, 2000; Eom & Kim, 2006; March & Hevner, 2007; Davenport, 2006; Delen & Demirkan, 2013; Watson, 2014). Table 2: Decision-Making Process vs. DMSS Capabilities by DMSS Type, summarizes the main capabilities provided by the different types of DMSS tools regarding the phases conducted in a generic Decision-Making process.

Hence, we propose that classic and modern DMSS can be very adequate tools for supporting a RM process in all domains of application. However, the design and utilization of DMSS, in general, are not seamless organizational paths despite the benefits of deploying them. Technical and organizational challenges must be addressed and solved for achieving a successful DMSS design, building and implementation (Setzerkon et al., 2002; Mora et al., 2013).

Table 2: Decision-Making Process vs. DMSS Capabilities by DMSS Type

DECISION-MAKING PROCESS PHASE	DSS	EIS	ES/KBS	DW/BI	DM/BA
INTELLIGENCE	-	Drill-Down, Roll-Out, Pivoting, Executive Dashboards on Data warehouses	Conceptual knowledge stored and logical inferences for analysis and synthesis tasks	Drill-Down, Roll-Out, Pivoting, Visual Reporting, Multiple Dashboards on Data warehouses	Descriptive Analytics for reporting and visualizing past facts
DESIGN	-	-		-	Predictive Analytics for exploring and elaborating forecasts models
CHOICE	What-if, Goal-Seeking, Sensitivity Analysis on Quantitative Models	-		-	Prescriptive Analytics for searching optimal or satisfactory solutions on Quantitative Models
IMPLEMENTATION	-	Drill-Down, Roll-Out, Pivoting, Executive Dashboards on Data warehouses		Drill-Down, Roll-Out, Pivoting, Visual Reporting, Multiple Dashboards on Data warehouses	Descriptive Analytics for reporting and visualizing past facts
LEARNING	-	-		-	Predictive Analytics for exploring and elaborating predictive models

THE UTILIZATION OF DMSS IN PROJECT RISK MANAGEMENT - SIX ILLUSTRATIVE CASES

In this section, six illustrative application examples of DMSS tools in the PRM domain are reported and summarized in Table 3: Case Studies of Project Risk Management Applications. Based on the comprehensive review, two extended and adapted DMSS models for project risks factors and phases are developed and presented thereafter.

Table 3: Case Studies of Project Risk Management Applications

Authors	Publication Date	Industry Area	Brief Description
Meyer et al.	1992	Underwriting life insurance	Linkage between business strategy and development of large-scale expert systems in the context of project selection and evolution
Dey	2001	A cross-country petroleum pipeline project	A quantitative approach to through analytic hierarchy process and decision tree analysis
Ngai, & Wat	2005	E-commerce (EC) development projects	A Web-based prototype fuzzy decision support system is designed and developed
Dey	2010	Oil pipeline construction project	Integrated framework using AHP and risk map for managing project risks
Fang & Marle	2012	A project of staging a musical show in Paris, France	An integrated DSS framework and simulation-based model
Choi et al.	2016	Industrial-based business systems, and operational risk management	Challenges and opportunities of big data analytics in industrial systems

Areas such as life insurance are driven by business strategy to provide fast, accurate decisions as a competitive advantage. Meyer et al. (1992) diagnosed the flow of information and the process of deciding on life insurance underwriting and found signification opportunities to improve the decision-making process. They evaluated two expert systems at two companies, John Hancock and Lincoln National. John Hancock approved 43% of applications within 24 hours, reducing expenses and permitting underwriters to focus on more complex cases. Lincoln

National formed a new organizational unit and a product that could be sold to other insurance companies. Both expert systems support the companies' strategic plans.

Dey (2001) utilized a quantitative approach of DMSS that incorporates AHP and decision tree analysis for PRM. Some examples of RM and uncertainty that affect projects are size of the project, the complexity of the project, the speed of project's constructions, the location of the project, lack of experience relating to the project, external factors that are beyond human control, external causes which limit resources availability, and the project's degree of unfamiliarity. To address all the complexity in PRM, using a case application of a cross-country petroleum pipeline project in India, Dey (2001) developed a project management model that enables the project management team to establish strong and adequate relationships among all the phases of the project. Some of the benefits of the project management model developed are: (1) to help forecast project achievement to build confidence of project team; (2) to help to make decision objectively with the help of reliable and accurate database; (3) to provide accurate and adequate information for effective and efficient project management; (4) to provide formal risk management method to help to manage projects relating to time, cost, and quality; (4) to help establish close co-operation among top executives and members of project teams; (5) to help identify all risk factors; (6) to help determine and compute the probability and severity of all risk factors; (7) to generate various alternative responses with cost implications for quantifying risks and (8) to derive monetary values for each alternative in a decision tree framework.

Ngai and Wat (2005) discussed a fuzzy decision support system (FDSS) that was a web-based prototype designed and developed to assist e-commerce project managers in identifying potential risk factors and the corresponding project risks. Evaluators can use the risk evaluation checklist and linguistic terms to evaluate the risk level by using a proposed risk analysis model incorporated into the FDSS. Although the results of an evaluation showed that the prototype performed to expectations, further research is needed to develop different ‘‘weightings’’ for different evaluators.

Dey (2010) used AHP to identify project level risks and select the least risky project. The method is applied to a cross-country oil pipeline project in western India comparing four pipeline routes. Dey attempted to capture both business risks and operational risks as determined by a team of experts. The author noted that although the method is tedious and time-consuming, the combined AHP and risk map is a useful technique for risk analysis and potential response planning.

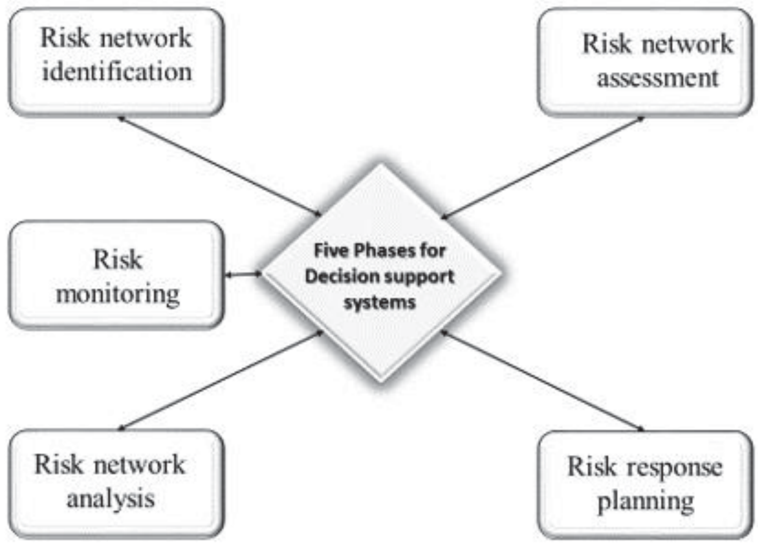
Fang and Marle (2012) argued that to effectively and efficiently manage the complexly interrelated risks associated with a project, that the DMSS should be able to integrate the multiple dimensions of risks (i.e., probability and impact) and to bring the modeling of risk interaction into the PRM process. They discussed that a simulation-based risk network model (i.e., integrated DMSS framework) for decision support in PRM should include these five phases: (1) risk network identification; (2) risk network assessment; (3) risk network analysis; (4) risk response planning and (5) risk monitoring and control. Some of the main advantages of the integrated DMSS framework are: (1) to re-evaluate risks; (2) to priorities risks; (3) to suggest and test mitigation actions and (4) to support top management in making better decisions relating to risks response actions.

Choi et al. (2016) discussed challenges in applying big data analytics to business applications. The authors presented an extensive references list and review of the literature to conclude that there is a myriad of opportunities; however, there are also significant challenges including cost, value, data quality and systems security in industrial system projects.

Based on the preceding discussions and comprehensive literature review of DMSS in PRM, we extend the earlier research work by developing the DMSS models for project risk factors as shown in Figure 3: DMSS model for project risk management phases. The arrows in Figure 2: DMSS model for project risk factors, posits that decision support systems must adapt to all its stakeholders and risks. The arrows in Figure 3 are also multidirectional and indicate that project managers and top management should have the skills, tools, and resources to adapt to all the risks as shown.

Figure 2: DMSS model for project risk factors

Figure 3: DMSS model for project risk management phases



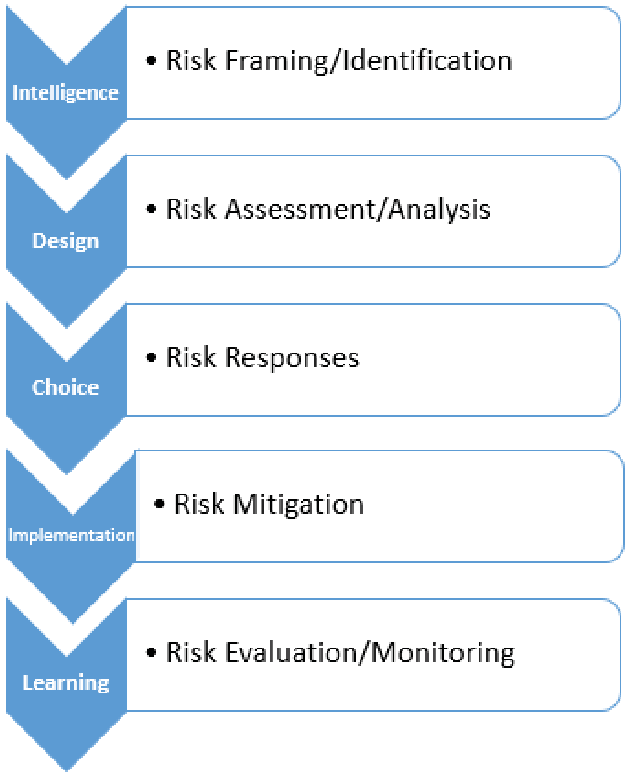
Figures 2 and 3 models address the limitations of current models relating to modeling complexity in PRM. In addition, the extended and adapted DMSS models provide insights and understanding to project managers and top management on how to effectively and efficiently manage project risks. The models also provide insights to project managers and top management on criteria to consider when developing or selecting DMSS tools that objectively measure the different kinds of project risks that exist in project development and assessments. The models enable project managers to save time in project processes relating to designing risk response plans, reducing the cost of dealing with contingencies and assessing and weighing all the risks that are associated with a specific project. Proactive RM can be achieved if project managers have the skills, tools, and resources to adapt and adjust to all the risks as illustrated in Figures 2 and 3.

Figures 2 and 3 equip project managers with the skills, tools, and structural procedure on a robust model to analyze and control project risks. The extended and adapted DMSS models encourage project managers and experts to actively participate in the entire process of DMSS to adapt to the different kinds of risks that exist and understand the relationship among these risks during decision support in PRM. Some of the benefits of the extended and adapted DMSS models are: (1) help to identify; refine and prioritize risks; (2) help to re-assign risk ownership; (3) help to provide effective and efficient mitigation strategies and actions and (4) help to adapt to the different kinds of risks that exist in the dynamic project environment.

DISCUSSIONS AND IMPLICATIONS ON THE LINKAGE BETWEEN DMSS AND RISK MANAGEMENT

In the previous sections, we have summarized the main concepts and characteristics of the RM process, reviewed the types and characteristics of DMSS tools, reported six illustrative application examples of DMSS tools in the PRM domain, and extended the current DMSS models for project risk factors and phases. We also identified the theoretical as well as practical linkages between the DMSS and RM processes. Figure 4: The theoretical mapping framework below illustrates the theoretical mapping framework between the generic decision-making processes and RM processes in the context of information technology standards.

Figure 4: The theoretical mapping framework



As illustrated in Figure 4, the essential risk management process and its sub-processes can be naturally mapped to the established generic decision-making process and the five phases therein. For instance, the *intelligence* phase of the generic decision-making process, when mapped to the essential risk management sub-process of Risk Framing, involves acquiring risk relevant contextual information and identifying possible risks in the identified organizational and/or project context. Next, in the *design* phase of the generic decision-making process, which is mapped to the essential risk management sub-process of Risk Assessment, decision makers or managers would identify or develop a relevant risk model to analyze and assess the identified risks in terms of their probabilities and consequences. Following that, generating recommended actions and responses and putting that into final decisions and strategies correspond to the *choice* phase and *implementation* phase respectively of the generic decision-making process, which is mapped to the essential risk management sub-process of Risk Responses and Risk Mitigation. Finally, in the *learning* phase of the generic decision-making process, which is mapped to the essential risk management sub-process of Risk Evaluation and Monitoring, decision makers or managers review, evaluate, and monitor the process and outcomes of implemented risk mitigation strategies for future organizational learning purpose.

Accordingly, DMSS, which are designed to support some or all phases of a decision-making process, can be considered as valuable and pertinent tools for supporting a Risk Management process. Table 4*:* The application framework of applying DMSS to Risk Management *sub*-processes, illustrates the application framework of applying relevant DMSS tools and capabilities to the specific RM sub-processes.

Table 4: The application framework of applying DMSS to PRM sub-processes

Risk Management Sub-Process	Main Tasks	Suggested DMSS Tools	Relevant DMSS Capabilities
Risk Assessment / Risk Framing	• Establish the high-level organizational context and decisional boundaries for Risk Management	• EIS • ES/KBS • DW/BI • DM/BA	• Drill-Down, Roll-Out, Pivoting, & Executive Dashboards • Visual, interactive, & real-time Reporting • Descriptive Analytics for reporting and visualizing
Risk Assessment / Risk Analysis	• Determine the specific organizational risks through the detailed analysis of threats, vulnerabilities, impacts, likelihoods and uncertainties	• ES/KBS • DM/BA	• Intelligent advice, qualitative reasoning, problem-solving assistance, and logical inferences • Predictive Analytics for exploring and elaborating forecasts models
Risk Mitigation / Risk Responses	• Prioritize, evaluate, and implement the appropriate risk-reducing controls recommended from the risk assessment process	• DSS • EIS • ES/KBS • DW/BI • DM/BA	• What-if, Goal-Seeking, Sensitivity Analysis on Quantitative Models • Prescriptive Analytics for searching optimal or satisfactory solutions on Quantitative Models
Risk Evaluation / Risk Monitoring	• Conduct an ongoing evaluation and monitoring of the implemented risk mitigation countermeasures, and watch over organizational and external changes	• ES/KBS • DM/BA	• Intelligent advice, qualitative reasoning, problem-solving assistance, and explanation of advices • Predictive Analytics for exploring and elaborating predictive models

From an applied perspective, executives and decision makers of worldwide organizations are facing challenges of effective PRM in the current era, and they look for practical DMSS solutions that can help them better identify, assess, mitigate, and monitor organizational risks. This integrative review paper adds to this body of knowledge and helps facilitate additional research and practical endeavors. An integrated and well-structured introduction to and review of the DMSS tools and PRM process were presented. Based on the richness and variety of such applications and research findings, we can present the following recommendations and cautions to executives and decision makers interested in applying relevant DMSS tools to managing organizational risks:

- All industry fields face the challenges of effective and efficient PRM and both the classic and modern DMSS solutions can help them better identify, assess, mitigate, and monitor organizational risks.
- To effectively address the PRM challenges, the classic DMSS solutions need to be equipped with modern and advanced techniques depending on the involved data types, sources, structures, constraints, priorities, uncertainty levels, and other organizational characteristics.
- Practitioners across different industries can utilize the presented application framework to adopt or develop the appropriate DMSS techniques and solutions to reduce risks and derive the intended benefits of making well informed decisions and risk strategies.
- Big data brings a new culture and way of decision making and PRM. Besides technical challenges, managerial challenges of DMSS for risk management in the current era must be addressed to reap the full benefits of that transition.

Hence, our study reveals that the PRM process has experienced drastic changes in its execution by organizations, its theoretical conceptualization by researchers and its practical implications in the current era. The main shift on the PRM process can be stated as an evolution from traditional analytical and statistical techniques with structured data sets in highly predictable and cooperative business environment, to data-driven discovery and highly proactive and creative decision-making utilizing modern and advanced analytical techniques with unstructured and massive data sources to cope with a highly dynamic global business environment in the current era.

CONCLUSIONS ON CONTRIBUTIONS, LIMITATIONS, TRENDS, AND OPEN CHALLENGES

Risk management has become a vital theme in both academia and practice over the past several decades (Wu et al., 2014). It refers to the overall process to identify, analyze and monitor risks as well as the design and implementation of risk strategies to identify risks. In this article, we have introduced and reviewed the characteristics of a set of DMSS as valuable tools for supporting the RM process. We have also reported and analyzed six illustrative application examples of DMSS in the PRM domain. Two extended and adapted DMSS models were developed and presented for project risk factors and phases based on the comprehensive literature review. We have also identified theoretical as well as practical linkages between the DMSS and PRM processes. A theoretical mapping framework between the generic decision-making processes and PRM processes was identified and depicted. In addition, we have developed an integrated application framework for applying relevant DMSS tools and capabilities to the specific PRM sub-processes.

This study also elicits a set of practical recommendations for executives, project managers and decision makers in organizations worldwide. Interpreting RM literature and applying a rich body of knowledge for business practitioners is of great importance. Currently, PM is gaining more attention due to technology and global competition. Successful projects have a greater number of on-time completions, stay within budget and customers are satisfied with the performance for the project. The main barriers to the achievement of completed projects are the changes in the project environment. Challenges may multiply with the size of the project as uncertainties. Given the significance and complexity of PRM and big data phenomenon, we encourage continued research in this area. It is necessary to investigate and establish advanced and updated conceptualizations including frameworks to comprehend this complex yet critical subject.

References

Adam, F., & Pomerol, J. (2002). Critical factors in the development of executive systems-leveraging the dashboard approach n *Decision-making Support Systems: Achievements and Challenges for the New Decade,* G. Forgionne, J. Gupta, & M. Mora (eds.), (305-330). IGI Global.

Aebi, V., Sabato G., & Schmid, M. (2012). Risk management, corporate governance, and bank performance in the financial crisis. *Journal of Banking & Finance*, *36*(12), 3213-3226.

Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research, 253*(1), 1-13.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems, 14*(1), 20-26.

Choi, T., Chan, H., & Yue, X. (2016). Recent Development in Big Data Analytics for Business Operations and Risk Management. *IEEE Transactions on Cybernetics,* Pre-print 1-12.

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review, 84*(1), 1-8.

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems, 55*(1), 359-363.

DeLong, D. W. & Rockart, J. F. (1986). Identifying the attributes of successful executive support system implementation (pp. 190-205). Center for Informations Systems Research, Sloan School of Management, MIT, USA.

Deshmulch, A. & Millet, I. (1999). An Analytic Hierarchy Process Approach to Assessing the Risk of Management fraud. *The Journal of Applied Business Research, 15*(1), 87-102.

Dey, P. (2001). Decision support system for risk management: a case study. *Management Decision, (39)*8, 634-649.

Dey, P. (2010). Managing project risk using combined analytic hierarchy process and risk map. *Applied Soft Computing, 10*(4), 990-1000.

Elam, J. J., & Konsynski, B. (1987). Using artificial intelligence techniques to enhance the capabilities of model management systems. *Decision Sciences, 18*(3), 487-502.

Elam, J. & Leidner, D. (1995). EIS adoption, use, and impact: the executive perspective. *Decision Support Systems, 14*(2), 89-103.

Eom, S. (1996). A survey of operational expert systems in business (1980–1993). *Interfaces, 26*(5), 50-70.

Eom, S. & Kim, E. (2006). A survey of decision support system applications (1995–2001). *Journal of the Operational Research Society, 57*(11), 1264-1278.

Fang, C. & Marle, F. (2012). A simulation-based risk network model for decision support in project risk management. *Decision Support Systems, 52*(3), 635-644.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM, 39*(11), 27-34.

Feigenbaum, E., McCorduck, P. & Nii, P. (1988). *The rise of the expert company*. Time Life.

Forgionne, G. (1991). Decision technology systems: A vehicle to consolidate decision making support. *Information Processing & Management, 27*(6), 679-697.

Forgionne, G. A., & Kohli, R. (1996). HMSS: a management support system for concurrent hospital decision making. *Decision Support Systems, 16*(3), 209-229.

Forgionne, G., Mora, M., Cervantes-Pérez, F., & Kohli, R. (2000). Development of integrated decision-making support systems: a practical approach. *AMCIS 2000 Proceedings*, 83.

Forgionne, G. A., & Kohli, R. (2000). Management Support System Effectiveness: Further Empirical Evidence. *Journal of the Association for Information Systems, 1*(1), 1-39.

Forgionne, G., Mora, M., Gupta, J. N., & Gelman, O. (2005). *Decision-Making Support Systems.* In Encyclopedia of Information Science and Technology, First Edition, M. Khosrow-Pour (Ed.), (pp. 759-765). IGI Global.

Gupta, J., Forgionne, G. & Mora, M. (2005). *Intelligent decision-making support systems*. Springer.

Ibbs, CW. & Kwak, YH. (2000). Assessing Project Management Maturity. *Project Management Journal.* March.

King, D. (1992). Intelligent support systems: art, augmentation, and agents. In *Decision support systems*, R. Sprague and H. Watson (Eds.), 137-159. Prentice-Hall, Inc.

Klein, M. & Methie, L. (1990). *Expert systems: A decision support approach: with applications in management and finance.* Addison Wesley Publishing Company.

Leven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research, 253,* 1–13.

Li, H., Apostolakis, G., Gifun, J., VanSchalkwyk, W., Leite, S. & Barber, D. (2009). Ranking the Risks from Multiple Hazards in a Small Community. *Risk Analysis, 29*(3), 438-356.

Liebowitz, J. (1990). *Expert systems for business & management*. Yourdon Press.

Liu, B., Siu, Y., & Mitchell, G. (2016). Hazard interaction analysis for multi-hazard risk assessment: a systematic classification based on hazard-forming environment. *Natural Hazards Earth System Science, 16*, 629–642.

March, S. T., & Hevner, A. R. (2007). Integrated decision support systems: A data warehousing perspective. *Decision Support Systems, 43*(3), 1031-1043.

Marzocchi, W., Mastellone, M., Di Ruocco, A., Novelli, P., Romeo, E., & Gasparini, P. (2009). Principles of Multi-Risk Assessment: Interactions Amongst Natural and Man-Induced Risks, European Commission, Directorate-General for Research, Environment Directorate, Luxembourg, 72.

Meyer, M., DeTore, A., Siegel, S. & Curley, K. (1992). The Strategic Use of Expert Systems for Risk Management in the Insurance Industry. *Expert Systems with Applications, 5*(1-2), 15-24.

Mora, M., Wang, F., & Gelman, O. (2013). A comparative study on the implementation inhibitors and facilitators of management information systems and integrated decision support systems: A perception of IT practitioners in Mexico. *Information Technology for Development, 19*(4), 319-346.

Mora, M., Phillips-Wren, G., Marx-Gomez, J., Wang, F., & Gelman, O. (2014). The role of decision-making support systems in IT service management processes. *Intelligent Decision Technologies, 8*(2), 147-163.

Ngai, E.W.T. & Wat, F.K.T. (2005). Fuzzy decision support system for risk analysis in e-commerce development. *Decision Support Systems, 40*(2), 235-255.

NIST (2002). NIST Special Publication 800-30 - Risk Management Guide for Information Technology Systems. National Institute of Standards and Technology Gaithersburg, MD, USA.

NIST (2012). Guide for Conducting Risk Assessments - NIST Special Publication 800-30 R1. National Institute of Standards and Technology Gaithersburg, MD, USA.

PMI (1996). A Guide to the Project Management Body of Knowledge. Project Management Institute

Raz, T. & Michael, E. (2001). Use and benefits of tools for project risk management. *International Journal of Project Management, 19*, 9-17.

Rainer Jr, R. K., & Watson, H. J. (1995). The keys to executive information system success. *Journal of Management Information Systems, 12*(2), 83-98.

Ramamurthy, K., Sen, A., & Sinha, A. P. (2008). Data warehousing infusion and organizational effectiveness. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 38*(4), 976-994.

Rockart, J. & Treacy, M. (1982). The CEO goes on-line. *Harvard Business Review, 60*(1), 82-89.

Schwalbe, K. (2016). *Information technology project management* (8th Ed.). Course Technology.

Setzekorn, K., Sugumaran, V., & Patnayakuni, N. (2002). A comparison of implementation resistance factors for DMSS versus other information systems. *Information Resources Management Journal, 15*(4), 48-63.

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision support systems, 33*(2), 111-126.

Sprague Jr, R. H. (1980). A framework for the development of decision support systems. *MIS Quarterly,* 1-26.

Sprague Jr, R. H., & Carlson, E. D. (1982). *Building effective decision support systems.* Prentice Hall Professional Technical Reference.

Torabi, S., Soufi, H. & Sahebjamnia, N. (2014). A new framework for business impact analysis in business continuity management (with a case study). *Safety Science, 68,* 329-323.

Turban, E., & Watkins, P. R. (1986). Integrating expert systems and decision support systems. *MIS Quarterly, 10*(2), 121-136.

Turban, E., & Watson, P. (1989). Integrating Expert Systems, Executive Information Systems and Decisions Support Systems. *DSS 89 Transactions*, 74-82.

Turban, E., & Aronson, E. (1998). *Decision support system and intelligent system*. Prentice-Hall.

Watson, H. J., Rainer Jr, R. K., & Koh, C. E. (1991). Executive information systems: a framework for development and a survey of current practices. *MIS Quarterly, 15*(1), 13-30.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems, 34*(1), 1247-1268.

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly, 25*(1), 17-41.

Wu, D., Chen, S., & Olson, D. (2014). Business intelligence in risk management: some recent Progresses. *Information Sciences, 256,* 1–7.

DECISION SCIENCES INSTITUTE

A Meta-Analysis of Environmental Sustainability in Small & Medium-Sized Enterprises: How has the research evolved over years?

Joseph Bakos
Ithaca College
Email: jbakos@ithaca.edu

Adalberto Orengo
Ithaca College
Email: aorengo@ithaca.edu

Michele Siu
Ithaca College
Email: msiu@ithaca.edu

Narges Kasiri
Ithaca College
Email: nkasiri@ithaca.edu

ABSTRACT

The need to consistently advance with environmentally sustainable practices in today's businesses is crucial. Businesses are demanded to be more environmentally sustainable every day. Our goal is to provide a review of common drivers, barriers and other similarities important to the future of implementing sustainability by small and medium sized enterprises. We investigated patterns in drivers and barriers to obtain the most recent trends by carefully examining 122 studies from 53 different journals published from 2013 to 2019. This meta-analysis can serve as guidelines to gain some insights and knowledge on sustainability strategies in small and medium sized businesses in order to produce a more sustainable and successful future for SMEs as well as evaluates research gaps for future studies.

KEYWORDS

SMEs, Small and Medium Enterprises, Barriers to Sustainability, Sustainability Drivers, Green Practices, Environmental Sustainability, Business Environment, Eco-Friendly

INTRODUCTION

Concerns over the environment have increased over the years, with the 2020 US election coming there are conversations among candidates on plans of enacting new legislation to help the environment. From afar, we might think that pollution exists from large corporations;

however, research shows that SMEs make up 90% of worldwide businesses and are reported to be responsible for about 70% of the global pollution (Tutterow, 2014; Langwell et al., 2014). Large businesses only contribute to around 30% of the pollution in the world today; they are far more qualified than SMEs to achieve environmental sustainability (Aghelie, 2017). SMEs lack experience, data, financial resources and the technical expertise to go green compared to their larger counterparts (Ghazilla et al., 2015). SMEs vary widely in the makeup of the business and commitment to going green (Parker et al., 2009). For example, the issue is that a good environmental plan that would work for a small farm in France would not be sufficient for a machine shop in the United States. Not just the size or make-up of an SME affects a good environmental plan, but it also needs to investigate environmental factors such as the economy and the government policies and regulations. An SME in underdeveloped or transition economies will have a different set of barriers than an SME in a developed country (Silajdzic et al., 2015). Transitional economies, for example, don't have government regulations like developed economies. Their educational institutions don't have the knowledge of green practices and the businesses themselves tend to only worry about profit because of the high failure rates of SMEs in their transitional economies (Silajdzic et al., 2015). Therefore, it would be left to assume that the research on SMEs going green needs to be specific to each of the variables of an SME operates within (Parker et al., 2009). Parker et al.'s (2009) indicates that all the research they found was vague and did not concentrate on the vast array of different SMEs. Now their research has hit the ten-year mark of reviewed journals on environmental sustainability in SMEs, we aim to get the most up to date trends in this research area.

A meta-analysis is a research process in which researchers systematically merge the findings of other independent studies using both qualitative and quantitative data from those studies. We used various meta-analysis studies such as Johnson & Schaltegger (2015) and Parker et al.'s (2009) in order to identify where they left off in their studies and where we should pick up. This meta-analysis is similar to Parker et al.'s(2009) as we found our research could build off of theirs and produce an updated version covering broader ideas.We used a systematic review and searched for studies published in high credibility and ranked journals. As part of our data extraction, we categorized the reviewed studies based on their research questions, methodology, sample size, conclusion and focus of the study. We reviewed many articles and those of best relevance were picked out and cross examined with each other to analyze patterns.

LITERATURE REVIEW

SMEs are important to focus on as larger businesses seem to have more resources, such as capital, knowledge, and technology to adjust and drive their business to become more environmentally sustainable. SMEs, on the other hand, face many barriers that are difficult to overcome. SMEs today have increasing demands from stakeholders that want more transparency and accountability from them about what they are doing to expand environmental sustainability. and accountability from them about what they are doing to expand environmental sustainability.

In Parker et al.'s study (2009), 50 peer-reviewed journal articles were analyzed using a deductive thematic approach in being selected for their study. They then analyzed the summaries and grouped the SME types and interventions based on the themes which emerged. Parker et al. (2009) categorized the data based on the method, countries and the size of the SMEs by employees. They found that a majority of the studies were from the Europeans. They also examined the drivers and barriers of sustainability in SMEs from these articles. They concluded it was naive to rely on the extremes of voluntary environmental agreements or regulation and legislation to engage SMEs in environmental improvement. Better yet, intervention strategies to assist SMEs to engage in environmental improvement need to be holistic and designed for the specific category of SME being targeted. They argued that a properly coordinated and mixed strategy intervention approach is likely to be more successful in engaging SMEs.

Albertini (2013) conducted a meta-analysis in which she focused on integrating prior research studying their relationship of environmental management and financial performance of businesses. Albertini (2013) investigated 52 studies over a 35-year period that confirms a positive relationship between environmental performance and financial performance. Moderators' analysis reveals that the relationship is significantly influenced by the environmental and financial performance measures, the regional differences, the activity sector and the duration of the sustainability practices. After discussing the theoretical and managerial implications, Albertini's (2013) meta-analysis tried to answer the question: "When and how does it pay to be green?" Using theories like Accounting-based indicators the results showed that financial performances are positively correlated with environmental management.

Johnson & Schaltegger (2015) conducted another meta-analysis focusing on sustainability management tools for SMEs. Their literature review was guided by specific questions such as: What sustainability management tools, including tools for corporate social responsibility and tools for environmental management, have been designed for and are applicable to SMEs? Their data collection was based on thematic analysis and included an explanation of how many articles were searched and were relevant. They also collected data on what journals these credible articles were from as well as where they were from geographically. Their results consist of explaining four particular areas of investigation, including the overview of proposed tools, the reasons for implementation, the barriers for implementation, and the key facilitating criteria to improve applicability of tools in SMEs.

Lastly, Alvarez et al. (2018) focused on strictly barriers to sustainability in small and medium enterprises. Their method was to analyze the top most cited studies. They were able to find 46 top most cited articles in helping them analyze the top barriers faced by small and medium sized enterprises when starting initiatives for environmental sustainability. They were able to find a total of 175 barriers to entry into environmental sustainability (Álvarez et al., 2018). The barriers that were the most frequent which showed up in 13 articles were lack of resources, high initial capital cost in implementing sustainability, which showed up 11 times and the third most frequent was lack of expertise which appeared in 9 papers. They also found that the countries with the highest amount of research publications that recalled barriers to sustainability for SMEs are United Kingdom, Germany, Italy, Malaysia, and Romania.

Common drivers known to date that promote sustainability in small businesses include financial gains, government regulations, resources, and other benefits. One of the most distinct factors in driving sustainable entrepreneurship were behavioral factors or habits that seem to be the norm within a business. Tur-Porcar et al. (2018) demonstrated that behavioral factors which are motivation, lifestyle, and metacognition scored 49.59% compared to environmental/surroundings at 6.55%, human relations 13.6% and business factors at 30.26%. They focused on the hierarchical design of the key factors that influence entrepreneurial sustainability. Their study sampled nine experts from three distinct professional areas: academia, business, and government. The results reflected that half the reasons why businesses go green is due to behavioral factors (Tur-Porcar et al., 2018). Other drivers that encourage sustainable acts include economic, external pressures and the role of personal values (Aboelmaged, 2018). The results underline that external pressures and entrepreneurs' attitudes are the most important predictors of environmental proactivity both for small and micro firms (Testa et al., 2016). In addition, the need for a fit between personal and professional values was essential in underpinning respondents' engagement with climate change and encouraging a sense of personal responsibility (Williams & Schaefer, 2013). William & Schaefer's (2013) study explored nine in-depth interviews with SME managers in the East of England between March and May 2009.

Other studies found that managers' environmental responsibility is the most important driver of environmental behavior (Shahlan et al., 2018; Iraldo et al., 2017). Top management and stakeholder pressure is a major influence in a business going green. A major conclusion in Pinget et al. 's (2015) study suggests that environmental regulation has a positive effect on the probability of adopting environmental impact (EI) practices. Firms that are conscious of their environmental impact are more actively adopting environmental solutions. Environmental stewardship was able to influence the firm's behavior (Hamann et al., 2016). Hamann et al. (2016) also found that managers' environmental responsibility is the most important driver of environmental behavior. They also showed that a clear and direct influence and the most significant drivers of going green are first local, national and international regulations and stakeholders, such as managers and employees. Their data collection consists of South African wine firms and 55 survey respondents.

Social responsibility and pro –environmental organizational culture and organizational support were found to have significant impacts on company's green initiative as well (Loke et al., 2014). Loke et al. (2014) hand delivered questionnaires to a sample of 1000 randomly selected companies located in the state of Perak. Based on these 1000 questionnaires originally distributed, a total of 571 of them were found completed and usable, yielding a response rate of 57.1%. Majority of the participating companies were from the electrical, machinery and apparatus industries (31.2%) followed by the food products and beverage industries (24.5%). Social, economic and formal practices dimensions of sustainability positively affect competitive advantage, mediated by corporate reputation, customer satisfaction, and organizational commitment (Madsen & Ulhøi,, 2015). Research on 40 Romanian articles showed that strong commitment from managers will increase sustainability strategies and implementation (Zbuchea & Pinzaru, 2017). Ghandi et al. (2018) reveal that top management commitment, technology up-gradation, current legislation, green brand image and future legislation are the five most important drivers for the implementation of integrated lean and green manufacturing in Indian manufacturing SMEs. A competitive advantage acts as a second-stage mediator that positively contributes to financial performance (Cantele & Zardini, 2018). Cantele and Zardini (2018) surveyed Italian Manufacturing businesses; survey of 348 Italian manufacturing small and medium-sized enterprises. Whether it's a corporate social responsibility (CSR), economic advantage, or legal compliance, drivers serve as a motivation to achieve sustainable environmentally friendly practices in SMEs. Companies who have sustainable practices have an advantage over those who don't, not only because they are helping the environment and their own resources but they are also branding an image target customers will want to support. This competitive advantage ultimately helps a business financially.

Triguero et al. (2013) conducted a survey of 5,222 European SMEs and discovered that companies can more easily develop eco-innovation when they work closely with institutions and colleges/government. An example is the European Innovation Program which brought together private stakeholders. The driver of compliance with regulation took the highest rank in Niknamfar et al.'s (2018) study as well. Caldera et al.'s (2019) study found enablers consisting of organizational support, stakeholder collaboration, business direction, lean and green relationship; yet another research article with a similar pattern of drivers repeated in our research. Just like Caldera et al., Raar's (2015) study mailed out surveys and received 360 usable response rate, 319 were directors, managers/owners, general managers, or accountants. Surveys found the same drivers: collaboration, continuous innovation, clear narrative and vision, profitability, foundation for sustainability and lastly external events. In contrast, in Ashton et al.'s study (2017), a majority of firms appear to be internally motivated to implement green practices driven primarily by cost and competitiveness concerns, more than by social responsibility concerns. External coercive pressure from government or customers does not appear to be a significant motivation for these SMEs. However, informal pressure through government incentives and support programs, as well as mimetic pressure through peer learning via industry associations, appear to be more effective in helping these SMEs to further. This is one of the very few articles focusing in the U.S.

Barriers prohibit or create an obstacle for SMEs to implement sustainable practices. For example, in Ghadge et al.'s study (2017), poor market structure, lack of appropriate logistic structure and underdeveloped environmental legislation are all barriers to achieving such practices they had aimed for. Culture plays a role in situations also (Rekik & Bergeron, 2017). In Pinget et al.'s study (2015), they focused on French SMEs. Lack of financial sources revealed to be another main obstacle many businesses seem to have. Research has proven that smaller companies have more problems and barriers because of lack of funds and lack of educated workers (Pinget et al., 2015). It is even more difficult to start a business in developing countries because loans and support from the government are just starting to take hold (Silajdzic et al., 2015). Many of these companies started with support from family because of lack of funding. The barriers don't stop there, when questioning the academics underdeveloped countries were unfamiliar with the statement of green and are just starting to learn about green ideas. So as education goes, they are lacking many resources. Interviews with 24 SMEs, 24 literature review and 3 business analysis in Bosnia and Herzegovina revealed that majority of the SMEs are not planning to introduce green practices even after they were educated about sustainable practices because the weight of running a business keeps them busy enough (Silajdzic et al., 2015).

In Ghadge et al.'s study (2017), demanding warehousing, rough distribution process and unorganized return management all affected the hardship of obtaining greener practices. This study focused on fresh dairy products in Greece that are costly to store and even harder to distribute, both processes require refrigeration and have a great effect on the environment but are a necessity to the dairy market. Using descriptive research, sensitivity analysis, questionnaires, and analytical hierarchy processes, they found that the problem of unorganized return management in the dairy industry also creates waste generated from the return of products. If they overcome this hurdle, they can significantly reduce their waste (Ghadge et al., 2017). One of the main barriers we found during this research is the lack of awareness on sustainability, commonly, Singh et al.'s research (2018) supported those findings. This similarity is again seen in Mangla et al.'s (2017) study in which they examined dimensions of barriers. Organizational barriers ranked first when it came to environmental management. Among the sub categories were lack of appropriate tools and techniques. SMEs foundries lack the resources and technology needed to go green (Wadhwa, 2015; Halme & Korpela, 2014).

Lack of awareness, lack of education, lack of expertise, cost which consisted of high cost of implementing the technology, lack of funds, lack of financial resources, risk, policies and regulations were present in Caldera et al.'s study (2019). Lack of government regulation is another relatively high barrier to SMEs. Because of the lack of monitoring and enforcing strong regulations it does not motivate or encourage SMEs to practice green practices. Luthra et al.'s (2016) findings also showed a lack of rewards and promotions to be a setback to greener SMEs.

Ghazilla et al.'s (2015) study found 39 drivers and 64 barriers in their literature review. The main barriers found out of the 64 were weak organizational structure, lack of empowerment to support, delay of high internal policies, and management resistance; all of which are consistent with researches mentioned before. The most ranked drivers were, voluntary regulations and standards, and compulsory regulation mandated by local government (hazardous and toxic regulation).

METHODOLOGY

We conducted a meta-analysis which involves a carefully selected review of journal articles in relation to SMEs and their journey to adopting these environmentally friendly practices. We strictly used peer-reviewed journals for higher quality from the past 6 years (2013-2019) to investigate the most recent studies. We found our articles by searching online databases (ILLiad, Science Direct, Wiley, ResearchGate, Google Scholar, Etc.) by using keywords in relation to our research (Table 1). We have searched through a total of 845 research articles discovered from our keywords (Table 2). After carefully examining these articles only 122 reached the criteria and needs for our research. The review was carried out using a deductive thematic approach, by examining the findings and categorizing the types of SMEs and the drivers and barriers described by the authors.

Table 1. Search Keywords

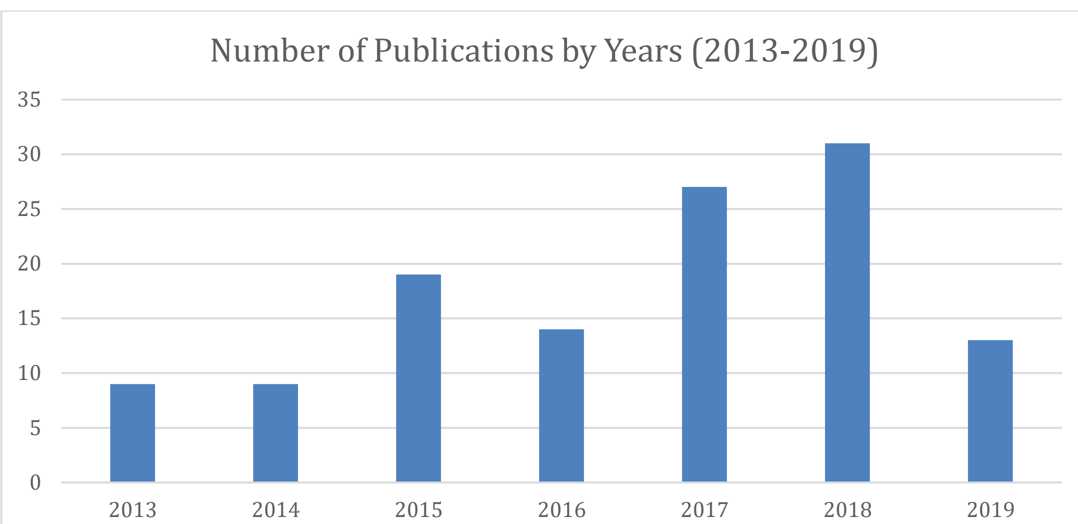
Search String				
"small and medium-sized enterprise (SME)"	AND	Environmental sustainability		
"small and medium-sized enterprise (SME)"	AND	Environmental sustainability	AND	Barriers
"small and medium-sized enterprise (SME)"	AND	Environmental sustainability	AND	Drivers
"small and medium-sized enterprise (SME)"	AND	Environmental sustainability	AND	USA
"small and medium-sized enterprise (SME)"	AND	Going Green		
"small and medium-sized enterprise (SME)"	AND	Going Green	AND	Barriers
"small and medium-sized enterprise (SME)"	AND	Going Green	AND	Drivers
"small and medium-sized enterprise (SME)"	AND	Going Green	AND	USA
"small and medium-sized enterprise (SME)"	AND	Eco-Friendly		
"small and medium-sized enterprise (SME)"	AND	Eco-Friendly	AND	Barriers
"small and medium-sized enterprise (SME)"	AND	Eco-Friendly	AND	Drivers
"small and medium-sized enterprise (SME)"	AND	Eco-Friendly	AND	USA
"small and medium-sized enterprise (SME)"				

Table 2. Keyword Search Results

Keyword Search Results	
Keywords	Count
SME's Barriers	225
SMEs Going Green	100
SMEs Drivers	205
SME's Environmental Sustainability	185
SMEs, Environmental Sustainability, USA	105
SMEs Eco Friendly	25
Total Count	845

Figure 1 compiles the years in which the articles analyzed in this research were published.

Figure 1. Publication by Years



For each article, we analyzed the region in which it took place, the sector, method, sample size, theories and their conclusion. We examined each article closely and analyzed patterns, trends, similarities and differences of all aspects of each research to help draw our findings.

RESULTS AND DISCUSSIONS

The published meta-analysis studies are not as extensive as our study. Our meta-analysis investigates and analyzes 122 articles from 53 different journals and sources (Table 3).

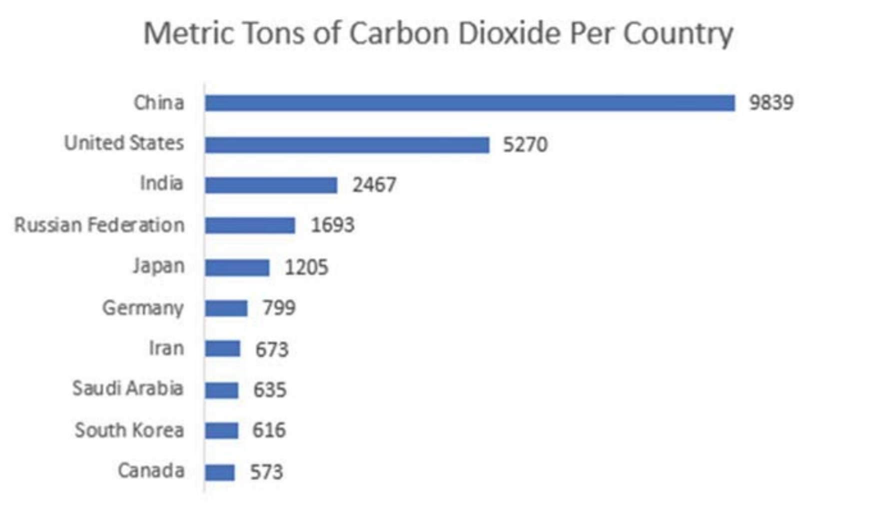
Table 3. List of studies Investigated

Studies	Journal
Silajdžic et al., 2015; Klewitz & Hansen,2014; Cantele &Zardini, 2018; Granley& Welo, 2014 ; Henriques and Catarino, 2015; Gear et al., 2018; Nunes et al., 2019; Long et al., 2018; Groot et al., 2019; Lambrechts et al., 2019; Bar, 2015 ; Cuerva et al., 2014; Gandhi et al.,2018; Testa et al., 2017; Font et al., 2016; Mangla et al.,2017; Mathiyazhagan et al., 2013; Seth et al.,2018; Aboelmaged, 2018; Caldera et al., 2019	Journal of Cleaner Production
Lopez-Perez et al., 2017; Kiefer et al., 2018; Malesios et al., 2018; Prieto-Sandoval et al., 2018; Halme & Korpela, 2014; Knight et al., 2019; Dey et al., 2019; Álvarez et al., 2018; Madsen& Ulhøi, 2016; Lewis et al., 2015; Doluca et al., 2018; Triguero et al., 2016; Jansson et al., 2017; Westman et al., 2018; Prieto-Sandoval et al., 2018	Business Strategy and the Environment
Tur-Porcar et al.,2018; Saez-Martinez et al., 2016; Falle et al. ,2016; Masocha, 2018; Sarango-Lalangui et al., 2018	Sustainability Journal
Wiesner et al., 2018; Tyler et al., 2018; Chassé& Boiral, 2017; Tounés et al., 2018; Albertini, 2013	Organization & Environment
Maniora, 2018; Amaeshi et al.,2016; Wood&Williamson,2014; Tang Zhi and Tang Jintong,2016; Abdelzaher and Abdelzaher, 2017; Leonidou et al., 2015; Olivier, 2019; Sarvaiya et al.,2018; Boiral et al.,2019	Journal of Business Ethics
Jamali et al., 2017; Wickert, 2016; Oliveira et al., 2017; Allet et al., 2017; Spence, 2016; Schaefer et al., 2018; Hamann et al., 2017	Business and Society
Muñoz et al.,2018; Nejati et al.,2017; Altinay et al., 2016; Mayr et al., 2017; Bilge and Leonidou, 2015; Li et al.,2016; Stoian and Gilman, 2016	Journal of Small Business Management
Lucato et al.,2017; Kornilaki & Font, 2019; Zaehringer et al.,2018	Journal of Environmental Management
Yacob and Wong, 2019; Huang et al., 2015; Thanki& Thakkar, 2018; Sajan et al., 2017	Journal of Manufacturing Technology Management
Testa et al., 2016; Girella et al., 2019	Corporate Social Responsibility and Environmental Management
Ghadge et al.,2017; Pinget et al., 2015; Girgenti et al.,2014; Aghelie, 2017; Heidrich and Tiwar, 2013; Triguero et al., 2013; Madsen & Ulhøi, 2015; Robinson & Stubberud, 2014; Singh and Musa, 2016;	All other journals with only one study

PIEDRA-MUÑOZ et al., 2016; Driafish et al.,2019; Galdeano-Go ´mez et al., 2015; Loke et al., 2014; Luthra et al., 2016; Pham, 2018; Wadhwa, 2015; Awasthi & Kannan, 2015; Niknamfar et al., 2018; Dou et al.,2014; Hasan,2016; Starcevic et al., 2017; Ciasullo and Troisi, 2013; Dhillon et al., 2016;Fall, 2013;Panigraphi and Rao, 2018; Zbuchea and Zaru, 2017; Gupta, 2017;Blundel et al., 2013; O'connor et al., 2017;Ashton et al., 2017;Skouloudis et al.,2017; Roehrich et al., 2017; Raar, 2015;Mahmood et al., 2017; Zeeshan et al., 2017; Alziady & Enayah, 2019; Aykol & Leonidou, 2015; Barzegar et al., 2018; Flynn & Davis, 2016; Rahbauer et al.,2018; Shahlan et al., 2018; Namagembe et al., 2016	

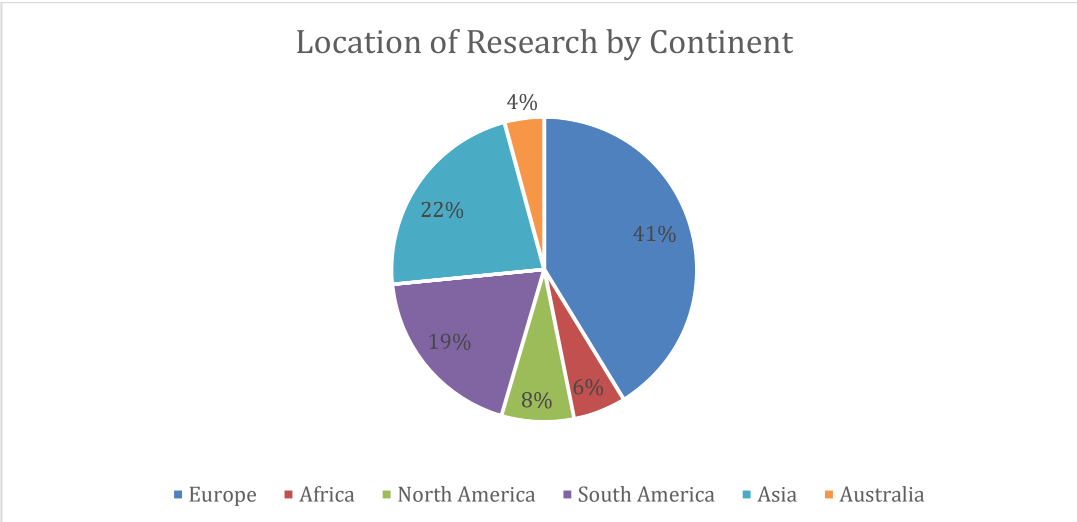
Most studies are not focused on SMEs in the United States. The pattern is disturbing because United States produces 30% of the earth's waste and is a consumer of 25% of the world's natural resources (Davis et al., 2010). The United States creates more pollution in the world and there is little research done on how to encourage SMEs to go green. China has overtaken the US as the world's largest producer of carbon dioxide but that does not downplay the United States involvement because the US receives a bulk of the imports from China (Rapier, 2018). Further research is needed to analyze the United States and regional specific drivers and barriers for environmental sustainability.

Figure 2. Metric Tons of Carbon Dioxide Per Country (Global Carbon Atlas, 2017)



China and the United States create a large amount of the greenhouse gases (Figure 2) and studies investigating how to improve emissions from SMEs fall behind in those countries. The next eight countries only add up to 57 % of the Carbon Dioxide of China and the United States. Figure 3 shows articles we reviewed broken into categories by continent.

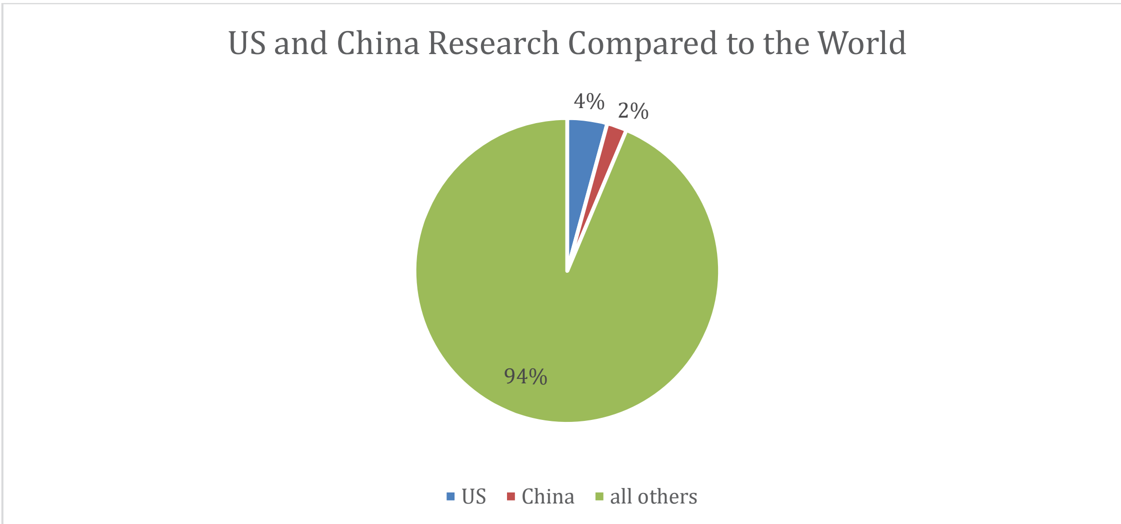
Figure 3. Location of research by continent



Most of the research studies we have found occur in Europe, which makes up 41% of the research we have analyzed, Asia accounts for 22% and South America at 19%. To compare, there was research done by Alvarez et al. (2018) which looked into barriers that were most frequent in SMEs. Alvarez et al. (2018) took 46 of the top cited articles to develop research and broke down what country the research oriented. Compared to our research, they investigated the 46 most cited articles and found 60% of the research oriented from Europe, which is well over half the researches done. Comparing to Parker's et al. (2009) study, it is still close to the same trend with Europe being the most concerned with environmental sustainability. There is more research coming from North America and Australia than what was observed in our latest research. Parker's et al. (2009) research was done on articles from 1999 to 2008 and the research we performed was from 2013 to 2019.

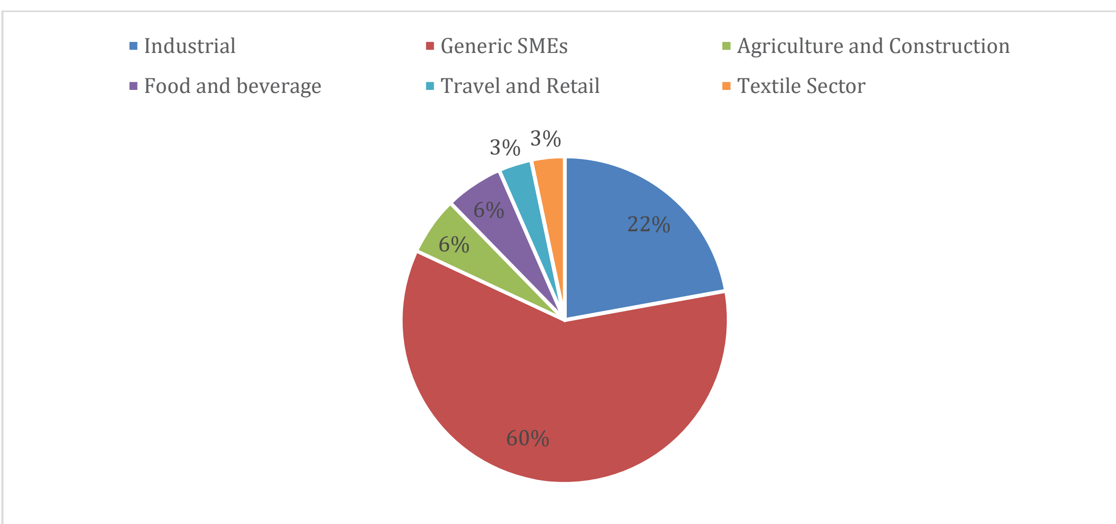
Figure 3 shows us the parts of the world that put an emphasis on environmental sustainability, but since the major contributors to pollution in the world are China and the United States, we broke it down further (Figure 4). China accounts for 2% of the research and the US accounts for 4% of the research. This leaves 94% of the research coming from countries that are not the major contributors to CO_2 gas in the world.

Figure 4. US and China Studies Compared to the World



Research shows that SMEs vary greatly between size, make-up, sector and location, all of which are factors influencing how an SME will develop an environmental sustainability business. Parker et al.'s (2009) recommended that researchers need to concentrate on industry-specific and other variables on which SMEs can be categorized. Based on our research we produced Figure 5 showing what sectors that have been researched from 2013-2019.

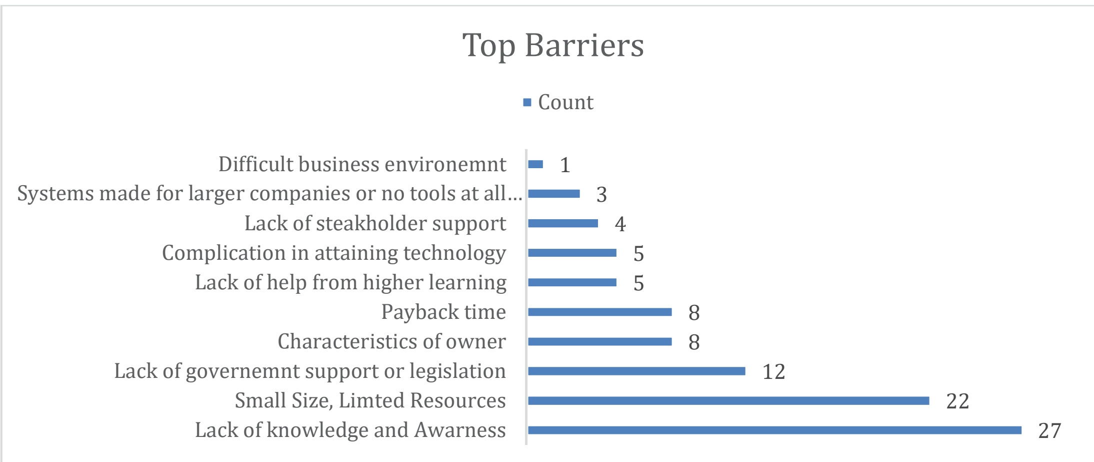
Figure 5. Research of SMEs by Sector



The research shows that 60% of articles focus on generic SMEs and is still not industry, sector or region specific. Further research should be done that is specific because SMEs vary greatly by sector.

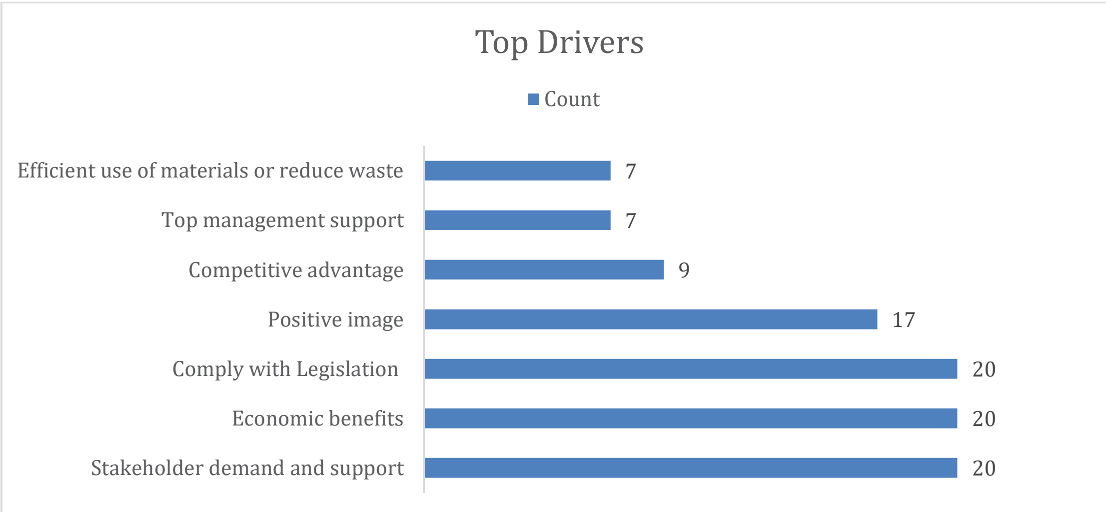
Drivers and barriers are another concern for companies going green. Barriers that seem to come up for SMEs are lack of resources, lack of knowledge, and lack of government support and legislation. These are all barriers faced by SMEs and not necessarily large corporations because they have the size, knowledge and resources. Figure 6 shows a breakdown of the number of studies in our research on barriers to environmentally sustainable SMEs. The major barriers are lack of knowledge and education, and limited resources in SMEs.

Figure 6. Top Barriers



Our analysis shows the drivers are concentrated in economic benefits, stakeholder demand and support and comply with legislation. Most of these were found in research based on SMEs. One unique driver we found was Eco-Islam and the 8 maxims. The authors use the idea of using the strength in the Muslim devotion to their religion to share the fact that the environment is a gift from Allah and should be protected (Abdelzaher, 2015). Figure 7 presents common drivers among the studies we analyzed.

Figure 7. Top Drivers



CONCLUSION

Our research shows that SMEs vary widely from sector, size and location and need to be addressed differently because of it. Papers we reviewed were focused on SMEs in 6 different sectors of Industrial, Generic SME's, Agriculture and Construction, Food and Beverage, Travel and Retail, and the Textile Sector. SMEs can also vary from 1 to 250 employees and are the majority of companies in the world. All SMEs are unique and so are their barriers and drivers. To help SMEs become the most environmentally sustainable they can be, future research needs to be addressed investigating the size, sector and location of the SMEs. One of the many barriers that was seen in almost all of the studies was lack of knowledge. We need future studies on training and support from the government to the public sectors. We need to use the vast knowledge at the colleges and universities around the world to do research and develop more specific tools and knowledge to help support green issues. SMEs don't have the time or the resources to educate themselves on environmental sustainability. Developing countries may not even have access to this information or anybody who was educated in the field. It is up to universities and governments of the world to help bridge this gap. ed in the field. It is up to universities and governments of the world to help bridge this gap. However, the United States is the second leading contributor to Carbon Dioxide output and has the knowledge and resources to lead this worldwide change. The US is leading when it comes to the standard of tertiary education and the top universities in the world are located in the United States. US universities need to focus on SME's in the United States and give them the knowledge and help they need to become green.

REFERENCES

Abdelzaher, D. M., & Abdelzaher, A. (2017). Beyond environmental regulations: Exploring the potential of "eco-Islam" in boosting environmental ethics within SMEs in Arab markets. *Journal of Business Ethics*, *145*(2), 357-371.

Aboelmaged, M. (2018). The drivers of sustainable manufacturing practices in Egyptian SMEs and their impact on competitive capabilities: A PLS-SEM model. *Journal of Cleaner Production,175*, 207-221.

Aghelie, A. (2017). Exploring drivers and barriers to sustainability green business practices within small medium sized enterprises: Primary findings. *International Journal of Business and Economic Development (IJBED), 5*(1), 41-48.

Albertini, E. (2013). Does Environmental Management Improve Financial Performance? A Meta-Analytical Review. Organization & Environment, 26(4), 431-457.

Allet, M., Jamali, D., Lund-Thomsen, P., & Jeppesen, S. (2017). Mitigating Environmental Risks in Microenterprises: A Case Study From El Salvador. *Business & Society, 56*(1), 57-91.

Altinay, L., Madanoglu, M., De Vita, G., Arasli, H., & Ekinci, Y. (2016). The interface between organizational learning capability, entrepreneurial orientation, and SME growth. *Journal of Small Business Management*, *54*(3), 871-891.

Álvarez Jaramillo, J., Zartha Sossa, J., & Orozco Mendoza, G. (2018). Barriers to sustainability for small and medium enterprises in the framework of sustainable development— Literature review. *Business Strategy and the Environment.*

Alziady, A. A., & Enayah, S. H. (2019). Studying the effect of institutional pressures on the intentions to continue green information technology usage. *Asian Journal of Sustainability and Social Responsibility,4*(1).

Amaeshi, K., Adegbite, E., Ogbechie, C., Idemudia, U., Kan, K., Issa, M., & Anakwue, O. (2016). Corporate Social Responsibility in SMEs: A Shift from Philanthropy to Institutional Works? *Journal of Business Ethics, 138*(2), 385-400.

Ashton, W., Russell, S., & Futch, E. (2017). The adoption of green business practices among small US Midwestern manufacturing enterprises. *Journal of Environmental Planning and Management,60*(12), 2133-2149.

Awasthi, A., & Kannan, G. (2016). Green supplier development program selection using NGT and VIKOR under fuzzy environment. *Computers & Industrial Engineering,91*, 100-108.

Aykol, B., & Leonidou, L. C. (2015). Researching the green practices of smaller service firms: A theoretical, methodological, and empirical assessment. *Journal of Small Business Management*, *53*(4), 1264-1288.

Bar, E. (2015). A case study of obstacles and enablers for green innovation within the fish processing equipment industry. *Journal of Cleaner Production, 90*, 234.

Barzegar, M., Rasi, R. E., & Niknamfar, A. H. (2018). Analyzing the Drivers of Green Manufacturing Using an Analytic Network Process Method: A Case Study. *International Journal of Research in Industrial Engineering,7*(1), 61-83.

Blundel, R., Monaghan, A., & Thomas, C. (2013). SMEs and environmental responsibility: A policy perspective. Business Ethics: A European Review, 22(3), 246-262.

Boiral, O., Ebrahimi, M., Kuyken, K., & Talbot, D. (2019). Greening remote SMEs: The case of small regional airports. *Journal of Business Ethics, 154*(3), 813-827.

Caldera, H., Desha, C., & Dawes, L. (2019). Evaluating the enablers and barriers for successful implementation of sustainable business practice in 'lean' SMEs. *Journal of Cleaner Production,218*, 575-590.

Cantele, S., & Zardini, A. (2018). Is sustainability a competitive advantage for small businesses? An empirical analysis of possible mediators in the sustainability–financial performance relationship. *Journal of Cleaner Production,182*, 166-176.

Chassé, S., & Boiral, O. (2017). Legitimizing Corporate (Un) Sustainability: A Case Study of Passive SMEs. *Organization & Environment, 30*(4), 324-345.

Ciasullo, M., & Troisi, O. (2013). Sustainable value creation in SMEs: A case study. *TQM Journal, 25*(1), 44-61.

Cuerva, Triguero-Cano, & Córcoles. (2014). Drivers of green and non-green innovation: Empirical evidence in Low-Tech SMEs. *Journal of Cleaner Production, 68*, 104-113.

Davis, S. J., Caldeira, K., & Matthews, H. D. (2010). Future CO2 emissions and climate change from existing energy infrastructure. *Science*, *329*(5997), 1330-1333.

Dey, P., Malesios, C., De, D., Chowdhury, S., & Abdelaziz, F. (2019). Could lean practices and process innovation enhance supply chain sustainability of small and medium-sized enterprises? *Business Strategy and the Environment, 28*(4), 582-598.

Dhillon, M., Bentley, Y., & Bukoye, T. (2016). Investigating Employee Attitudes Towards Adopting Green Supply Chain Practices in Indian SMEs - Using Qualitative Methods: Literature Review and Research Methodology. *European Conference on Research Methodology for Business and Management Studies,* 367-374.

Doluca, H., Wagner, M., & Block, J. (2018). Sustainability and Environmental Behaviour in Family Firms: A Longitudinal Analysis of Environment-Related Activities, Innovation and Performance. *Business Strategy and the Environment, 27*(1), 152-172.

Dou, Y., Zhu, Q., & Sarkis, J. (2014). Evaluating green supplier development programs with a grey-analytical network process-based methodology. *European Journal of Operational Research,233*(2), 420-431.

Falle, S., Rauter, R., Engert, S., & Baumgartner, R. (2016). Sustainability Management with the Sustainability Balanced Scorecard in SMEs: Findings from an Austrian Case Study. *Sustainability, 8*(6), 545.

Flynn, A., & Davis, P. (2016). The policy–practice divide and SME-friendly public procurement. *Environment and Planning C: Government and Policy,34*(3), 559-578.

Font, Garay, & Jones. (2016). Sustainability motivations and practices in small tourism enterprises in European protected areas. Journal of Cleaner Production, 137(C), 1439-1448.

Galdeano-Gómez, E., Pérez-Mesa, J. C., & Godoy-Durán, Á. (2015). The social dimension as a driver of sustainable development: The case of family farms in southeast Spain. *Sustainability Science,11*(2), 349-362.

Gandhi, Thanki, & Thakkar. (2018). Ranking of drivers for integrated lean-green manufacturing for Indian manufacturing SMEs. Journal of Cleaner Production, 171, 675-689.

Gasbarro, F., Rizzi, F., & Frey, M. (2018). Sustainable institutional entrepreneurship in practice. *International Journal of Entrepreneurial Behaviour & Research, 24*(2), 476-498.

Gear, Sadhukhan, Thorpe, Clift, Seville, & Keast. (2018). A life cycle assessment data analysis toolkit for the design of novel processes – A case study for a thermal cracking process for mixed plastic waste. *Journal of Cleaner Production, 180*(C), 735-747.

Genc, E., Dayan, M., & Genc, O. F. (2019). The impact of SME internationalization on innovation: The mediating role of market and entrepreneurial orientation. *Industrial Marketing Management*.

Ghadge, A., Kaklamanou, M., Choudhary, S., & Bourlakis, M. (2017). Implementing environmental practices within the Greek dairy supply chain: Drivers and barriers for SMEs. Industrial Management & Data Systems, 117(9), 1995-2014.

Ghazilla, R. A., Sakundarini, N., Abdul-Rashid, S. H., Ayub, N. S., Olugu, E. U., & Musa, S. N. (2015). Drivers and Barriers Analysis for Green Manufacturing Practices in Malaysian SMEs: A Preliminary Findings. *Procedia CIRP,26*, 658-663.

Girella, L., Zambon, S., & Rossi, P. (2019). Reporting on sustainable development: A comparison of three Italian small and medium-sized enterprises. Corporate Social Responsibility and Environmental Management, Corporate Social Responsibility and Environmental Management.

Girgenti, Vincenzo, Peano, Cristiana, Baudino, Claudio, & Tecco, Nadia. (2014). From "farm to fork" strawberry system: Current realities and potential innovative scenarios from life cycle assessment of non-renewable energy use and greenhouse gas emissions. The Science of the Total Environment, 473 474, 48

Global Carbon Atlas. (n.d.). Retrieved March 18, 2019, from http://www.globalcarbonatlas.org/en/CO2-emissions

Granly, & Welo. (2014). EMS and sustainability: Experiences with ISO 14001 and Eco-Lighthouse in Norwegian metal processing SMEs. *Journal of Cleaner Production, 64*, 194-204.

Groot, A., Bolt, J., Jat, H., Jat, M., Kumar, M., Agarwal, T., & Blok, V. (2019). Business models of SMEs as a mechanism for scaling climate smart technologies: The case of

Gupta, H. (2017). Integration of Quality and Innovation Practices for Global Sustainability: An Empirical Study of Indian SMEs. *Global Business Review, 18*(1), 210-225.

Halme, M., & Korpela, M. (2014). Responsible Innovation Toward Sustainable Development in Small and Medium-Sized Enterprises: A Resource Perspective. *Business Strategy and the Environment, 23*(8), 547-566.

Hamann, R., Smith, J., Tashman, P., & Marshall, R. S. (2016). Why Do SMEs Go Green? An Analysis of Wine Firms in South Africa. *Business & Society,56*(1), 23-56.

Hasan, M. N. (2016). Measuring and understanding the engagement of Bangladeshi SMEs with sustainable and socially responsible business practices: An ISO 26000 perspective. *Social Responsibility Journal,12*(3), 584-610.

Heidrich, O., & Tiwary, A. (2013). Environmental appraisal of green production systems: Challenges faced by small companies using life cycle assessment. International Journal of Production Research, 51(19), 5884–5896.

Henriques, J., & Catarino, J. (2015). Sustainable Value and Cleaner Production - research and application in 19 Portuguese SME. *Journal of Cleaner Production, 96*, 379.

Hertel, M., & Menrad, K. (2015). Adoption of energy-efficient technologies in German SMEs of the horticultural sector—the moderating role of personal and social factors. *Energy Efficiency,9*(3), 791-806.

Huang, X., Tan, B., & Ding, X. (2015). An exploratory survey of green supply chain management in Chinese manufacturing small and medium-sized enterprises. *Journal of Manufacturing Technology Management, 26*(1), 80-103.

Iraldo, F., Testa, F., Lanzini, P., & Battaglia, M. (2017). Greening competitiveness for hotels and restaurants. *Journal of Small Business and Enterprise Development, 24*(3), 607-628.

Jamali, D., Lund-Thomsen, P., & Jeppesen, S. (2017). SMEs and CSR in Developing Countries. *Business & Society, 56*(1), 11-22.

Jansson, J., Nilsson, J., Modig, F., & Hed Vall, G. (2017). Commitment to Sustainability in Small and Medium-Sized Enterprises: The Influence of Strategic Orientations and Management Values. *Business Strategy and the Environment, 26*(1), 69-83.

Johnson, M. P., & Schaltegger, S. (2015). Two Decades of Sustainability Management Tools for SMEs: How Far Have We Come? Journal of Small Business Management, 54(2), 481-505.

Kiefer, C. P., González, P. D., & Carrillo-Hermosilla, J. (2018). Drivers and barriers of eco-innovation types for sustainable transitions: A quantitative perspective. *Business Strategy and the Environment,28*(1), 155-172.

Klewitz, J., & Hansen, E. G. (2014). Sustainability-oriented innovation of SMEs: A systematic review. Journal of Cleaner Production, 65, 57-75.

Knight, H, Megicks, P., Agarwal, S., & Leenders, M. (2019). Firm resources and the development of environmental sustainability among small and medium-sized enterprises: Evidence from the Australian wine industry. *Business Strategy and The Environment, 28*(1), 25-39.

Kornilaki, M., & Font, X. (2019). Normative influences: How socio-cultural and industrial norms influence the adoption of sustainability practices. A grounded theory of Cretan, small tourism firms. *Journal of Environmental Management,230*, 183-189.

Lambrechts, Gelderman, Semeijn, & Verhoeven. (2019). The role of individual sustainability competences in eco-design building projects. *Journal of Cleaner Production, 208*, 1631-1641.

Langwell, C., Heaton, Dennis P., Fisher, David, McCollum, Bruce, & Rands, Gordan. (2013). *Using Human Resource Activities to Implement Sustainability In Small and Medium-Sized Organizations,* ProQuest Dissertations and Theses.

Leonidou, L. C., Christodoulides, P., Kyrgidou, L. P., & Palihawadana, D. (2015). Internal Drivers and Performance Consequences of Small Firm Green Business Strategy: The Moderating Role of External Forces. *Journal of Business Ethics,140*(3), 607-607.

Lewis, K., Cassells, S., & Roxas, H. (2015). SMEs and the Potential for A Collaborative Path to Environmental Responsibility. *Business Strategy and the Environment, 24*(8), 750-764.

Li, N., Toppinen, A., & Lantta, M. (2016). Managerial Perceptions of SMEs in the Wood Industry Supply Chain on Corporate Responsibility and Competitive Advantage: Evidence from China and Finland. *Journal of Small Business Management, 54*(1), 162-186.

Loke, S., Khalizani, K., Rohati, S., & Sayaka, A. (2014). Drivers and Barriers for Going Green: Perceptions from the Business Practitioners in Malaysia. ASEAN Journal on Science and Technology for Development, 31(2), 49.

Long, T. B., Looijen, A., & Blok, V. (2018). Critical success factors for the transition to business models for sustainability in the food and beverage industry in the Netherlands. *Journal of Cleaner Production,175*, 82-95.

Lopez-Perez, M. E., Melero, I., & Javier Sese, F. (2017). Management for sustainable development and its impact on firm value in the SME context: Does size matter?. *Business Strategy and the Environment*, *26*(7), 985-999.

Lucato, W. C., Costa, E. M., & de Oliveira Neto, G. C. (2017). The environmental performance of SMEs in the Brazilian textile industry and the relationship with their financial performance. *Journal of environmental management*, *203*, 550-556.

Luthra, S., Mangla, S. K., Xu, L., & Diabat, A. (2016). Using AHP to evaluate barriers in adopting sustainable consumption and production initiatives in a supply chain. International Journal of Production Economics*,181*, 342-349.

Lynch-Wood, Gary, & Williamson, David. (2014). Civil Regulation, the Environment and the Compliance Orientations of SMEs. *Journal of Business Ethics, 125*(3), 467-480.

Madsen, H., & Ulhøi, J. (2016). Corporate Environmental Initiatives in Small and Medium Sized Enterprises and Their Outcomes: A Longitudinal Study. *Business Strategy and the Environment, 25*(2), 92-101.

Madsen, H., & Ulhøi, J. P. (2015). Stakeholder pressures, environmental impact and managerial initiatives of SMEs: A longitudinal study. *The Journal of Transdisciplinary Environmental Studies*, *14*(1), 13.

Mahmood, Z., Kouser, R., & Iqbal, Z. (2017). Why Pakistani small and medium enterprises are not reporting on sustainability practices?. *Pakistan Journal of Commerce and Social Sciences (PJCSS)*, *11*(1), 389-405.

Malesios, C., Skouloudis, A., Dey, P. K., Abdelaziz, F. B., Kantartzis, A., & Evangelinos, K. (2018). Impact of small- and medium-sized enterprises sustainability practices and performance on economic growth from a managerial perspective: Modeling considerations and empirical analysis results. *Business Strategy and the Environment,27*(7), 960-972.

Mangla, S. K., Govindan, K., & Luthra, S. (2017). Prioritizing the barriers to achieve sustainable consumption and production trends in supply chains using fuzzy Analytical Hierarchy Process. *Journal of Cleaner Production,151*, 509-525.

Maniora, J. (2018). Mismanagement of Sustainability: What Business Strategy Makes the Difference? Empirical Evidence from the USA. *Journal of Business Ethics,152*(4), 931-947.

Mathiyazhagan, Govindan, Noorulhaq, & Geng. (2013). An ISM approach for the barrier analysis in implementing green supply chain management. *Journal of Cleaner Production, 47*, 283-297.

Mayr, Stefan, Christine Mitter, and Andrea Aichmayr. "Corporate Crisis and Sustainable Reorganization: Evidence from Bankrupt Austrian SMEs." *Journal of Small Business Management* 55.1 (2017): 108-27.

McEwen, T. (2013). An Examination of the Barriers that Impact the Implementation of Environmental Sustainability Practices in Small Businesses. *Journal of Business and Entrepreneurship, 25*(1), 117-139.

Muñoz, P., & Cohen, B. (2018). Entrepreneurial Narratives in Sustainable Venturing: Beyond People, Profit, and Planet. *Journal of Small Business Management, 56*(S1), 154-176.

Namagembe, S., Sridharan, R., & Ryan, S. (2016). Green supply chain management practice adoption in Ugandan SME manufacturing firms. *World Journal of Science, Technology and Sustainable Development,13*(3), 154-173.

Nejati, M., Quazi, A., Amran, A., & Ahmad, N. (2017). Social Responsibility and Performance: Does Strategic Orientation Matter for Small Businesses? *Journal of Small Business Management, 55*(S1), 43-59.

Niknamfar, Amir & Barzegar, Majid & Ehtesham Rasi, Reza. (2018). Analyzing the Drivers of Green Manufacturing using an Analytic Network Process Method: A Case Study. The International Journal of Industrial Engineering: Theory, Applications and Practice. 4.

Nunes, Silva, Moris, & Giannetti. (2019). Cleaner Production in small companies: Proposal of a management methodology. *Journal of Cleaner Production, 218*, 357-366.

O'connor, A., Parcha, J., & Tulibaski, K. (2017). The Institutionalization of Corporate Social Responsibility Communication: An Intra-Industry Comparison of MNCs' and SMEs' CSR Reports. Management Communication Quarterly, 31(4), 503-532.

Pacheco, D. A., Caten, C. S., Jung, C. F., Navas, H. V., & Cruz-Machado, V. A. (2018). Eco-innovation determinants in manufacturing SMEs from emerging markets: Systematic literature review and challenges. *Journal of Engineering and Technology Management,48*, 44-63.

Panigrahi, S. S., & Rao, N. S. (2018). A stakeholders' perspective on barriers to adopt sustainable practices in MSME supply chain: Issues and challenges in the textile sector. *Research Journal of Textile and Apparel, 22*(1), 59-76.

Parker, C., Bellucci, E., Zutshi, A., Torlina, L., & Fraunholz, B. (2015). SME stakeholder relationship descriptions in website CSR communications. Social Responsibility Journal, 11(2), 364-386.

Parker, C. M., Redmond, J., & Simpson, M. (2009). A Review of Interventions to Encourage SMEs to Make Environmental Improvements. Environment and Planning C: Government and Policy, 27(2), 279–301.

Sarango-Lalangui, P., Álvarez-García, J., & del Río-Rama, M. (2018). Sustainable practices in small and medium-sized enterprises in Ecuador. *Sustainability*, *10*(6), 2105.

Pham, L. (2018). Does It Pay for SMEs in Developing Countries to Go Green? Evidence from Vietnam. *Journal of Strategic Innovation and Sustainability*, *13*(3).

Piedra-Munoz, L., Vega-Lopez, L., Galdeano-Gomez, E., & Zepeda-Zepeda, J. (2018). Drivers for Efficient Water Use in Agriculture: An Empirical Analysis of Family Farms in Almeria Spain. *Experimental Agriculture, 54*(1), 31-44.

Pinget, A., Bocquet, R., & Mothe, C. (2015). Barriers to Environmental Innovation in SMEs: Empirical Evidence from French Firms. M@n@gement, 18(2), 132-155.

Prieto-Sandoval, V., Ormazabal, M., Jaca, C., & Viles, E. (2018). Key elements in assessing circular economy implementation in small and medium-sized enterprises. *Business Strategy and the Environment*, *27*(8), 1525-1534.

Puppim de Oliveira, J., Jabbour, C., Jamali, D., Lund-Thomsen, P., & Jeppesen, S. (2017). Environmental Management, Climate Change, CSR, and Governance in Clusters of Small Firms in Developing Countries: Toward an Integrated Analytical Framework. *Business & Society, 56*(1), 130-151.

Raar, J. (2015). SMEs, environmental management and global warming: A fusion of influencing factors? *Journal of Small Business and Enterprise Development, 22*(3), 528-548.

Rahbauer, S., Menapace, L., Menrad, K., & Lang, H. (2018). Determinants for the adoption of green electricity by German SMEs – An empirical examination. *Energy Policy,123*, 533-543.

Rapier R. (2018). China Emits More Carbon Dioxide Than The U.S. and EU Combined. Forbes, July 2018. https://www.forbes.com/sites/rrapier/2018/07/01/china-emits-more-carbon-dioxide-than-the-u-s-and-eu-combined/#77a318f2628c accessed May 2019.

Masocha, R. (2018). Does Environmental Sustainability Impact Innovation, Ecological and Social Measures of Firm Performance of SMEs? Evidence from South Africa. *Sustainability*, *10*(11), 3855.

Rekik, L., & Bergeron, F. (2017). Green Practice Motivators and Performance in SMEs: A Qualitative Comparative Analysis. *Journal of Small Business Strategy, 27*(1), 1-15.

Robinson, S., & Stubberud, H. A. (2014). INNOVATION MOTIVATION: REDUCING ENVIRONMENTAL IMPACT. Allied Academies International Conference.Academy for Studies in International Business.Proceedings, 14(1), 7-12.

Roehrich, J., Hoejmose, S., & Overland, V. (2017). Driving green supply chain management performance through supplier selection and value internalisation: A self-determination theory perspective. *International Journal of Operations & Production Management, 37*(4), 489-509.

Sáez-Martínez, F., Díaz-García, C., & González-Moreno, Á. (2016). Factors Promoting Environmental Responsibility in European SMEs: The Effect on Performance. *Sustainability,8*(9), 898.

Sajan, M., Ramesh, A., & Biju, A. (2017). Lean manufacturing practices in Indian manufacturing SMEs and their effect on sustainability performance. *Journal of Manufacturing Technology Management, 28*(6), 772-793.

Sarvaiya, H., Eweje, G., & Arrowsmith, J. (2018). The roles of HRM in CSR: Strategic partnership or operational support? Journal of Business Ethics, 153(3), 825-837.

Schaefer, A., Williams, S., & Blundel, R. (2018). Individual Values and SME Environmental Engagement. *Business & Society,* 000765031775013.

Seth, D., Rehman, M. A., & Shrivastava, R. L. (2018). Green manufacturing drivers and their relationships for small and medium(SME) and large industries. *Journal of Cleaner Production,198*, 1381-1405.

Shahlan, M. Z., Sidek, A. A., Suffian, S. A., Hazza, M. H., & Daud, M. R. (2018). An examination on the influence of small and medium enterprise (SME) stakeholder on green supply chain management practices. *IOP Conference Series: Materials Science and Engineering,290*, 012059.

Skouloudis, A., Jones, N., Roumeliotis, S., Issac, D., Greig, A., & Evangelinos, K. (2017). Industrial pollution, spatial stigma and economic decline: The case of Asopos river basin through the lens of local small business owners. Journal of Environmental Planning and Management, 60(9), 1575-1600.

Silajdžić, Kurtagić, & Vučijak. (2015). Green entrepreneurship in transition economies: A case study of Bosnia and Herzegovina. Journal of Cleaner Production, 88, 376-384.
Singh, Olugu, & Musa. (2016). Development of Sustainable Manufacturing Performance Evaluation Expert System for Small and Medium Enterprises. *Procedia CIRP, 40*, 608-613.

Singh, M. D., & Thakar, G. D. (2018). Green Manufacturing Practices In SMEs Of India -A Literature Review. *Industrial Engineering Journal,11*(3).

Spence, L. (2016). Small Business Social Responsibility: Expanding Core CSR Theory. Business & Society, 55(1), 23-55.

Starčević, D. P., Zrnić, A., & Mijoč, J. (2017). Is it Worth Going Green in Croatia? Empirical Evidence from SMEs. *EconViews,*141-154.

Stoian, C., & Gilman, M. (2016). Corporate Social Responsibility That “Pays”: A Strategic Approach to CSR for SMEs. *Journal of Small Business Management,55*(1), 5-31.
Tang, Zhi, & Tang, Jintong. (2016). The Impact of Competitors-Firm Power Divergence on Chinese SMES' Environmental and Financial Performance. *Journal of Business Ethics, 136*(1), 147-165.

Testa, F., Gusmerottia, N., Corsini, F., Passetti, E., & Iraldo, F. (2016). Factors Affecting Environmental Management by Small and Micro Firms: The Importance of Entrepreneurs' Attitudes and Environmental Investment. *Corporate Social Responsibility and Environmental Management,23*(6), 373-385.

Testa, F., Nucci, B., Iraldo, F., Appolloni, A., & Daddi, T. (2017). Removing obstacles to the implementation of LCA among SMEs: A collective strategy for exploiting recycled wool. *Journal of Cleaner Production*, *156*, 923-931.
Thanki, S., & Thakkar, J. (2018). Interdependence analysis of lean-green implementation challenges: A case of Indian SMEs. *Journal of Manufacturing Technology Management, 29*(2), 295-328.

Torkkeli, L., Kuivalainen, O., Saarenketo, S., & Puumalainen, K. (2019). Institutional environment and network competence in successful SME internationalisation. *International Marketing Review,36*(1), 31-55.

Tounés, A., Tornikoski, E., & Gribaa, F. (2018). The Formation of Environmentally Friendly Intentions of SME Owner-Managers in an Emerging Country: The Case of Tunisian's Textile–Clothing Industry. *Organization & Environment,*108602661876426.

Triguero, A., Moreno-Mondéjar, L., & Davia, M. A. (2013). Drivers of different types of eco-innovation in European SMEs. *Ecological economics*, *92*, 25-33.

Triguero, A., Moreno-Mondéjar, L., & Davia, M. (2016). Leaders and Laggards in Environmental Innovation: An Empirical Analysis of SMEs in Europe. *Business Strategy and the Environment,25*(1), 28-39.

Tur-Porcar, A., Roig-Tierno, N., & Mestre, A. L. (2018). Factors Affecting Entrepreneurship and Business Sustainability. *Sustainability,10*(2), 452.

Tutterow, V. (2014), "Embedding sustainability into small & medium-sized manufacturing enterprises (SMEs)" Virginia Tech, Center for Leadership in Global Sustainability.

Tyler, B., Lahneman, B., Beukel, K., Cerrato, D., Minciullo, M., Spielmann, N., & Discua Cruz, A. (2018). SME Managers' Perceptions of Competitive Pressure and the Adoption of Environmental Practices in Fragmented Industries: A Multi-Country Study in the Wine Industry. *Organization & Environment,* 108602661880372.

Wadhwa, R. (2015). Sustainable manufacturing in SMEs: Technology options. *International Journal of Computer Science Issues (IJCSI), 12*(6), 107-112.

Wiesner, R., Chadee, D., & Best, P. (2018). Managing Change Toward Environmental Sustainability: A Conceptual Model in Small and Medium Enterprises. *Organization & Environment, 31*(2), 152-177.

Westman, L., Luederitz, C., Kundurpi, A., Mercado, A., Weber, O., & Burch, S. (2018). Conceptualizing businesses as social actors: A framework for understanding sustainability actions in small- and medium-sized enterprises. Business Strategy and the Environment, Business Strategy and the Environment, 12/27/2018.

Wickert, C. (2016). "Political" Corporate Social Responsibility in Small- and Medium-Sized Enterprises: A Conceptual Framework. *Business & Society, 55*(6), 792-824.

Williams, Sarah & Schaefer, Anja. (2013). Small and Medium-Sized Enterprises and Sustainability: Managers' Values and Engagement with Environmental and Climate Change Issues. *Business Strategy and the Environment*. 22.

Yacob, P., & Wong, L. (2019). An empirical investigation of green initiatives and environmental sustainability for manufacturing SMEs. *Journal of Manufacturing Technology Management, 30*(1), 2-25.

Zaehringer, J. G., Wambugu, G., Kiteme, B., & Eckert, S. (2018). How do large-scale agricultural investments affect land use and the environment on the western slopes of Mount Kenya? Empirical evidence based on small-scale farmers' perceptions and remote sensing. *Journal of environmental management*, *213*, 79-89.

Zbuchea, A., & Pinzaru, F. (2017). Tailoring CSR Strategy to Company Size. *Management Dynamics in the Knowledge Economy,5*(3), 415-437.

DECISION SCIENCES INSTITUTE

Rhetoric in young entrepreneurs' firms as an instrument to embrace sustainability

Christian B. Bautista M.
The University of Texas Rio Grande Valley
Email: christian.bautista01@utrgv.edu

ABSTRACT

Corporate Sustainability (CS) has been developed and deployed in established firms by implementing multiple sustainability business models. Lately, this business tendency has been extrapolated among young entrepreneurial firms, this investigation uses the tenets of Neo-Institutional Theory to explain how young entrepreneurial startup firms achieve legitimacy and external resource support. Factors such as sustainable business rhetoric and competitive advantage in entrepreneurial firms can help to achieve organizational legitimacy among startups towards achieving financial performance.

KEYWORDS: Sustainability, Sustainable embracement, Sustainable entrepreneurship, SMEs, Sustainability-oriented innovation, Sustainable supply chain management.

INTRODUCTION

Western industrial societies have experienced severe problems over the last three decades, and probably these problems will be transferred to future generations, as the scale of the environmental and natural resource or sustainability problems they face has become increasingly clear (Elkington, 1994). As a consequence, sustainability and multiple sustainable strategies & business models have been adopted and embraced by multiple firms worldwide in order to maintain profit and growth, including environmental performance or social equity (Bansal, 2002). This economic impact is around US$2.15 trillion, or 3.6% of global GDP, according to the United Nations Environment Program Finance Initiative in 2008, and includes leading public companies globally such as fresh water usage, pollution control, waste and land-use change (Sukhdev, 2012). Additionally, corporate sustainable responsibility extends more than just producing consciously and maybe collect waste created by those produced goods; corporate advertising plays a key role also, since it converts insecurities into a chain of desires and excessive demand of manufactured products or services that create a massive carbon footprint that exceed our planet ability to produce resources and absorb those emissions (Sukhdev, 2012; WWF, 2012).

Nowadays, corporations are considered the main players creating and developing environmental and social issues, as a consequence, they are pointed out as the generators of lack of sustainability in our society (Schaltegger & Wagner, 2011). Likewise, corporations at all levels believe that natural resources or externalities, clean air, fresh water, carbon storage, capacity of forest, among others, are a taken from granted capital and public goods; this is because those commodities are not being traded in any market and are mostly available for free (Sukhdev, 2009). Moreover, natural resources availability for the upcoming future has been a enormous concern for organizational and societal actors such as firm's executives, shareholders, policy-makers, academicians, environmentalists, among others (Bansal, 2002; Nandi & Kaynak, 2017), this fear affects big players such as corporations, however, small players such as small and medium sized enterprises (SME's) are concerned as well since the impact they might have is small, but cumulatively their impact is significant as institution (Hardy, 2011; Lawrence, Collins, Pavlovich,

& Arunachalam, 2006). Concurrently, the career related to reach economic, environmental and social sustainability is attracting more and more industry players, independently of their size, and becoming a focal objective for enterprises along different industries worldwide (Arevalo et al., 2011). Furthermore, corporations, established and young firms, are bodies that the main responsibility is create corporate profit, business growth, shareholder value, market share, economic performance, and innovation (Bansal, 2002; Goldratt & Cox, 2004); putting sustainability efforts not as priority until firms reach enough surplus to invest in (Berrone, Fosfuri, Gelabert, & Gomez-Mejia, 2013) due to higher operational expenditure and as a consequence lower margin profits at year end (Hart, 1995). Moreover, imbedding sustainability in firms is a difficult task since the internal and external system dynamics that surrounds those corporations is not well-defined in order to meet top notch sustainability standards (Arevalo et al., 2011). Additionally, firms' management is facing overwhelming challenges to meet regulations and requirements to meet new sustainability standards that are updating continuously, underestimating and distorting the role of companies in the society (Schaltegger & Wagner, 2011).

Lately, the sustainable development concept has been embraced more vividly by entrepreneurial firms, even developing a new set of streams of thought such as ecopreneurship; social entrepreneurship; sustainable entrepreneurship; and in an indirect way also, institutional entrepreneurship; each with different motivation, goals, and organization (Schaltegger & Wagner, 2011). Similarly, entrepreneurial firms have been recognized as a major contributor for innovative sustainable products and processes, among developing new ventures concerned for social and environmental issues that can resolve environmental problems of global socio-economic systems that current private and public institutions are not effectively attending (Dean & McMullen, 2007; Hall, Daneke, & Lenox, 2010; Tilley & Young, 2009). However, entrepreneurs might ask themselves what is needed to create a venture capable to meet the sustainable criteria nowadays. As a result, research has revealed five key principles of organizational design to have success for sustainability-driven entrepreneurial firms in a competitive market context are: purpose, efficiency, tradeoffs, criteria, and inducement (Parrish, 2010).

Furthermore, to create, and even maintain, sustainability legitimacy, reputation and image, firms need to open their operations for public scrutiny, emphasizing and focalizing transparency and openness, among design significant changes in order to reduce ecological impacts. (Hart, 1995; Sharma & Henriques, 2005). On the other hand, management literature supports that internal and competitive approach at the firm might not be sufficient to implement sustainability practices because issues of firm's external or social legitimacy and reputation are also extremely important (Hart, 1995). The purpose of this article is to focus on understanding the sustainable business rhetoric and competitive advantage in entrepreneurial firms and how they can achieve organizational legitimacy among startups towards achieving financial performance using the tenets of Neo-Institutional Theory.

LITERATURE REVIEW

Fundamentals

Table 1 illustrates the audience with fundamental concepts that will be used during the development of this research. It is important to review them before proceeding further with this document.

Table 1 - Fundamentals for this research

Term	Definition	Source
Sustainability or Sustainable development	This term has been accepted and developed by one of the previous United Nations entities in 1987, The World Commission on Economic Development, and its meaning is "development that meets the needs of the present without compromising the ability of future generations to meet their own needs".	(Development, 1987; Elkington, 1994; Kroll, Abebe, & Dadanlar, 2017; Shaker, 2015)
Rhetoric	The art of persuasion by words and has a long history in the humanities; it can overpass logic as a mode of assessing truth by reflecting message source's implicit intension and consequent actions.	(Aristotle, 1991; Burke, 1969; Kennedy, 1991; Richards, 1936; Searle, 1995; Suddaby & Greenwood, 2005; Zald, 2008)
Legitimacy	A generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions.	(Suchman, 1995)
Organizational Legitimacy	Organizational legitimacy refers to the extent to which the array of established cultural accounts explains organizational existence .	(Castelló & Lozano, 2011; R. W. Scott & Meyer, 1983)
Environmental Violation	Thus, environmental violations are disciplined largely through legal and regulatory penalties, not through reputational penalties.	(Karpoff, Lott, Jr., & Wehrly, 2005)
Business Model	Is a strategic management tool, a combination of material, existing structures and processes at a business organization to improve a company's value chain.	(Tikkanen, Lamberg, Parvinen, & Kallunki, 2005)
Sustainable Business Model	Are traditional business models with incorporated concepts, principles, and goals aiming sustainability, among integrating it in their value proposition, creation and mechanisms.	(Geissdoerfer, Vladimirova, & Evans, 2018)
Corporate Social Responsibility (CSR)	Contains some other constructs such as: corporate citizenship, issues management, and cause-related marketing, and goes beyond the compliance with the legal and basic moral rules of society.	(Montiel, 2008; Palazzo & Richter, 2005)
Entrepreneurship	Involving unique individuals who create value by engaging in the discovery and exploitation of profitable opportunities (Shane and Venkataraman, 2000: 217).	(Gartner, 1990; Ruebottom, 2013; Shane & Venkataraman, 2007)

THEORETICAL DEVELOPMENT/MODEL

Neo-Institutional Theory (NIT): a mean to reach sustainability in young firms

Institutions are embedded and influenced in an open system, exposed to external environment with resources exchanged with other entities, additionally, institutions can be bigger than organizations (W. R. Scott & Davis, 2007). Furthermore, institutions are the rules of the game and organizations are the players (North, 1990); even more important, the specific institutional constraints dictate the margins at which organizations to operate, shaped by broader social-political-cultural processes, and hence make intelligible the interplay between the rules of the game, institutional elements, and the behavior of the actors, institutional agents such as government and professions (North, 1990; W. R. Scott & Davis, 2007). Neo-Institutional Theory (NIT) present a guise with a bold argument: Organizations structures are far from reflecting the forces of task demands and competitive/efficiency requirements, Organizations are more responsive to a set of social-cultural pressures (Dimaggio & Powell, 1983; Meyer & Rowan, 1977; W. R. Scott & Davis, 2007) in this case to embrace sustainability. NIT main supporting pillars are legitimacy and isomorphism (Dimaggio & Powell, 1983; W. R. Scott & Davis, 2007; Suchman, 1995), and both become important at firm level since after implementing a norm or practice, the path that the norm's implementation path follows is: isomorphism, later to conform, this leads to legitimacy, and ultimately leads to resource acquisition (Meyer & Rowan, 1977; Suchman, 1995). Likewise, according to NIT, the resources provision depends on firm's legitimacy levels in an open system such as the institutional environment (Dimaggio & Powell, 1983; Suchman, 1995). As a result, in order to implement a corporate sustainable strategy, a firms need to review and assess the industry structure and norms initially, then the intended strategy to implement within the institution, and finally the expected performance (Dimaggio & Powell, 1983; Hart, 1995; Porter, 1981).

Moreover, NIT focus on how social influence toward conformity shapes organizations' actions, makes local firms to engage in environmental innovation given the regulatory pressures from governments, and normative pressures from non-governmental organizations such as activist (Berrone et al., 2013). Precisely, sustainability practices can develop systematic approaches at different levels that can be defined as sustainable business models (SBM) (Stubbs & Cocklin, 2008) in particular social enterprises, business models that aim at social impact by generating profits from economic activity or reinvesting them entirely (Defourny & Nyssens, 2010).

Young firms' resources acquisition to embrace sustainability challenges

Precisely, strategic and dynamic resources and its attainment is what a firm, particularly young entrepreneur firms, need to sustain a competitive advantage and subsequently have positive financial performance and success in a Schumpeterian economy (Barney, 1991; Dierickx & Cool, 1989; Eisenhardt & Jeffrey, 2000; Moore & Manring, 2009; Priem & Butler, 2001; Teece, Pisano, & Shuen, 1997). According to Bansal (2005), Resource-based rationales adopts corporate sustainable development (CSD) due to its firm influence (Hart & Ahuja, 2002; Waddock & Graves, 1997), substantial investment in capital and human resources (Sharma & Vredenburg, 2002), among changes in technology, legislation, market forces, and in a fast pace or Schumpeterian world with innovation-based competition and creative destruction of existing competences (Eisenhardt & Jeffrey, 2000; Porter & van del Linde, 1995; Teece et al., 1997). Among resources, some other factors are needed to impact venture subsequent growth such as goals, self-efficacy, and communicated vision, these mediated effect of passion, tenacity and new resource skills (J. R. Baum & Locke, 2004). Additionally, three firm motivations are required to take ecological responsive initiatives such as: competitiveness, legitimation, and ecological responsibility, these induced by field cohesion, issue salience, and individual concern (Bansal & Roth, 2000). Embed these efforts and practices, Bansal (2005) proposed three principles of sustainable development need to be included such as: *environmental integrity*, human activities do not erode the earth's land, air, and water resources through environmental management; *social equity*, all members of society have equal access to resources and opportunities by having social responsibility; and *economic prosperity*, a reasonable quality of life through the productive capacity of organizations and individuals in society and create value.

Nonetheless, multiple challenges need to be faced when adopting and embracing sustainability (Epstein & Buhovac, 2010) and the natural-resource-based view theory proposition offers a connection between the environmental challenge and firm resources operationalized through three interconnected strategic capabilities: pollution prevention, product stewardship, and sustainable development (Hart, 1995). This is supported by Sustainable, ethical, entrepreneurial (SEE) enterprises since they are seeking to regenerate the environment and drive positive societal changes rather than only minimizing harm (Markman, Russo, Lumpkin, Jennings, & Mair, 2016); and sustainability-oriented innovations (SOIs), that is, the integration of ecological and social aspects into products, processes, and organizational structures, are increasingly recognized as central contributors to sustainable development particularly in SMEs (Klewitz & Hansen, 2014).

Rhetoric and Sustainability

Rhetoric has been identified as a key tool in building legitimacy (Suchman, 1995). Likewise, a rhetoric perspective and perception can be gained by using legitimacy language, specifically using the conventional positive protagonist and the negative antagonist characters, having an impact on agents working for the firm, and the introduction of new ideas within it (Ruebottom, 2013). Furthermore, if the firm wants to implement as a strategy to implement new ideas to the corporation, a rhetorical strategy, persuasive language to build congruence, needs to be followed (Suddaby & Greenwood, 2005). Additionally, research indicates that rhetorical strategy and language per se has been underused and it can be developed as a core competency by entrepreneurs to pursue stakeholders (Ruebottom, 2013).

Three rhetoric types can be implemented: strategic, uses the scientific-economic paradigm; institutional, based on corporate social responsibility; and dialectic, improves the fuzzy communications channel between firms and stakeholders (Castelló & Lozano, 2011). And within those strategies, two main elements that needs to be present are institutional vocabulary, which means words that have a meaning and reference to the institutional logic, history and professionalism; as well as theorization of change by those characters who are proposing innovation changes and scenarios for firm's improvement, the five types of this kind of changes are: teleological, historical, cosmological, ontological, and value- based (Suddaby & Greenwood, 2005).

Moreover, by using rhetoric, entrepreneurs can use these strategies to create symbolism to connect new ideas to established cultural framework (R. W. Scott & Meyer, 1983), additionally, entrepreneurs can use rhetoric and their sensemaking ability to penetrate existing cultural schemas with cognitive legitimacy (Suddaby & Greenwood, 2005; Van de Ven & Garud, 1993). It would be appropriate to remember that entrepreneurs are looking to disrupt major institutions and create their own systemic change and vision (Ruebottom, 2013; Suddaby & Greenwood, 2005). These entrepreneurs' rhetoric and persuasion skills can be used to implement their vision and schema about sustainability in any realm they have penetration and governance. One example would be social enterprises, since they are the first mode of organization to respect natural system ecologies providing theoretical foundation for sustainable rhetoric in new small firms that has proliferated within the latest decades (Trexler, 2014). Moreover, social entrepreneurs who use rhetoric compel industry members to trend to adopt new practices, including sustainability (Markman et al., 2016; Waldron, Fisher, & Pfarrer, 2016). And precisely with sustainability, its corporate rhetoric needs to be maintained from journey from storytelling a metaphor to argument that sustainability has been achieved, contrary to common believe that this rhetoric is a one-time action (Ihlen, 2014; Ihlen & Roper, 2014; Milne, 2015),

Environmental violations leading to sustainability

In the corporate world a firm perceives that severe legal penalties and castigations are enforced against an environmental laws violator, other companies within the same industry truly learn and respond by themselves in order to be sustainable compliant, creating isomorphism inside their members (Deephouse, 1996; Thornton, Gunningham, & Kagan, 2005). This is because this general deterrence develops a perceived threat of legal punishment and reassurance that compliance is essential and reliable for those firms (Thornton et al., 2005). Furthermore, firms that violate environmental laws are hurt with substantial losses in the market value of firm's equity and profit (Karpoff et al., 2005); media attention can play an important role since it can expose enviromental efforts or issues that can lead to a institution-building process, shape the norms for acceptable and legitimate sustainable development practices (Bansal, 2005). However, pollution violation regulatory punishment by governments is the main pillar that sustains each developed economy country environmental policy, giving policy-makers a key role to develop effective pollution regulations, inspections and sanctions (Gray & Shimshack, 2011).

Likewise, complying with the society, not only with law regulations, becomes critical since negative publicity related to environmental harm potentionally can triggered corporate isolation with the firm's host community, leading to evil firm's visibility, market share loss, and future more closer and scrutinious inspections by regulators (Bansal, 2005; Thornton et al., 2005); creating a "good soldier" behavior within corporate stake and shareholders (Bolino, 1999; Thornton et al., 2005). This mimetic and normative isomorphism of sustainability image and behavior improvement, resulting in a enriched reputation, needs to be well-defined and inspected since it can be developed only as impression enhancement and self-serving (Bolino, 1999; Dimaggio & Powell, 1983); these practices can lead to a concept named: "greenwashing" and it is being used mainly to respond to negative media coverage and outsider stakeholders, such as organizations and government, that demand fundamental corporate conduct sustainable changes and business practices, but multiple times are only representative (Bansal, 2005; David, Bloom, & Hillman, 2007; Greer & Bruno, 1996; Kim & Lyon, 2015; Kroll et al., 2017; Meyer & Rowan, 1977; Thompson, 1967).

Young entrepreneurs' firm's performance by adopting sustainable rhetoric business model

Figure 1 below present and proposed a theoretical model that converges sustainable rhetoric among young entrepreneurial firms and its performance; as well as two factors that moderate positively such performance as low frequency of environmental violations, and high presence of activism; mediated by organizational legitimacy in order to reach firm's performance.

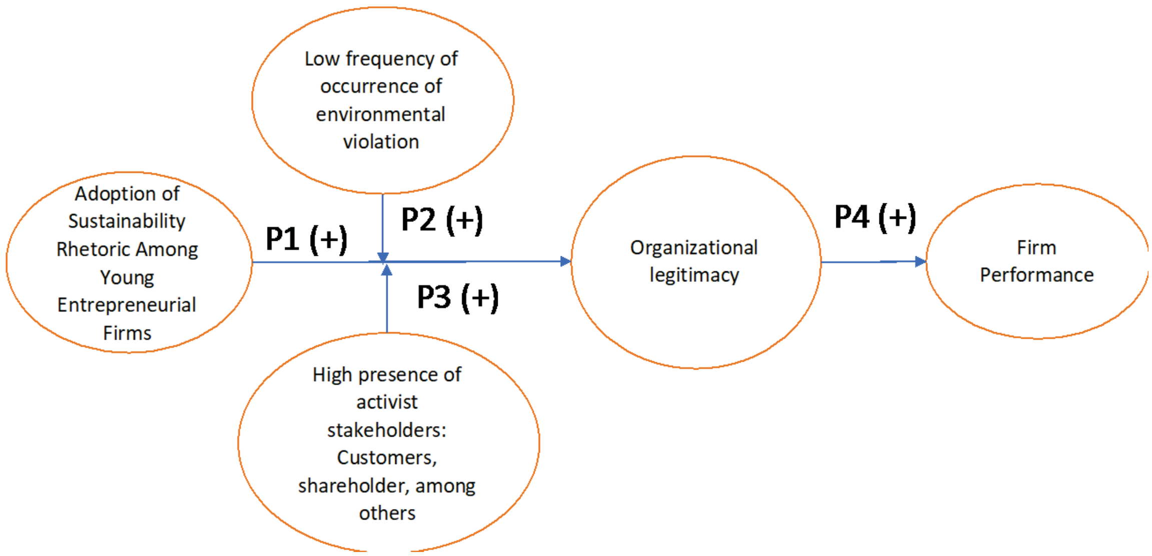


Figure 1 - Sustainable Rhetoric Adoption as Firm Performance Enhancer Model

Sustainability rhetoric adoption and its bond to organizational legitimacy

Rhetoric studies can be perceived as a superficial speech that can be given by a leader in order to persuade subordinates for their action; however, these studies are concerned with capturing the deep structure or the implicit categories of meaning (Berg, 2004; Castelló & Lozano, 2011). Furthermore, the rhetorical devices capable to create a positive narrative in order to produce a protagonist theme and finally organization legitimacy are: Contrastive typecasting, presenting a contrast between positive protagonist and a negative antagonist; Problematizing and theorizing, why the protagonist achieve those outcomes; Alignment, aligning supporters; and Evaluative Statements, must be positive and emphatic for the protagonist (Ruebottom, 2013).

Defining social enterprise is quite difficult since multiple authors, scholars, companies, among other have significant differences in its definition, however, one general conclusion is that it is an enterprise that its self-sustained and its profits benefit its members and their community (Trexler, 2014). This definition becomes important since the leaders of those enterprises, social entrepreneurs, facilitate the adoption of new industry practices including sustainability and ethics since the moment a company is conceived, and that leadership action is being pushed by using a rhetoric (Waldron et al., 2016). Furthermore, the psychological research to date has established that knowledge structures affect information processing in predictable ways and information environment can be interpreted and represented, or used as rhetoric, as knowledge structure that can be used to create consequences or impact at organizational level (Walsh, 1995). This reality interpretation or sensemaking, through leaders vision implemented by a rhetoric tactic, helps to extend and execute business practices (Markman et al., 2016; Waldron et al., 2016). The previous statement has been tested successfully by Thomas, Clark, & Gioia (1993) who investigated the strategic "sensemaking" processes of scanning, interpretation, and action and how those activities are linked to organizational performance; one of the main findings were that high information use strongly influences strategic interpretation, attention to a wide array of information tends to influence the interpretation of strategic issues positively. Even though, rhetoric techniques and persuasion are a great vehicle to move entrepreneurial firms to organizational legitimacy and finally to improve firm performance, those techniques depend on firm logics, norms, and differences between advocates and incumbents (Markman et al., 2016). Based on the above arguments, the following proposition has been conceived:

P1: Adoption of sustainability rhetoric drive firms to organizational legitimacy

Reduction in environmental violations improves legitimacy in a firm

Environmental violations becoming public affects firms reputation, as a consequence reputation repair following past noncompliance, can create organizational motivation to embrace sustainability, resulting in social legitimacy and reputation improvement and credibility at corporate level (Bansal, 2002; Hart, 1995). Furthermore, there is empirical evidence that propose that monitoring and enforcement are significant factors to prevent pollution, reduce emissions, and improve environmental compliance at firms, even those with high compliance levels, among generating substantial specific deterrence and reducing future violations (Gray & Shimshack, 2011).

Firms strategically respond to institutions such as government, professional associations, and particularly media, acknowledging that by complying with those institutions norms, it would provide access to resources (Bansal, 2005; Oliver, 1991; Suchman, 1995). Specifically, media coverage creates firm's visibility, resulting in more public attention and scrutiny, that in case of negative reviews, this coverage can be coercive pressure to implement sustainability and can erode firm's legitimacy (Bansal, 2005; Dimaggio & Powell, 1983). Thus, extrapolating the improving reputation by reducing environmental violation principle to SMEs, becomes an incentive for young entrepreneurial firms to optimize sustainability for the following reasons (Moore & Manring, 2009):

- Becoming valuable sustainable investment targets for larger firms;
- Creating highly competitive networks of sustainable SMEs in market spaces where large enterprises are less successful;
- Becoming highly efficient suppliers in global supply chains through sustainable practices.

One sustainable business model practice than can improve environmental reputation and obtain legitimacy is technological innovation (Boons & Lüdeke-Freund, 2013), this is helpful since is critical renewing competences to achieve good performance in a changing business environment (Teece et al., 1997). This model consist in offering current products in new selling and distribution ways, in order to reduce carbon footprint, for example the new internet based giant retailer: Amazon; the main challenge is convincing current customer about the innovative system that will make their products arrive to them (Boons & Lüdeke-Freund, 2013). Research shows eco-innovation is stronger than innovation from a dimentions economic, social and environmental on SMEs (Klewitz & Hansen, 2014). Based on the above arguments, the following proposition is stated:

P2: Frequency of occurrence of environmental violation moderates positively organizational legitimacy

Activist stakeholders push sustainable rhetoric in order to improve firm's legitimacy

Stakeholders theory is defined as an organizational management and business ethics that addresses morals and values in managing an organization (Hörisch et al., 2014). Furthermore, multiple parties around the firm are considered stakeholders such as employees, customers, suppliers, investors, communities, government, political groups, trade associations and unions, even competitors (Freeman & Reed, 1983); these organizational and societal actors, in addition to corporate governance and sustainability disclosure policies can function as legitimacy instruments for the firm (Michelon & Parbonetti, 2012). Particularly, stakeholder's activism may pull the firm in towards adopting sustainable practices, resulting in multiple times in stakeholders benefits such as return for invertors solutions as well as to preserve ecosystems and promote social equity (Sharma & Henriques, 2005); however, this relationship can create stakeholders dependency with activist stakeholders and even activist shareholders, the most appropriate strategy to handle dependency is: Direct-withholding influence strategy (Sharma & Henriques, 2005). Furthermore, stakeholder management proposed that by having transparent operations, representing stakeholders' interest in decision-making, and distributing the value created by firms equitably among all relevant stakeholders, relationship with stakeholders can be built strongly (Bansal, 2005). For instance, many sustainable entrepreneur firms increasingly engage with external stakeholders, especially, suppliers, buyers, NGO's and activist with the firm purpose to shape young entrepreneurial firm's policies (Markman et al., 2016; Waldron et al., 2016). Additionally, with no stakeholders legitimacy, a firm might have some issues to continue operating and its grow can be limited (Castelló & Lozano, 2011).

Sustainability strategies create many synergistic effects for SME's working collaboratively, as well as systemic benefits for the stakeholders in the value chain such as highly efficient suppliers, industrial ecology, enterprise resilience and global supply chain sustainability (Moore & Manring, 2009); additionally, this business social engagement looks for co-creating and co-engaging startups in create value by innovation and scalation (Crossan & Apaydin, 2010). Constant stakeholders inspection and audit along the value chain, particularly in developing economies countries producing low cost products for developed economies countries, need to be performed in order to be conscious about manufacturing socially undesirable products and not to involve with unethical partners, for example employing child labor (Bansal, 2005). Based on the above arguments, the following statement is being propositioned:

P3: Presence of activist stakeholders' moderates positively organizational legitimacy

Developing legitimacy at the firm enhance performance

The rise of institutional norms, either formal or informal, and social constructed practices is being supported by environments and environmental domains which have institutionalized a greater number of rational myths generate more formal organization (Meyer & Rowan, 1977). As a consequence, institutional rules function as myths which organizations incorporate, gaining legitimacy, resources, stability, and enhanced survival prospects (Meyer & Rowan, 1977) and organizations look similar overtime through Coercive, mimetic, or normative isomorphism (Tolbert et al., 1983). This explains how firms and organizations are being structured, and it is counterintuitive in the sense that decision at corporate level are not made rationally, multiple times decisions are made with the firm objective to imitate and copy what either competitors or some other similar firms are developing and deploying in the market with no logical or rational reason (Dimaggio & Powell, 1983; Meyer & Rowan, 1977; Suchman, 1995; Tolbert et al., 1983). This principle can be though that only applies to established firms, however, young organizations have a higher propensity to die than old organizations if they don't follow similar approach (Stinchcombe, 1965). Research shows support for the external legitimacy than for the internal coordination process and lack of institutional support are prominent underlie the liability of the newness (Singh, Tucker, & House, 1986).

Legitimacy becomes a critical resource because it can bring capital, human, among other important resources to young firm'x, as a consequence those resources can enhance firm's financial performance (Meyer & Rowan, 1977; Suchman, 1995). Moreover, firms of any size always try to be innovative when it comes to seek new forms of legitimacy (Castelló & Lozano, 2011; Palazzo & Scherer, 2006). That hypothetical limit expansion would be given by a bundle of resources that would remain until the enterprise reach its full potential (Lockett, Thompson, & Morgenstern, 2009), developing strong relationships along the supply chain all firm's stakeholders such as suppliers, distributors, wholesalers, retailers, even customers (Muratcan, Kaynak, & Montiel, 2015). Based on the above arguments, the following proposition is stated:

P4: Organizational legitimacy leads to firm performance.

DISCUSSION AND IMPLICATIONS

The combination of rhetoric adaptation framed by NIT for SME's can result on a proper channel of proliferation of sustainability in young entrepreneurial firms. Additionally, the power of narrative style is an potential powerful strategy when implementing changes within or around the corporation, particularly for entrepreneurial endeavors that need to be exposed (Ruebottom, 2013). Furthermore, during the last decades where NIT has accentuated the requirement or need to comply with external pressure, from either stakeholders or peer companies, to adapt industry rules and norms (Berrone et al., 2013; Dimaggio & Powell, 1983; Meyer & Rowan, 1977). Moreover, firms are ambitious of legitimacy with multiple stakeholders and this can be achieved through isomorphism that can be motivated by either intrinsic or extrinsic factors (M. Høgevold et al., 2014).Similarly, the role of rhetoric in successful implementation changes cases in corporations has its contribution, however, it needs to take in count that this rhetoric has its own interaction with some other dynamics around the firm (Suddaby & Greenwood, 2005). Dynamic capabilities are necessary, but not sufficient, conditions for competitive advantage (Eisenhardt and Martin, 2000).

However, rhetorical strategies often have some "missing parts", or enthymemes in a rhetorical syllogism context, and ideas or directions are only partially expressed; and the audience has to complete those enthymemes and its logic (Castelló & Lozano, 2011). This can lead to have a rhetoric ideational bricolage that can be manipulated to support controversial initial changes or modifications in a firm (Ruebottom, 2013), particularly in an start up with limited amount of employees. Within rhetorical tactics implementation and in order to make and give sense at startups, two cognitive structures need to be present such as identity and power, both affect how messages are framed to persuade social entrepreneurial firms to adopt new practices as well as how these firms customize messages related to the new practices (Markman et al., 2016; Waldron et al., 2016). These has been probed by Berrone, Cruz, Gomez-Mejia, & Larraza-Kintana, (2010) when they compared family and nonfamily corporations; these finding indicated that family-controlled firms have better environmental practices adoption predisposition to engage in substantive institutional compliance with environmental regulations depending on external stakeholders, such as government or customers, compared with those with nonfamily nature. For social entrepreneurs who seek to change existing community practices, the difficulties in building legitimacy may pose a challenge that compromises their ability to create sustainable institutional change by using rhetoric.

Additionally, greater environmental regulatory and normative policies can pull and influence corporations to embrace sustainability innovation; however, this needs to be systematically and gradually since these kind of organizational changes can overwhelm the current corporate structure creating a temporary increase in hazard organizational failure (Hannan & Freeman, 1984). For these implementation changes and introduction of new ideas, rhetoric perspective can help in order to keep a positive narrative within the firm (Ruebottom, 2013).

Implications for Research:

Stakeholders are entities that can enable resources and alternatives for learning about massive organizational changes such as implementing and integrating sustainability into firm's business model, this in order to obtain credibility and legitimacy (Arevalo et al., 2011). Furthermore, stakeholders need to have empirical foundations in order to implement any corporate sustainable responsible development model coming from research development stages (Arevalo et al., 2011; Harrison & Freeman, 1999; Palazzo & Richter, 2005). Likewise, stakeholder rhetoric represent a future research that would demonstrate the effectiveness of the protagonist/antagonist rhetorical strategy (Ruebottom, 2013), particularly those with multiple influence in a firm, and mainly in a young entrepreneurial firm. Probably, digging into the stakeholder's area, research as well needs to be done about stakeholders' heterogeneity since there is little known about principals' motivators, preferences, and values influencing the firm for institutional demands (Berrone et al., 2010; Birkin, Polesie, & Lewis, 2009). Firm's stakeholders need to recognize the value of the ecosystem, and understand that is an issue that has social, economic and policy dimensions (Sukhdev, 2011).

On the other hand, the role of culture has not been directed properly when reviewing rhetoric and corporations, finding only a few literature related (Conaway & Wardrope, 2010). This phenomenon is even more rare for that rhetoric including the constructs entrepreneurship and sustainability. An example of this are the high cultural complexity found in Latin-American rhetoric activity contained in CEO's annual reports compared with American corporations (Conaway & Wardrope, 2010). Likewise, finding empirical evidence related to forms and processes of legitimacy-building relations in CSR is a problem for researchers (Castelló & Lozano, 2011; Palazzo & Richter, 2005).

Implications for practice:

Firstly, corporations are using their annual report as a rhetoric strategy with narrative presentation and style, keeping accuracy, clarity and honesty, to convince stakeholders that they are improving in CSR activities (Castelló & Lozano, 2011; Conaway & Wardrope, 2010; Suddaby & Greenwood, 2005). This kind of reports is important in order to gain legitimacy has been confirmed by research in different cultural environments (Conaway & Wardrope, 2010). For that reason, if entrepreneurs are looking for legitimacy in order to obtain resources and firm performance, they need to add rhetoric to their main activities. Similarly, embracing and working with sustainability can provide revenue opportunities, cost savings and risk mitigation (Ihlen, 2014).

Secondly, rhetoric needs to include that natural resources are limited and they should not be taken for granted under any circumstance by any corporation independently of their size (Sukhdev, 2011). In addition, shared economy needs to be embraced in order reduce emissions and operational cost, among boost innovation, especially for production and consumption of goods and services firms (Heinrichs, 2013). These two propositions can be used to promote between startups young firms a rhetoric of low initial capital investment, and at the same time embrace sustainability with environmental management systems while they are growing up, besides commit those entrepreneurial firms to engage in an eco-efficiency model that can provide product stewardship to minimize environmental impacts on the product life cycle (Bansal, 2002). Moreover, this rhetoric to stakeholders needs to include in its vocabulary sets arguments such as: opportunities; risk; creativity and creation; passion, dedication, and importance of the people and multiple human resources; early stage products and spin-offs (Ruebottom, 2013).

Finally, an important, and strong, advice for corporations would be to use rhetoric to embrace sustainability in a short term; however, being accountable in social and environmental issues will help in the medium and long term (Ihlen, 2014).

CONCLUSION

Sustainable development is still under development, and trying to convince multiple audiences in order to being adopted as a daily basis task without leaving human needs fulfilment and most importantly respecting the environment is not an easy endeavor (Lorek & Spangenberg, 2014). This paper explored and analyzed if the adoption of sustainability rhetoric in entrepreneurial firms can help to obtain organizational legitimacy, thus, achieve high firm performance by having low frequency of environmental violations (Gray & Shimshack, 2011; Karpoff et al., 2005) and high present of stakeholders (Freeman & Reed, 1983; Harrison & Freeman, 1999) as moderators. As a result of this analysis, it can be seen that effectively, adopting a sustainability rhetoric, by using the tenets of neo-institutional theory (Castelló & Lozano, 2011; Dimaggio & Powell, 1983; Ihlen, 2014), can lead to obtain legitimacy, particularly critical in entrepreneurial firms (Ruebottom, 2013; Singh et al., 1986; Suddaby & Greenwood, 2005), and finally achieving firm performance. Furthermore, corporate rhetoric needs to formal and informal in order to convince firm's internal and external stakeholders about future plans, specifically, embracing sustainability.

REFERENCES

Arevalo, J. A., Castelló, I., De Colle, S., Lenssen, G., Neumann, K., & Zollo, M. (2011). Introduction to the special issue: Integrating sustainability in business models. *Journal of Management Development*, *30*(10), 941–954. https://doi.org/10.1108/02621711111182466

Aristotle. (1991). *On Rhetoric*. (G. A. Kennedy, Ed.). New York: Oxford Uni- versity Press.

Bansal, P. (2002). The corporate challenges of sustainable development. *Academy of Management Executive*, *16*(2), 122–131. https://doi.org/10.5465/AME.2002.7173572

Bansal, P. (2005). Evolving sustainably: A longitudinal study of corporate sustainable development. *Strategic Management Journal*, *26*(3), 197–218. https://doi.org/10.1002/smj.441

Bansal, P., & Roth, K. (2000). Why Companies Go Green : A Model of Ecological Responsiveness. *The Academy of Management Review*, *43*(4), 717–736. Retrieved from https://www.jstor.org/stable/1556363

Barney, J. B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*. https://doi.org/10.1177/014920639101700108

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-9010.89.4.587

Berg, B. (2004). *Qualitative Research Methods for the Social Science*. Toronto: Allyn and Bacon.

Berrone, P., Cruz, C., Gomez-Mejia, L. R., & Larraza-Kintana, M. (2010). Socioemotional Wealth and Corporate Responses to Institutional Pressures : Firms Pollute Less ? *Administrative Science Quarterly*, *55*(1), 82–113. https://doi.org/10.2307/27856089

Berrone, P., Fosfuri, A., Gelabert, L., & Gomez-Mejia, L. R. (2013). NECESSITY AS THE MOTHER OF 'GREEN' INVENTIONS: INSTITUTIONAL PRESSURES AND ENVIRONMENTAL INNOVATIONS. *Strategic Management Journal*, *34*(8), 891–909. https://doi.org/10.1002/smj

Birkin, F., Polesie, T., & Lewis, L. (2009). A new business model for sustainable development: An exploratory study using the theory of constraints in nordic organizations. *Business Strategy and the Environment*, *18*(5), 277–290. https://doi.org/10.1002/bse.581

Bolino, M. C. (1999). Citizenship and impression management: Good soldiers or good actors? *Academy of Management Review*, *24*(1), 82–98. https://doi.org/10.5465/AMR.1999.1580442

Boons, F., & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: State-of-the-art and steps towards a research agenda. *Journal of Cleaner Production*, *45*, 9–19. https://doi.org/10.1016/j.jclepro.2012.07.007

Burke, K. (1969). *A Grammar of Motives*. Berkeley, CA: University of California Press.

Castelló, I., & Lozano, J. M. (2011). Searching for New Forms of Legitimacy Through Corporate Responsibility Rhetoric. *Journal of Business Ethics*, *100*(1), 11–29. https://doi.org/10.1007/s10551-011-0770-8

Conaway, R. N., & Wardrope, W. J. (2010). DO THEIR WORDS REALLY MATTER? THEMATIC ANALYSIS OF U.S. AND LATIN AMERICAn CEO LETTERS. *Journal of Business Communication*, *47*(2), 141–168.

Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*. https://doi.org/10.1111/j.1467-6486.2009.00880.x

David, P., Bloom, M., & Hillman, A. J. (2007). Investor activism, managerial responsiveness, and corporate social performance. *Strategic Management Journal*. https://doi.org/10.1002/smj.571

Dean, T. J., & McMullen, J. S. (2007). Toward a theory of sustainable entrepreneurship: Reducing environmental degradation through entrepreneurial action. *Journal of Business Venturing*, *22*(1), 50–76. https://doi.org/10.1016/j.jbusvent.2005.09.003

Deephouse, D. L. (1996). DOES ISOMORPHISM LEGITIMATE ? *Academy of Management Journal*, *39*(4), 1024–1039.

Defourny, J., & Nyssens, M. (2010). Conceptions of social enterprise and social entrepreneurship in Europe and the United States: Convergences and divergences. *Journal of Social Entrepreneurship*, *1*(1), 32–53. https://doi.org/10.1080/19420670903442053

Development, W. C. on E. and. (1987). *Our Common Future*. *Medicine, Conflict and Survival*. https://doi.org/10.1080/07488008808408783

Dierickx, I., & Cool, K. (1989). Asset Stock Accumulation and the Sustainability of Competitive Advantage: Reply. *Management Science*, *35*(12), 1514–1514. https://doi.org/10.1287/mnsc.35.12.1514

Dimaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited : Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, *48*(2), 147–160.

Eisenhardt, K. M., & Jeffrey, M. A. (2000). Dynamic capabilities : What are they ? *Strategic Management Journal*, *21*, 1105–1121.

Elkington, J. (1994). Towards the Sustainable Corporation: Win-Win-Win Business Strategies for Sustainable Development. *California Management Review*, *36*(2), 90–100. https://doi.org/10.2307/41165746

Epstein, M. J., & Buhovac, A. R. (2010). Solving the sustainability implementation challenge. *Organizational Dynamics*, *39*(4), 306–315. https://doi.org/10.1016/j.orgdyn.2010.07.003

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston, MA: Pitman.

Freeman, R. E., & Reed, D. L. (1983). Stockholders and Stakeholders: A New Perspective on Corporate Governance. *California Management Review*, *25*(3), 88–106. https://doi.org/10.2307/41165018

Gartner, W. B. (1990). What are we talking about when we talk about entrepreneurship? *Journal of Business Venturing*. https://doi.org/10.1016/0883-9026(90)90023-M

Geissdoerfer, M., Vladimirova, D., & Evans, S. (2018). Sustainable business model innovation: A review. *Journal of Cleaner Production*, *198*, 401–416. https://doi.org/10.1016/j.jclepro.2018.06.240

Goldratt, E. M., & Cox, J. (2004). *The Goal: A Process of Ongoing Improvement*. *North River Press*. https://doi.org/10.2307/3184217

Gray, W. B., & Shimshack, J. P. (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, *5*(1), 3–24. https://doi.org/10.1093/reep/req017

Greer, J., & Bruno, K. (1996). *Green wash: The reality behind Corporate Environmentalism*. *New York: Apex Press.*

Hall, J. K., Daneke, G. A., & Lenox, M. J. (2010). Sustainable development and entrepreneurship: Past contributions and future directions. *Journal of Business Venturing*, *25*(5), 439–448. https://doi.org/10.1016/j.jbusvent.2010.01.002

Hannan, M. T., & Freeman, J. (1984). Structural Inertia and Organizational Change. *American Sociological Review*, *49*(2), 149–164. https://doi.org/10.2307/2095567

Hardy, D. (2011). *The Compound Effect*. Vanguard Press.

Harrison, J. S., & Freeman, R. E. (1999). Stakeholders , Social Responsibility , and Performance : Empirical Evidence and Theoretical Perspectives. *The Academy of Management Journal*, *42*(5), 479–485. https://doi.org/10.4028/www.scientific.net/DDF.309-310.255

Hart, S. L. (1995). A NATURAL-RESOURCE-BASED VIEW OF THE FIRM. *Academy of Management Review*, *20*(4), 986–1014. https://doi.org/10.1128/JCM.00534-09

Hart, S. L., & Ahuja, G. (2002). DOES IT PAY TO BE GREEN? AN EMPIRICAL EXAMINATION OF THE RELATIONSHIP BETWEEN EMISSION REDUCTION AND FIRM PERFORMANCE. *Business Strategy and the Environment*. https://doi.org/10.1002/(sici)1099-0836(199603)5:1<30::aid-bse38>3.3.co;2-h

Heinrichs, H. (2013). Sharing Economy: a Potential New Pathway to Sustaintability. *Gaia*, *22*(4), 228–231.

Hörisch, J., Freeman, R. E., & Schaltegger, S. (2014). Applying Stakeholder Theory in Sustainability Management: Links, Similarities, Dissimilarities, and a Conceptual Framework. *Organization and Environment*, *27*(4), 328–346. https://doi.org/10.1177/1086026614535786

Ihlen, Ø. (2014). "It Is Five Minutes to Midnight and All Is Quiet": Corporate Rhetoric and Sustainability. *Management Communication Quarterly*, *29*(1), 145–152.

Ihlen, Ø., & Roper, J. (2014). Corporate reports on sustainability and sustainable development: "We have arrived." *Sustainable Development*. https://doi.org/10.1002/sd.524

Karpoff, J. M., Lott, Jr., J. R., & Wehrly, E. W. (2005). The Reputational Penalties for Environmental Violations: Empirical Evidence. *The Journal of Law and Economics*, *48*(2), 653–675. https://doi.org/10.1086/430806

Kennedy, G. A. (1991). *Aristotle on Rhetoric. A theory of Civic Discourse*. Oxford: Oxford University Press.

Kim, E.-H., & Lyon, T. P. (2015). *Greenwash vs. Brownwash: Exaggeration and Undue Modesty in Corporate Sustainability Disclosure*. *SSRN*. https://doi.org/10.2139/ssrn.2546497

Klewitz, J., & Hansen, E. G. (2014). Sustainability-oriented innovation of SMEs: A systematic review. *Journal of Cleaner Production*, *65*, 57–75. https://doi.org/10.1016/j.jclepro.2013.07.017

Kroll, M., Abebe, M., & Dadanlar, H. H. (2017). What Are The Odds? Stakeholder and Institutional Challenges of Achieving Sustainable Growth. *Academy of Management Proceedings*, *1*(13971), 1–40. Retrieved from https://doi.org/10.5465/ambpp.2017.13971abstract

Lawrence, S. R., Collins, E., Pavlovich, K., & Arunachalam, M. (2006). Sustainability practices of SMEs: The case of NZ. *Business Strategy and the Environment*, *15*(4), 242–257. https://doi.org/10.1002/bse.533

Lockett, A., Thompson, S., & Morgenstern, U. (2009). The development of the resource-based view of the firm: A critical appraisal. *International Journal of Management Reviews*, *11*(1), 9–28. https://doi.org/10.1111/j.1468-2370.2008.00252.x

Lorek, S., & Spangenberg, J. H. (2014). Sustainable consumption within a sustainable economy e beyond green growth and green economies. *Journal of Cleaner Production*, *63*, 33–44. https://doi.org/10.1016/j.jclepro.2013.08.045

M. Høgevold, N., Svensson, G., Wagner, B., J. Petzer, D., Klopper, H. B., Carlos Sosa Varela, J., … Ferro, C. (2014). Sustainable business models. *Baltic Journal of Management*, *9*(3), 357–380. https://doi.org/10.1108/BJM-09-2013-0147

Markman, G. D., Russo, M., Lumpkin, G. T., Jennings, P. D. D., & Mair, J. (2016). Entrepreneurship as a Platform for Pursuing Multiple Goals: A Special Issue on Sustainability, Ethics, and Entrepreneurship. *Journal of Management Studies*, *53*(5), 673–694. https://doi.org/10.1111/joms.12214

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations : Formal Structure as Myth and Ceremony. *The American Journal of Sociology*, *83*(2), 340–363.

Michelon, G., & Parbonetti, A. (2012). The effect of corporate governance on sustainability disclosure. *Journal of Management and Governance*, *16*(3), 477–509. https://doi.org/10.1007/s10997-010-9160-3

Milne, M. J. (2015). Phantasmagoria, sustain-a-babbling in social and environmental reporting. In *The Routledge Companion to Accounting Communication*. https://doi.org/10.4324/9780203593493.ch9

Montiel, I. (2008). CSR AND CORPORATE SUSTAINABILITY: seperate pasts, common futures. *Organization and Environment*, *21*(3), 245–269. https://doi.org/10.1177/1086026608321329

Moore, S. B., & Manring, S. L. (2009). Strategy development in small and medium sized enterprises for sustainability and increased value creation. *Journal of Cleaner Production*, *17*(2), 276–282. https://doi.org/10.1016/j.jclepro.2008.06.004

Muratcan, E., Kaynak, H., & Montiel, I. (2015). Supplier relations and sustainable operations: the roles of codes of conduct and human resource development. *Academy of Management Annual Meeting Proceedings*, *9*(3), 225–249. https://doi.org/10.5465/AMBPP.2009.44256538

Nandi, S., & Kaynak, H. (2017). A Circular Business Model (CBM) Acceptance Framework. *Development*, *1*(1291716), 1–15.

Näsi, J. (1995). *Understanding stakeholder thinking*. Helsinki, Finland: LSR-Julkaisut Oy.

North, D. C. (1990). *Institutional Change and Economic Performance*. *Cambridge University Press*. https://doi.org/10.2307/2234910

Oliver, C. (1991). STRATEGIC RESPONSES TO INSTITUTIONAL PROCESSES. *Academy of Management Review*. https://doi.org/10.5465/amr.1991.4279002

Palazzo, G., & Richter, U. (2005). CSR business as usual? the case of the tobacco industry. *Journal of Business Ethics*, *61*(4), 387–401. https://doi.org/10.1007/s10551-005-7444-3

Palazzo, G., & Scherer, A. (2006). Corporative Legiti- macy as Deliberation: A Communicative Framework. *Journal of Business Ethics*, *66*, 71–88.

Parrish, B. D. (2010). Sustainability-driven entrepreneurship: Principles of organization design. *Journal of Business Venturing*, *25*(5), 510–523. https://doi.org/10.1016/j.jbusvent.2009.05.005

Porter, M. E. (1981). The Contributions of Industrial Organization To Strategic Management. *Academy of Management Review*, *6*(4), 609–620. https://doi.org/10.5465/AMR.1981.4285706

Porter, M. E., & van del Linde, C. (1995). Green and competitive: ending the stalemate. *Harvard Business Review*, *73*(5), 120–134.

Priem, R. L., & Butler, J. E. (2001). IS THE REOURCE-BASED “VIEW” A USEFUL PERSPECTIVE FOR STRATEGIC MANAGEMENT RESEARCH? *Academy of Management Review*, *26*(1), 22–40.

Richards, I. A. (1936). *The Philosophy of Rhetoric*. (Romanelli, Ed.). New York: Oxford University Press,.

Ruebottom, T. (2013). The microstructures of rhetorical strategy in social entrepreneurship: Building legitimacy through heroes and villains. *Journal of Business Venturing*, *28*, 98–116.

Schaltegger, S., & Wagner, M. (2011). Sustainable Entrepreneurship and Sustainability Innovation: Categories and Interactions. *Business Strategy and the Environment*, *237*(04), 222–237. https://doi.org/10.1002/bse.682

Scott, R. W., & Meyer, J. W. (1983). The Organization of Societal Sectors. In *Organizational Environments: Ritual and Rationality*.

Scott, W. R., & Davis, G. (2007). Organization of the environment. In *Organizations and Organizing: Rational, Natural and Open System Perspectives*. Upper Saddle River, NJ: Pearson Prentice-Hall.

Searle, J. R. (1995). *Construction of Social Reality*. *The Free Press*.

Shaker, R. R. (2015). The spatial distribution of development in Europe and its underlying sustainability correlations. *Applied Geography*, *63*, 304–314. https://doi.org/10.1016/j.apgeog.2015.07.009

Shane, S., & Venkataraman, S. (2007). The promise of entrepreneurship as a field of research. In *Entrepreneurship: Concepts, Theory and Perspective*. https://doi.org/10.1007/978-3-540-48543-8_8

Sharma, S., & Henriques, I. (2005). Stakeholder influences on sustainability practices in the Canadian forest products industry. *Strategic Management Journal*, *26*(2), 159–180. https://doi.org/10.1002/smj.439

Sharma, S., & Vredenburg, H. (2002). Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. *Strategic Management Journal*. https://doi.org/10.1002/(sici)1097-0266(199808)19:8<729::aid-smj967>3.3.co;2-w

Singh, J. V, Tucker, D. J., & House, R. J. (1986). Organizational Legitimacy and the Liability of Newness. *Administrative Science Quarterly*, *31*(2), 171–193.

Stinchcombe, A. L. (1965). Organizations and social structure. *Handbook of Organizations*, *44*(2), 142–193.

Stubbs, W., & Cocklin, C. (2008). Conceptualizing a “sustainability business model.” *Organization and Environment*, *21*(2), 103–127. https://doi.org/10.1177/1086026608318042

Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *The Academy of Management Review*, *20*(3), 571–610.

Suddaby, R., & Greenwood, R. (2005). Rhetorical strategies of legitimation. *Administrative Science Quarterly*, *50*, 35–67.

Sukhdev, P. (2009). Costing the Earth. *Nature*, *462*(November), 2009.

Sukhdev, P. (2011). Putting a Price on Nature: The Economics of Ecosystems and Biodiversity. *Solutions*, *1*(6), 34–43. Retrieved from http://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/8370/Putting a Price on Nature_ The Economics of Ecosystems and Biodiversity.pdf?sequence=1&isAllowed=y

Sukhdev, P. (2012). The corporate climate overhaul. *Nature*, *486*(7401), 27–28. https://doi.org/10.1038/486027a

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533. https://doi.org/Article

Thomas, J. B., Clark, S. M., & Gioia, D. A. (1993). Strategic sensemaking and organizational performance: linkages among scanning, interpretation, action, and outcomes. *Academy of Management Journal*, *36*(2), 239–270. https://doi.org/10.2307/256522

Thompson, J. (1967). Organizations in Action: Social Science Bases of Administrative Theory. In *Organizations in Action: Social Science Bases of Administrative Theory*.

Thornton, D., Gunningham, N. A., & Kagan, R. A. (2005). General Deterrence and Corporate Behaviour. *Law and Policy*, *27*(2), 262–288.

Tikkanen, H., Lamberg, J. A., Parvinen, P., & Kallunki, J. P. (2005). Managerial cognition, action and the business model of the firm. *Management Decision*, *43*(6), 789–809. https://doi.org/10.1108/00251740510603565

Tilley, F., & Young, W. (2009). Sustainable Entrepreneurs: Are They the True Wealth Generators of the Future? *Greener Management International*, *55*(1), 79–92.

Tolbert, P. S., Zucker, L. G., Quarterly, A. S., Mar, N., Tolbert, P. S., & Zucker, L. G. (1983). Institutional Sources of Change in the Formal Structure of Organizations : The Diffusion of Civil Service Reform , 1880-1935. *Administrative Science Quarterly*, *28*(1), 22–39.

Trexler, J. (2014). Social Entrepreneurship as Algorithm: Is Social Enterprise Sustainable? *Emergence Compelxity & Organization*, *10*(3), 65–85. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1307752

Van de Ven, A. H., & Garud, R. (1993). The co-evolution of technical and institutional events in the development of an innovation. In J. A. C. Baum & J. Singh (Eds.), *Evolutionary Dynamics of Organizations* (pp. 425–443). New York: Oxford University Press.

Waddock, S. A., & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal*. https://doi.org/10.1002/(SICI)1097-0266(199704)18:4<303::AID-SMJ869>3.0.CO;2-G

Waldron, T. L., Fisher, G., & Pfarrer, M. (2016). How Social Entrepreneurs Facilitate the Adoption of New Industry Practices. *Journal of Management Studies*. https://doi.org/10.1111/joms.12187

Walsh, J. P. (1995). Managerial and organizational cognition: notes from a trip down memory lane. *Organization Science*, *6*(3), 2701–2707. https://doi.org/10.1055/s-0030-1259018

WWF. (2012). *WWF Living Planet Report 2012*.

Zald, M. N. (2008). Organization Studies as a Scientific and Humanistic Enterprise: Toward a Reconceptualization of the Foundations of the Field. *Organization Science*. https://doi.org/10.1287/orsc.4.4.513

DECISION SCIENCES INSTITUTE

Seven Step Model: Traditional to Green Supply Chain Management Transformation

Jeff Adams
Roosevelt University
Email: jadams32@roosevelt.edu

Ganesh Vaidyanathan
Roosevelt University
Email: gvaidyanathan@roosevelt.edu

Asghar Sabbaghi
Roosevelt University
Email: asabbaghi@roosevelt.edu

ABSTRACT

Organizations are often unclear how to transform themselves from a traditional supply chain management to a green supply chain model. We interviewed several suppliers to design a transformation process. We illustrate a seven-step model ground in Stakeholder Theory on how those suppliers transformed themselves to be a green supplier.

KEYWORDS: Green supply chain, management, transformation, model

INTRODUCTION

Technological advances and intensified global competition in the market for products and services have contributed to numerous studies in examining the effectiveness and efficiency in supply chain management. Green supply chain management (GSCM), a concept that attempts to address these environmental concerns and efficiency improvements in supply chain processes, have been the subject of many recent studies. Due to environmental concerns, regulators, communities and employees, are demanding efficient actions from firms (Paulraj et al., 2017). Moreover, there is a consistent growth in the evaluation of green supply chain management practices and performance (Tseng, Islam, Karia, Fauzi, & Afrin. (2019).

Because of the growing awareness regarding environmental sustainability issues by consumers it is becoming important for manufacturers to adopt green supply chain management practices to remain competitive. The environmental awareness of customers has pushed manufacturers to adopt cleaner production techniques and implement GSCM practices. Mathiyazhagan, Govindan, Hag & Geng 2013). GSCM activities not only have to be adopted within the organization but in addition need to be coordinated with suppliers and customers (Green, Zelbst, Meacham & Bhadaurin, 21012).GSCM can result in improved organizational performance, Past studies have shown that the adoption of GSCM practices leads to better performance particularly economic performance. These improvements are across all size organizations and industries. The improved economic performance is because of better sales, profits, and market share. Other potential benefits of GSCM are better operational efficiency with reduced energy consumption lower scrap rates and inventory levels. (Geng, Mansoiuri & Aktas, 2017).

In 1989, Walmart launched one of the first major retail campaigns to sell environmentally safe products in recyclable or biodegradable packaging. The corporation promoted these eco-friendly products by labeling them with green-colored shelf tags. The idea was to improve the company's impact on the environment through a commitment to three ambitious goals including 100 percent renewable energy, zero waste, and sell products that sustain the resources and the environment (Scott, 2005). Walmart also cultivated deeper relationships with its suppliers in addressing the environmental concerns. In the past, retail textile buyers selected manufacturers based on the cost and quality of their products. Walmart, however, interacted with suppliers, more often, more directly, and for a greater duration than ever before. These closer relationships were necessary to sustain initiatives like the organic cotton project (Plambeck & Denend, 2008).

Walmart has forged ties further upstream to become more efficient and to reduce costs. The company used to buy cotton from Turkey, ship it to China for spinning and knitting, and then ship it again to Guatemala to be cut and sewn. Recently, they are finding opportunities to eliminate the shipment to China and have all processing done in Guatemala. Going directly to Guatemala not only saves time and money for Wal-Mart, but also further reduces the company's impact on the environment by lessening the amount of fuel and other resources used in shipping (Plambeck & Denend, 2008). According to Plambeck and Denend (2008), at the heart of Walmart's business sustainability strategy is a shift from generating value through price-based, transactional interactions toward generating value from longer-term, collaborative relationships with nonprofits, suppliers, and other external stakeholders. Through its sustainable value networks, Walmart gains a whole-system perspective that helps the retailer find profitable ways to address environmental issues such as fishery depletion, climate change, and pollution. Walmart set the goal of a 5 percent reduction in packaging by 2013. It has been widely reported that the retail giant expects the cut in packaging will save 667,000 metric tons of carbon dioxide from entering the atmosphere.

With increasing consumer awareness and regulatory compliances, organizations with GSCM practices will have a competitive advantage over others who are either lagging or reluctant to embrace GSCM. Many organizations are reluctant because they are not convinced how to deploy GSCM in ways that can benefit them from their desired cost savings and creation of value. An earlier study (Beamon, 1999) recognized five issues of environmental management in GSCM. The first issue is how to find out the environmental factors driving the green supply chain. The second is to understand the differences between the traditional supply chain and the green supply chain. The third is to understand the risks arise from practicing the green supply chain. The fourth is how to measure the performance for the green supply chain and the final problem is to develop a process that would create an effective green supply chain.

This paper illustrates how companies must collaborate and partner with their suppliers in order to produce a more effective green supply chain process. A seven-step model, grounded in the theoretical perspectives of the Stakeholder Theory, is developed to illustrate how suppliers can transform themselves from a traditional supplier to be a green supplier.

LITERATURE REVIEW

Since Carter and Rogers (2008) introduced the concept of sustainability —the integration of environmental, social, and economic criteria that allow an organization to achieve long-term economic viability —the environmental concerns in supply chain have received more attention in the literature and in business in recent years. Green supply chain has recently received much attention among researchers and practitioners of operations and supply chain management and has been the subject of numerous extensive studies in recent years due to environmental concerns and awareness. In a traditional supply chain process, the resources from suppliers are used by the manufactures to produce and sell products to retailers and consumers (Beamon, 1999). But the green supply chain would integrate the environmental elements into the supply chain process, and the goal is to address the influences and relationships between supply chain management and the natural environment. This would enlarge the consideration of whole manufacturing process to relate and include environment influences by products. For example, in April 2010, IBM announced that it will require its 28,000 suppliers in more than 90 countries to install management systems to gather data on their energy use, greenhouse gas emissions and waste and recycling. Those companies in turn must ask their subcontractors to do the same if their products or services end up as a significant part of IBM's $40 billion global supply chain. The suppliers must also set environmental goals and make public their progress in meeting those objectives (IBM. 2009).

Traditionally, environmental management in business has been limited to complying with the environmental laws and regulations, and their financial consequences. However, environmental awareness has been associated with competitive advantage in business, particularly when the companies taking a constructive approach to environmental issues, look beyond their current process to find and eliminate sources of waste. in this approach, the focus is on the value embodied in the product and process and maximization of the benefit attained from environmental initiatives (Walton, Handfield & Melnyk 1998). Still, there is another approach that integrate total quality environmental management (TQEM) into its planning and operations processes that includes not just executives and workers, but customers, suppliers, and neighbors (Makower, 1994).

Beamon (1999) explores Lamming and Hampson's definition of green supply chain as the environmental impact of the production process from the original resource to the end of the production including the disposal process. Zhu and Cote (2004) define the green supply chain as finding a 'win-win' strategy for both suppliers and distributors on marketing performance and environmental issues. In order to meet these requirements, the company needs to connect the suppliers to import environmental products.

Zhu and Sarkis (2004) notes various definitions of the green supply chain. Green, Morton, and New (1996) define it in the context of the environment in purchasing and innovations. Narasimhan and Carter (1998) define green supply chain as the activities that involve recycling, reduction, reuse, and substitution of materials. Godfrey (1998) defines the green supply chain as monitoring and improving environmental performance of the supply chain management. Messelbeck and Whaley (1999) assert that the green supply chain would be researching, developing, manufacturing, storing, transporting, and using a product, as well as disposing of the product's waste.

GREEN SUPPLY CHAIN BENEFITS

In the green world, every production process should meet the requirement of environmental protection policy. The public is aware that they need to protect the environment. In order to respond to the demands of the consumers and environmental groups, the manufacturers need to use any necessary means to improve the production flow process to reduce any pollution by any possible measure (Fiksel, 1996). On the other hand, the green supply chain needs to help the manufacturers to benefit from these changing processes. The companies could decrease the cost of the materials, maintenance, operation, and repairing by checking and investing the inventory operation. In addition, the companies could benefit from recycling materials and disposes products and decrease the risks to the workers' heaths by using environmental protection technologies during the production processes.

The benefits of green supply chain are much more than environmental protection. The green supply chain does not only provide the benefits by reducing hazardous wastes or purifying toxic water but also has the potential of sustainably social improvement, utilization of resources usage, and ethical perception development. Green supply chain provides both direct and indirect benefits. Direct benefits include reduction of air emission, waste water and solid wastes, decreasing of consumption for hazardous/harmful/toxic materials, decreasing of frequency for environmental accidents, improving an enterprise's environmental situation, decreasing of cost for materials purchasing, cost for energy consumption, fee for waste treatment, and fee for waste discharge, decreasing of fine for environmental accidents, reduction of cost of supply chain and cost of production, reducing cost of ownership of customer, and lowering resource consumption. On the other hand, GSCM gives indirect benefits on implanting the conscious of environmental protection, building up stronger connection among stakeholders, satisfying customer needs, and developing the relationship of society. Furthermore, a company should consider adopting of the green supply chain since the GSCM will give great benefits on customer relationship, resources management, cost reduction, product differentiation, competitive advantage, supply chain management, regulation and risk management, branding, return on investment, employee regulations, and ethics of business (Khiewnavawongsa and Schmidt, 2008).

THE GREEN SUPPLY CHAIN IMPROVEMENT PROCESS

The evolution to a greener supply chain is a continuous process with no final end point. Because it is an endless nature of becoming greener an appropriate method for developing a green supply chain is the continuous improvement methodology. Kazan or continuous improvement focusses on the reduction of waste and improving efficiency. The purpose of continuous improvement is to eliminate waste (muda) in all areas of an organization and the supply chain. Waste is considered anything that customers are not willing to pay for. (Bhuiyan and Bahel, 2005). Waste can take the form of pollution. With a continuous improvement process in mind, a seven-step model for greening the supply chain was developed.

The stakeholder Theory (ST) asserts that managers have a fiduciary duty not merely to the corporation's stockholders, but to the corporation's stakeholders— to those groups that are either vital to the survival and success of the corporation or whose interests are vitally affected by the firm (Smith & Hasnas, 1999). One of the four theories offered by Donaldson and Preston (1995) states that ST is managerial in that it recommends attitudes, structures, and practices and requires that simultaneous attention be given to the interests of all legitimate stakeholders. The seven-step model integrates the view of the stakeholders of a firm to develop a green supply chain strategy.
This seven-step approach was developed after interviews with business managers involved in the green supply chain process in companies located in Midwest US. The seven steps are listed below for transforming an organization's supply chain process to a greener or more sustainable supply chain:

- Calculate total cost of ownership and identify the green aspect of cost of ownership.
- Identify the touchpoints between product lifecycle and greening process;
- Evaluate all bases of touchpoints;

- Obtain “buy-in” from stakeholders;
- Develop green strategy for product lifecycle;
- Integrate (change management, process reengineering, and technology) and implement green strategy;
- Assess metrics and setup a continuous improvement process

Calculate Total Cost of Ownership and Identify the Green Aspect of Cost of Ownership
The first step in developing a greener organization is to document all the costs and causes of waste and pollution within the organization’s operations and supply chain. In conventional supply chains, the most visible costs are considered while ignoring less apparent costs, pollution, and waste. An example of this is when selecting a supplier only considering the purchase cost while not considering other less visible factors such as the environmental performance of the supplier. In a cost analysis, organizations would normally review all their purchases as a first step towards developing a strategic sourcing plan (Monczka, Handfield, Giunipero, & Patterson, 2014). In considering environmental aspects of purchases, all costs will be considered including not only purchase costs, but also the generation of waste and pollution. All steps and activities in the supply chain need to be considered and costs, waste, and sources of pollution identified. Once the costs and causes of pollution and waste are identified a plan can be developed to reduce the sources of waste and develop a greener supply chain in the following six steps.

Identify the Touchpoints Between Product Lifecycle and Greening Process
The second step in this process is to identify touchpoints or responsible parties in the supply chain. The product life cycle includes the introduction, growth, maturity, saturation, and decline stages. Touchpoints would include suppliers and logistics providers that could be sources of waste and pollution in those stages. This step identifies specific responsible parties where each source of cost and waste take place.

Evaluate All Bases of Touchpoints
Once the major touchpoints are identified each specific point is evaluated to see how it is responsible for costs, waste, and pollution. Possible changes that can be made to make the overall process greener are identified at each touch-point location. Touchpoint evaluations for example could identify alternative suppliers that are closer and require less energy costs for transportation.

Obtain “Buy-in” from Stakeholders
The next step is to identify and obtain buy-in from stakeholders that are impacted by the greening process. Stakeholders would include customers, suppliers, employees, stockholders, government agencies, and local communities. In this step the organization works with each of the stakeholders to make sure the stakeholder’s needs are met in the process of developing a greener supply chain.

Develop Green Strategy for Product Lifecycle
After working with each of the touch points and stakeholders the next step is to develop a holistic strategy for implementing a more greener supply chain and product lifecycle. Individual alternative changes to each of the touchpoints are evaluated for their impact on the performance of the whole supply chain. The basic principles of green strategy include repair, recycle, recover, repair, regeneration, remanufacture, reduction, and waste management (Glavic & Lukman, 2007).

Integrate and Implement Green Strategy
A detailed plan to transform the supply chain is developed based on information from all the previous steps. Change management, process reengineering, and technology strategies should be included in the integration. The transformation plan that includes those strategies should be implemented. Key metrics are developed and used to determine the effectiveness of the change. These metrics will serve as a yardstick to determine progress made in the supply chain transformation process.

Assess Metrics and Setup a Continuous Improvement Process
After changes have been implemented key measures are evaluated to determine the effectiveness of the transformation. Since this is a continuous process the organization returns to step 1 to develop further process improvements.

CONCLUSION AND FUTURE RESEARCH

We are planning to use qualitative research methods to analyze several consumer product manufacturing companies in Midwest. It focuses on five fields including materials used in product design for the environment, product design processes, supplier process improvement, supplier evaluation, and inbound logistics processes to find out the relationship between those processes.

A study by Cote, Lopez, Marche, Perron, and Wright (2008) examined three medium-sized companies in Burnside Industrial Park, Nova Scotia, and evaluated the challenges and opportunities for reducing greenhouse gas emissions and solid waste which are of significant concern to the Nova Scotia federal government. The study reviewed the procurement policies and practices of large corporations which serve as both suppliers and customers and the results show that time and financial resources are the greatest factors that limit many companies to employ environmental-based processes.

After analyzing the supply chain process for these companies, the paper shows that the design department and purchasing department should work together to make an environmental product from the resource directly because a lot of materials need to be concerned with the recycling issue and avoid any impact on the environment after producing the merchandise. We proposed that the company should use its purchasing power to influence the suppliers to let them evaluate their process into the green process. A qualitative study using the interview data is being planned.

REFERENCES

Beamon, B.M. (1999). Designing the green supply chain. *Logistics Information Management, 12*, 332–342.
Bhuiyan, N., & Baghel, A. (2005). An overview of continuous improvement: from the past to the present. *Management Decision, 43*(5), 761-771.
Carter, C., & Rogers, D. (2008). A Framework of Sustainable Supply Chain Management: Moving Toward New Theory. *International Journal of Physical Distribution & Logistics Management, 38*(5), 360-387.

Cote, R. P., Lopez, J., Marche, S., Perron, G. M., & Wright, R. (2008). Influences, practices and opportunities for environmental supply chain management in Nova Scotia SMEs. *Journal of Cleaner Production, 16*(15), 1561-1570.
Donaldson, T., & Preston, L. (1995). The stakeholder theory of the modern corporation: Concepts, evidence and implications. *Academy of Management Review, 20*, 65-91.
Fiksel, J. (1996). *Design for environment: Creating eco-efficient products and processes.* New York: McGraw-Hill.
Geng, R., Mangeiuri, A., & Aktas, E. (2017) The relationship between green supply chain management and performance: A meta-analysis of empirical evidence in Asian emerging economies. *International Journal of Production Economics, 183*, 245-258.
Glavič, P., &, Lukman, R. (2007), Review of sustainability terms and their definitions, *Journal of Cleaner Production, 15*(18), 1875-1885.
Green, K., Morton, B., & New, S. 1996. Purchasing and environmental management: interaction, policies and opportunities. *Business Strategy & the Environment, 5*, 188-97.
Green, K., Zelbst, P., Meacham, J., & Bhadauria, V. (2102) Green supply chain management practices: impact on performance. *Supply Chain Management: An International Journal, 17*(3). 290-305.
Godfrey, R. (1998). Ethical purchasing: Declining the supply chain beyond the environment in Russel, T. (ed) *Greener Purchasing: Opportunities and Innovations,* 244-251.
IBM. 2009. Cosco leverages supply chain insights to keep its distribution center network in balance. October 14. Retrieved from: ftp://ftp.software.ibm.com/software/ solutions/pdfs/ on March 18, 2019.
Khiewnavawongsa, S., & Schmidt, E. K. (2008). Green power to the supply chain. *Proceedings of the Annual Meeting of the Association of Collegiate Marketing Educators.* Houston, TX. March 4-8, 2008.
Paulraj, A., Chen, I. J., & Blome, C. (2017). Motives and performance outcomes of sustainable supply chain management practices: A multi-theoretical perspective. *Journal of Business Ethics, 145*(2), 239-258.
Plambeck, E.L., & Denend, L. (2008). The greening of Wal-Mart. Stanford Social Innovation Review, Retrieved from https://ssir.org/articles/entry/the_greening_of_wal_mart on May 10, 2019.
Makower, J. (1994). *The E-Factor: The bottom-line approach to environmentally responsible business.* New York: Times Books.
Messelbeck, J., & Whaley, M. (1999). Greening the health care supply chain: triggers of change, models for success. *Corporate Environmental Strategy, 6*(1): 39–45.
Mathiyazhagan, K., Govindan, K., Hag, A., & Geng, Y., (2013). An ISM approach for the barrier analysis in implementing green supply chain management. *Journal of Cleaner Production, 47*, 283-297.
Monczka, R.M., Handfield, R.B., Giunipero, L.C., & Patterson, J.L. (2014). *Purchasing and Supply Chain Management.* 6th Edition, New York: McGraw-Hill.
Narasimhan, R., & Carter, J.R. (1998). *Environmental Supply Chain Management.* The Center for Advanced Purchasing Studies, Arizona State University Tempe, AZ.
Scott, L. (2005). Twenty First Century Leadership. Retrieved from https://corporate.walmart.com/_news_/executive-viewpoints/twenty-first-century-leadership on May 10, 2019.
Smith, H. J., & Hasnas, J. (1999). Ethics and information systems: The corporate domain. *MIS Quarterly, 23*(1). 109-127.
Tseng, M. L., Islam, M. S., Karia, N., Fauzi, F. A., & Afrin, S. (2019). A literature review on green supply chain management: Trends and future challenges. *Resources, Conservation and Recycling, 141*, 145-162.

Walton, S.V., Handfield, R.B., & Melnyk, S.A. (1998). The green supply chain: integrating suppliers into environmental management process. *International Journal of Purchasing and Materials Management, 34*(2), 2-11.

Zhu, Q., & Cote, R. (2004) Integrating green supply chain management into an embryonic eco-industrial development: a case study of the Guitang Group. *Journal of Cleaner Production, 12*(8-10), 1025-35.

Zhu, Q., & Sarkis, J. (2004). Relationships between operational practices and performance among early adopters of green supply chain management practices in Chinese manufacturing enterprises. *Journal of Operations Management, 22*(3), 265-89.

LEVERAGING ONLINE SOCIAL RELATIONS TO IMPROVE THE DIVERSITY OF PERSONALIZED RECOMMENDATION LIST

(Full Paper Submission)

Xiaoyun He
Auburn University at Montgomery
xhe@aum.edu

ABSTRACT

Recommender systems typically focus on using accuracy as the key metric to evaluate the performance of the recommendations. However, recent studies suggest that recommending a diverse list of products enhances user satisfaction and is positively associated with customer retention rates. In this study, we propose to leverage the social relations of a user in an online environment to improve the diversity of personalized recommendation list. Our preliminary empirical results indicate that our proposed approach performs well in increasing the recommendation diversity while maintaining comparable level of accuracy.

KEYWORDS: social relations, the diversity, online recommendation, personalized list

INTRODUCTION

With the trend of consumers spending more time and money on online shopping than ever before, recommender systems have been widely used by online businesses to recommend products that are most likely to be of interest to users (Jannach et al. 2010). According to a Forrester Research report, one third of online customers that notice these recommendations wind up purchasing the recommended product (Mulpuru 2006). Given the increasingly intensive competition in online retailing, improving the quality of product recommendations becomes a critical factor for companies to gain sustainable advantage.

Traditional recommender systems have typically used the metrics such as accuracy and coverage to measure the performance of the recommendations (Adomavicius & Tuzhilin 2005). For example, the 1 million Netflix Prize (netflixprize.com) sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences. However, recent studies have argued that the quality of recommendations should go beyond accuracy. As one of the goals of recommender systems is to provide a user with highly idiosyncratic or personalized items, more diverse recommendation list would result in more opportunities for users to be recommended such items (e.g., Adomavicius & Kwon 2009; Zhang & Hurley 2008; Zhang, et al. 2012; Ziegler et al. 2005). Park & Han (2013) find that recommending a diverse list of products enhances user satisfaction and is positively associated with customer retention rates. These, in turn, would lead to the boost of sales revenue and the bottom line.

In a nutshell, a basic recommendation method seeks to predict the 'preference' or 'rating' that a user would give to an item, such as music, books, movies, etc. that has not been seen by the user. A rating indicates how a particular user liked a particular item, e.g., Jane Betty gave the book ``Thinking, Fast and Slow'' the rating of 4 (out of 5). These ``particular'' recommendations are often referred as personalized recommendations (Kim et al. 2003). The task of personalized recommendation requires the ability to predict which items will be considered interesting by the user. The personalized recommendations have been viewed as an important source to assist and augment the natural social recommendation process - in our everyday life, we rely on recommendations from our social circles by word of mouth (Shardanand & Maes 1995).

The research on social relations and consumer decision making indicates that social influence is an important factor of affecting people's decision making and consumer decisions are best understood in the social contexts in which these decisions are made (e.g., Deutsch & Gerard 1955; Simpson et al. 2012). For instance, we tend to be influenced by our friends in terms of products purchase. Social influence not only occurs among the directly connected acquaintances, also propagates via the chain of online connections in a social network (Richardson & Domingos 2002). In addition, recent studies suggest that online social relations influence users' both product choices and ratings (e.g., He et al. 2017). Within an informal community of users and social context, it is therefore natural to incorporate the factor of social relations into recommendation model to enhance the diversity of the personalized recommendation list.

In this study, we utilize a data set collected from an online Web 2.0 site. The website integrates both online social networking and online product rating functions. In order to achieve recommendation diversity while keeping up with accuracy, we put more weight on the products rated by the focus user's social circle while limiting the degree of dissimilarity in user preferences. The preliminary empirical evaluation results show that our proposed approach performs in increasing the diversity while maintaining comparable level of accuracy in the recommendations.

LITERATURE REVIEW AND RELATED WORK

Collaborative Filtering based Recommendation Techniques

Over the past two decades, collaborative filtering techniques have been considered as the most commonly adopted in developing the recommender systems (Adomavicius & Tuzhilin 2005). These techniques rely on the past user behavior (e.g., their previous purchase transactions or product ratings) to predict what a user may like. The collaborative filtering methods can be generally classified into two categories: neighborhood-based methods and model-based methods.

The neighborhood-based approaches are centered on finding the relationship between users or, alternatively, between items. The initial form of neighborhood-based approaches is user-based (e.g., Herlocker et al. 1999). In the user-oriented case, the system predicts a user's rating based on the known ratings of similar users, i.e., the like-minded users. Later, an analogous item-based approach was proposed (e.g., Deshpande & Karypis 2004). In this case, a rating is computed based on past ratings made by the same user on similar items.

The model-based approaches typically involve two steps. First, training data consisting of past user-item ratings is used to train a predefined model. Then, the model is used to predict the unknown ratings. In this category, various models have been proposed, including clustering model (Merialdo 1999), latent semantic model (Hofmann 2004), ranking-based model (Liu & Yang 2008), to name a few.

Recommendation Based on Social Relations

Exploiting social networks to enhance recommendation performance has been studied by the researchers (e.g., Kautz et al. 1997; Guy et al. 2009; Liu & Lee 2010; Ma et al. 2011). Kautz et al. (1997) propose to build an interactive ReferralWeb system that uses social network to make recommendation / search more focused and effective. They argue that the referral chain arising from a social network enables the utilization of experts' expertise and the evaluation of trustworthiness of the expert. The explorative study in Guy et al. (2009) demonstrates a clear superiority of users' familiar network (e.g., being connected in a social network) over similarity network (e.g., having similar tastes as expressed in item ratings) as a basis for recommendation. Liu and Lee point out that their evaluation results on recommendations indicate that more accurate prediction algorithm can be developed by incorporating social network information into collaborative filtering (Liu & Lee 2010). Ma et al. (2011) propose a matrix factorization framework to include the social information as regularization terms into recommendation systems. The similarity between a pair of user in the social network is still based on the past ratings of the users. That is, the social information is based on users' similarity network.

The Diversity of the Recommendations and User Satisfaction

There has been considerable understanding that high recommendation accuracy alone does not necessarily provide users of recommender systems satisfying experience (Herlocker et al. 2004; Ziegler et al. 2005). As one of the goals of recommender systems is to provide a user with highly idiosyncratic or personalized items, more diverse recommendation list would result in more opportunities for users to be recommended such items (e.g., Adomavicius & Kwon 2009; Zhang & Hurley 2008; Zhang, et al. 2012; Ziegler et al. 2005). Diversification is defined as the process of maximizing the variety of items in a recommendation list. Diversity may refer to all kinds of features, e.g., genre, author, and other discerning characteristics (Ziegler et al. 2005). Park & Han (2013) find that recommending a diverse list of products enhances user satisfaction and is positively associated with customer retention rates.

Some recent studies have examined the principle of diversity to improve user satisfaction (Adamopoulos & Tuzhilin, 2015). For example, Ziegler et al. (2005) propose an approach that intends to balance and diversify personalized recommendation lists in order to reflect the user’s spectrum of interest. They show that the user’s overall liking of recommendation lists goes beyond accuracy and involves other factors, e.g., the user’s perceived diversity of the list. Hurley and Zhang (2011) argue that the motivation of diversity research is to increase the probability of retrieving unusual or novel items which are relevant to the user and introduce a methodology to evaluate their performance in terms of novel item retrieval. Adomavicius and Kwon (2009) define the concept of aggregated diversity as the ability of a system to recommend across all users as many different items as possible through a controlled promotion of less popular items toward the top of the recommendation lists while limiting the loss of accuracy.

RESEARCH CONTEXT AND METHODOLOGY

Research Context and Data

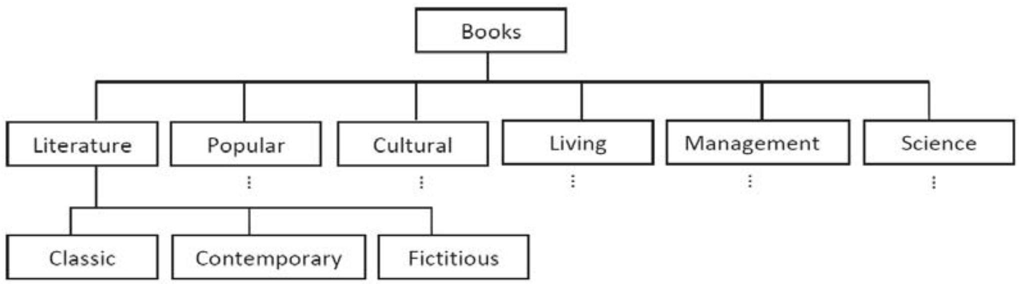
The data was obtained from a Web 2.0 site, which wishes to remain anonymous. The website offers the integrated social networking and online rating services to the users. It provides a unique platform for users to establish online social relations and to participate in rating and reviewing the products that they have consumed. The data of users' relationships establishment and rating participations generated from this platform enables us to empirically leverage the effects of online social relations so as to improve personalized product recommendation.

At the web site, a user may use a valid email address to register and create a screen name as the user identifier that is shown to the other registered users. The screen name does not have to be a user's real name, thus it offers certain level of anonymity. Over 95% of the registered users chose to use a screen name that is different from their real name. Registered users can rate and write reviews for specific product items. These ratings are available to be viewed by all other registered users. The products fall into one of three product categories - books, music, and movies. The rating scale is from 1 to 5 stars with 1 being the lowest and 5 being the highest level. In addition, the web site also supports online social network service that a registered user can choose to follow another registered user by simply clicking a "Follow" button. Unlike other well-known social networking sites such as Facebook or LinkedIn, the establishment of "following" relations does not need to be consented / invited by the followed user. Based on a survey of the registered users, about 87% of the respondents stated that they follow other users whom they do not know in person. In other words, the online "following" relations in this context are largely different from those in the physical offline world.

Consequently, different from the undirected social relations in a typical social network, the "following" relationships created at the website are directed, which forms a directed social network. Specifically, given a pair of users, namely user u_1 and user u_2, if user u_1 chooses to follow user u_2, then it represents that there is a directed link from u_1 to u_2. However, if user u_2 does not follow user u_1, then there is no directed link from u_2 to u_1. Then, user u_2 may have potential social influence on user u_1 in terms of product ratings and adoption. Such online social network can be modeled by a directed graph, where each node represents a user and directed links represent the established "following" relations between users. User at the tail follows the user at the head of arrow as a friend. The extracted data include the users' following relationships as well as review ratings.

The web site provides relatively simplified two-level hierarchical classification taxonomies for the product categories. For example, books are organized into 6 first-level categories - Literature, Popular, Cultural, Living, Management, and Science; each of these categories is further classified into second-level sub-categories. For instance, Literature is classified into Classic, Contemporary, Fictitious, and so on. See Figure 1 for a part of the taxonomy for books. Music is classified in terms of genre, country/region, artists, and year produced. Movies are classified in terms of genre, country/region, and artists.

Figure 1. A Sample Taxonomy for Books



The final sample data span the period from January 2008 to August 2011. The data that we used contain the following data sets: (1) Books: the rating data for books; (2) Music: the rating data for music; (3) Movies: the rating data for movies; (4) OnlineSN: the online social network that describes that “following” relationships among users including user identifiers and date when the relation was established. For each of the categories of items, the rating data include the following attributes: itemID (the specific item that was rated); userID (the user who rated the item); rating (an integer number between 1 and 5 assigned by the user to rate the item); date (when the rating was provided); (5) the taxonomies for books, music, and movies.

Proposed Approach

In this section, we first provide background knowledge and information model related to recommendation method that we use. Then, we discuss the detailed steps to generate the top-N recommendation list, including how social relations are leveraged to enhance diversity of the list while maintaining comparable level of accuracy.

Preliminaries and Collaborative Filtering Recommendation

For clarity and simplicity, we use the following notations to represent the various information. For the users, we use $U = \{u_1, u_2, \ldots, u_m\}$ to represent all the users. The list of product items are notated by $P = \{p_1, p_2, \ldots, p_n\}$. Each user u_i has a set of product ratings represented by R_i. Our approach utilizes the product taxonomy (see Figure 1. as a partial example) to create content-based user profiles. Such product taxonomy T is represented by $C = \{ c_1, c_2, \ldots, c_l\}$. Set T contains the categories of product classification. Graphically, the taxonomy has a tree structure. Each c_i represents the specific category that product p_i falls into. At the bottom of the product taxonomy, the category is the narrowest. From bottom up, the category becomes broader. For any pair of categories c_i and c_j $(i \neq j)$, $T(c_i) \cap T(c_j) = Ø$ holds.

In general, in order to discover like-minded users, we compute content-based user profiles that is consistent with the “collaboration via content” in recommendations (Ziegler et al. 2004). The similarity of the users are based on user profiles. Each user’s neighborhood with the most similar users is formed. The recommender focuses on products rated by those neighbors and generates top-N recommendation lists. The product ranking depends on the proximity of its classification categories with respect to a user profile. Specific details are provided in the next section.

The Steps to Generate Diverse Recommendation List

Below, we describe the detailed steps of our approach that leverages social relations to generate a diverse recommendation list while without losing much accuracy.

Step 1. Create user profiles. Traditional collaborative filtering techniques represent user profiles by the vectors of their product ratings. Then, the computation of the similarity between a pair of users is to apply Pearson correlation formula to their respective rating vectors. As our objective is to leverage the social relations to enhance the diversity of recommendations, we follow the work of Ziegler et al. (2004) to generate user profiles that are based on product taxonomy instead. Specifically, a user profile is represented by a vector of interest scores. Given a user u_i, the interest scores vector $sv_i = (sc_1, sc_2, \ldots, sc_l)$. Depending on the product categories in which the products rated by a user fall, each of the scores measures the degree of interest in a particular product category that the user has. Hence, the size of the vector sv_i equals to the number of distinct product categories in the product taxonomy, denoted by *l*. Score assigned to the upper-categories decays with increasing distance from leaf node in the taxonomy tree. To ensure a common scale, the user profile vectors are normalized to a fixed overall score *s*. In order to reflect the relationship between categories in the tree taxonomy, the interest score assigned to each of the product category along the tree path has to meet certain criteria. Given a category c_i, let $(p_0, p_1, \ldots, p_q)$ denote the path from its topmost super-category to its immediate super-category, then $\Sigma sc(p_m) = sc(c_i)$, where $0 \le m \le q$. In addition, the interest score $sc(p_m)$ depends on the score of its sub-category p_{m+1} and the number of its sub-categories, denoted by n_s. Specifically, $sc(p_m) = \lambda\ (sc(p_{m+1})/\ (n_s+1))$, where λ is a propagation factor to fine tune the process of profile generation.

Step 2. Form the neighborhood. First, for a given user u_i, a proximity measure is used to find the users that are most similar to u_i. Specifically, we use Pearson Correlation to compute the similarity between a pair of users, say u_i and u_j with the profile score vector sv_i and sv_j respectively. We use $SIM(u_i, u_j)$ to denote the similarity. Then, $SIM(u_i, u_j) = \Sigma(sv_k – M(sv_i))(sv_k – M(sv_j))\ /sqrt(\Sigma(sv_k – M(sv_i))^2(sv_k – M(sv_j))^2)$, where $0 \le k \le l$, and $M(sv_i)$ and $M(sv_j)$ are the mean value for vector sv_j and vector sv_j respectively. Because we normalize the user profile vectors, $M(sv_i) = M(sv_j) = s\ /\ l$. Values for $SIM(u_i, u_j)$ range from -1 to +1, where positive values indicate positive correlation, and negative values means negative correlation. For generic collaborative filtering, users who have rated many common products would have high similarity. In our case, as user profile generation is based on product taxonomy, a pair of users may have high similarity even though they have little or no common rated products. The more score values two profiles sv_i and sv_j have accumulated in the same paths of the taxonomy, the more similar they are. Then, the neighborhood of user u_i, denoted by $NB(u_i)$, is formed by including the top-M users based on the SIM values in descending order.

Step 3. Generate the diverse recommendation list. First, given a focus user u_i, if a product p_k is rated by one of u_i's neighborhood users derived from the previous step but not yet by the user u_i, then p_k is added to the candidate list. The relevance level of product p_k to user u_i is then computed. The relevance level may dependent on multiple factors such as the similarity between user u_i and the users that have rated p_k, how close the product p_k is considered to be close to the user u_i's interest profile. The latter is an important measure that is supplemental to the user similarity, as a user may be interested in something that is beyond what is specified in the focus user's profile. The computation of product similarity $SIM(u_i, p_k)$ follows that of user similarity through the creation of a dummy user u_d. That is, $SIM(u_i, p_k) = SIM(u_i, u_d)$. We use the same formula described in Ziegler et al. (2004) to compute the relevance level of product p_k for the focus user u_i. The relevance level is defined as $w_i(p_k)$ = q SIM(u_i, p_k) ΣSIM(u_i, u_j)/(|NB(u_i, p_k)| + Y_R), where u_j Є NB(u_i), |NB(u_i, p_k)| denotes the number of users that are in the set of NB(u_i) and have rated product p_k, q = (1.0 + $N_k T_\Delta$) (N_k is the number of categories in which product p_k falls), and T_Δ and Y_R are fine-tuning parameters that allow for customizing the recommendation process. T_Δ rewards those products that fall under relatively large number of categories in the taxonomy. Large Y_R makes popular products acquire higher relevance level, which makes it possible for users to be recommended those popular items, while low Y_R treats popular and new items in the same manner, thus helping to alleviate the cold start problem in recommendations.

Next, once all the relevance levels $w_i(p_k)$ are computed for the focus user u_i, they are ranked in descending order, and their corresponding products are considered as the baseline recommendation list, denoted by $L_{(w_i)}$. In order to diversify the baseline recommendation list, Ziegler et al. (2004) proposes a greedy re-ranking algorithm by iteratively re-ranking items based on the dissimilarity to the preceding items in the list. One issue with this approach is that the diversity is gained at the expense of recommendation accuracy. To measure the diversity of a given recommendation list $L_{(w_i)}$, intra-list similarity ILS($L_{(w_i)}$) is defined as follows: ILS($L_{(w_i)}$) = ($\Sigma p_k \Sigma p_e$ SIM(p_k, p_e))/2, where p_k, p_e Є $L_{(w_i)}$ and $p_k \neq p_e$. Higher value indicates lower diversity.

Instead, we propose to leverage the focus user's social circles to enhance the diversity of the baseline recommendation list. Specifically, we use CIR(u_i, FL) to denote user u_i's online circle of friends (e.g., either being followed or following) at the level FL. If FL = 1, then CIR(u_i, 1) indicates there is a direct following or being followed relationship between the focus user u_i and the friends in the set of CIR(u_i, 1); FL = 2 means that CIR(u_i, 2) is CIR(u_i, 1) plus the friend's friend. In other words, the relation is indirect with path length of 2; FL = 3 then denotes CIR(u_i, 2) plus friend's friends' friend, and so on. The procedure below describes the basic idea of our approach.

Input: Users' profiles, baseline recommendation list $L_{(w_i)}$, parameters: N, FL, Z, and ε.
Output: Diversified top-N recommendation list $DL_{(w_i)}$.

Identify the product list P_f = {p_1, p_2, … } that the users in CIR(u_i, FL) have rated but not
    being ranked as top-N in $L_{(w_i)}$;
Compute the similarity between each product p_f and the focus user u_i's profile;
Rank the above similarities in the descending order to create a list DSIM;
for each of the top-Z corresponding products in DSIM do
    Temporarily substitute it for the last ranking product in top-N list of $L_{(w_i)}$;
    Compute the new intra-list similarity ILS($L_{(w_i)}$)*;
    If (ILS($L_{(w_i)}$) - ILS($L_{(w_i)}$)*) > ε then
        Make the substitute permanent in the Top-N list of $DL_{(w_i)}$;
    end
end

We note that the parameters M and ε provide us certain degree of flexibility in balancing the trade-off between diversity and accuracy in the process of generating the top-N recommendation list.

EXPERIMENTAL EVALUATION AND PRELIMINARY RESULTS

In this section, we describe the metrics used for performance evaluation, present the evaluation results, and discuss the preliminary findings from our experimental evaluations. Evaluation metrics are essential in order to measure the quality and performance of recommendation approaches. We consider accuracy measurements such as precision and recall as well as non-accuracy evaluation metrics such as intra-list similarity.

As our approach does not actually predict the ratings, we cannot use the accuracy metrics that measure how close predicted ratings come to true user ratings. Such metrics include mean absolute error (MAE), the Root Mean Square Error (RMSE), etc. Instead, we use precision and recall. Recall indicates the percentage of relevant items that were included in the recommendation list, and precision gives the percentage of recommended products that are relevant. Suppose that the total number of products in test set is n_t, and out of these products, n_p of them appear in the top-N recommendation list, then recall is defined as: Recall = $100 \cdot (n_p / n_t)$. Accordingly, precision represents the percentage of test set products occurring in the recommendation list with respect to N (i.e., the size of the recommendation list): Precision = $100 \cdot (n_p / N)$. The diversity of a given recommendation list $L_{(w'i)}$ is measured by intra-list similarity $ILS(L_{(w'i)})$ defined in the previous section. In order to obtain "global" values, we average the respective metric values for all evaluated users.

We perform 5-fold cross validation in our experimental evaluations. In each fold, we use 80% of randomly selected data as training set and the remaining 20% as the test set. For the Books, Music, and Movies data, the observations for relative rankings are 1,052, 1,237, and 833 respectively. For our preliminary experiments, N is set to 20, FL is set to 3, Z = 30, and ε = 0.10.

In order to demonstrate the effectiveness of our proposed approach, we compare it with the baseline non-diversified collaborative filtering method. Table 1. Presents the preliminary results.

Table 1. Preliminary Experimental Results

Category	Methods	Recall	Precision	ILS
Books	Proposed	6.02	3.73	9.29
	Baseline	6.11	3.82	14.78
Music	Proposed	7.53	4.42	13.80
	Baseline	7.66	4.51	16.18
Movies	Proposed	5.84	3.46	21.32
	Baseline	6.02	3.77	30.65

At this point, we were only able to obtain the results from one set of parameters and the datasets are also relatively small. Thus, in terms of accuracy and intra-list similarity measures, we cannot draw definite conclusion yet on whether the differences between our proposed method and baseline approach are statistically significant. However, based on the closeness of the accuracies (i.e., measured by recall and precision) with lower inter-list similarity for our approach across all three product categories, it seems to be a good indication that our proposed approach might be able to perform well in enhancing the diversity of recommendation list while maintaining comparable level of accuracy. That is, our approach could be promising.

As we put more weight on the rated products from a user's online social circle, the above results suggest that leveraging the social relations data from online social circles may help to improve the diversity of personalized recommendation. In the context of online social networks, the ratings from a user's online social circle can be considered as a proxy for online social influence. Thus we believe that our findings might have practical implications for online retailers to take advantage of online social relations.

CONCLUSION AND FUTURE WORK

There has been considerable understanding that high recommendation accuracy alone does not necessarily provide users of recommender systems satisfying experience. Recent studies suggest that recommending a diverse list of products enhances user satisfaction and is positively associated with customer retention rates. In this study, we propose to leverage the social relations of a user in an online environment to improve the diversity of personalized recommendation list.

Our initial experiments show the promise of our proposed approach. However, the dataset that we have obtained so far are relatively small, and we have only tested our approach on one set of parameters. Obtaining and preprocessing the data is a time consuming process. We plan to expand our dataset to include more users, and also apply our approach to different levels of social circles. In the addition, we would like to experiment with different sets of parameters to find out how they might affect the accuracy and diversity of the recommendations. We also intend to compare our approach with the existing approaches to demonstrate whether our approach is computationally more efficient. It is worth noting that the quality of recommendations can be reflected in different dimensions, such as novelty, unexpectedness (e.g., Adamopoulos & Tuzhilin 2015; Hurley & Zhang 2011), etc. As this study only represents our initial effort in developing a novel approach, our primary focus is on the diversity. In this regard, future direction may include exploring the new dimensions of recommendation quality.

ACKNOWLEGEMENTS

We would like to express our gratitude to Auburn University at Montgomery for financial support. We also thank Qiyuan Peng for providing research assistance.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Project Evaluation for a Media Company — A Case Study

Massoud M. Saghafi, Department of Marketing, Fowler College of Business
(619) 594-6358, msaghafi@sdsu.edu

Feraidoon (Fred) Raafat, Department of Management Information Systems,
Fowler College of Business (619) 594-5685, fred.raafat@sdsu.edu

Nathan Oestreich, Charles W. Lamden School of Accountancy,
Fowler College of Business (619) 594-5070, drno@sdsu.edu
San Diego State University, 5500 Campanile Drive, San Diego, CA 92182

ABSTRACT"

Trends toward digital news media have driven readers away from most printed media causing many of them to reassess their products and processes. To spur growth in readership and revenue, media companies need to develop strong niche media products and companion digital components that are standardized and scalable for evaluating the potential success or failure of new product ideas to efficiently supplement their product portfolios. This paper describes a study conducted for a regional newspaper to achieve this objective.

KEYWORDS: printed media, digital media, market analysis, and new product development

INTRODUCTION

A California based media company (CBMC) was interested to explore and identify approaches to increase its revenue and readership as recent trends toward digital news media had driven readers away from printed media, causing this company to reassess its products and processes, which it had impacted its readership. This company had achieved a market share advantage within its region for printed news media and was a low-cost leader with extensive distribution reach. CBMC asked our university team to assist in market analysis and developing an easy, scalable method to evaluate new proposals for various products in its portfolio. Our team consisted of graduating MBA along with faculty who undertook to provide guidance. This paper describes our attempts in assisting this company to achieve its objective.

BACKGROUND

CBMC is a Pulitzer Prize winning daily newspaper published California. Its ownership has changed hands several times in the past two decades. It currently has about 250,000 daily and 400,000 Sunday readers. Its reach extends with its many locally distributed daily and weekly newspapers, entertainment guides, Spanish language editions, and website. It offers information on various topics such as sports, business, lifestyle, military, opinion, videos, obituaries, deals and classifieds, jobs, real estate, etc. With the exponential rise in digital media, the print medium has suffered a significant decline of 8-10% annually. Also, the majority of print media customers are the Baby Boomer generation (50 – 70-year-old). To drive growth in readership and revenues CBMC was looking to launch new products for its print media, as well as its digital medium that would be relevant to its targeted markets. In the past, when it had developed products, it tried to achieve a 100% market share, however, due to dwindling readership, the company is considering other print products to capitalize on its market position.

Historically, the news media industry tended to be very hierarchical in nature, resulting in inefficient and inconsistent decision making and not adept at managing change. CBMC's existing processes for new product development were not well defined or standardized. It did not offer an open platform or equal opportunity for everyone to pitch new ideas. Only staff members with most influence were able to push new product ideas with no requirements for data collection and analysis. Management was receiving about three to four new product proposals every month with varying degree of completeness, lacking clear understanding of the proposal requirements. To review these proposals, the management would rely on their own intuition as to the appeal of the proposal, the conviction of the presenter, and the perceived feasibility to decide on an ad hoc basis, and often without the underlying supporting data. This had resulted in failed product ideas. While intuition-based product decisions had worked in the past when CBMC enjoyed a dominant market share, its management decided to systematize its approach based on data-driven decision processes.

METHODOLOGY

To achieve the goals of the project, the following activities were pursued:

1. Formulated a functional and consistent new product development process to identify and prioritize key assessment criteria. The process was then used to evaluate the potential profitability of a new product's go-to-market plan.
2. Created a templated to input the data obtained in accordance with the new product process criteria and applies weights based on the priorities determined in step 1.
3. Demonstrated the functionality of the new product development process by empirically testing and calibrating the template using recent CBMC home improvement media product idea.

We first identified and researched the target market in which CBMC intended to serve and established standards for determining if a product would generate a sufficient primary and secondary audience. At this stage the basic data, such as demographic, ethnographic, and behavioral were researched and applied to the process. Then we considered plausible revenues that can be generated by analyzing CBMC's advertising revenue and product payback/profitability targets to establish a standardized methodology for projecting a new product's likely revenue generation. We also considered project startup and recurring costs, by reviewing CBMC's product portfolio to create cost projections for each of the Company's various product categories. Using these metrics, we established a template for projecting a new product's associated costs. Finally, we provided guidance for new product market release by examining CBMC's current portfolio success metrics to establish go-to-market guidance.

CBMC NEW PRODUCT DEVELOPMENT PROCESS

1. Application: CBMC Team Members who wish to propose a new product idea to the Vice President of Strategy and Operations.

2. Objective: To establish the key criteria and data inputs required to present a new product idea, in order to create a standardized, data driven evaluation environment.

3. Responsibilities:

   a. Idea Lead: The creator of the new product idea, who is responsible for collecting the standard data required to propose a new product.
   b. Research Lead: The research department member responsible for assisting the Idea Lead with any background research required to pitch a new product idea.
   c. Management Lead: The management representative responsible for verifying new product idea assumptions and using the excel model to make cost and revenue projections. Ultimately, the Management Lead will make the final go-to market determination based on the objective data.
   d. Marketing Lead: The marketing representative responsible for securing advertising sales prior to initial product release.

4. Outputs (Available upon request):

   a. New Product Idea Submission Form.
   b. New Product Evaluation Model: Excel

5. Process

   a. Idea Submission: The Idea Lead, who wishes to propose a new product idea, must first fill out the New Product Idea Submission Form. The proposer might already have that data on hand or may need to obtain it through relevant sources, with the assistance of the research lead. The required data is as follows:
      1. Develop a new product idea: The Idea Lead must identify the details of the proposed new product, including what the new product/service offers and what consumer needs it fulfills.
      2. Develop a consumer profile: The Idea Lead must create the optimal profile and characteristics of the anticipated new product user. This includes gender, age group, income group, education, geographical location, family/marital status etc. Surveys may be required to assist with profiling potential consumers. The parameters for any required surveys will be coordinated by the Idea Lead through the Research Lead.
      3. Define consumer target market: The Idea Lead must work with the Research Lead to funnel the population into a specialized target market in terms of households or page views.
      4. Define advertising market: The Idea Lead will gather data regarding the potential advertising market in the target category. Relevant data includes number of advertisers, spending by media product category, and any potential advertising leads for the new product. The Research Lead will assist with gathering any advertising spending surveys and/or data.
      5. Define the category competitive landscape: The idea lead will gather data regarding the competitive landscape within the proposed new product media category. This data includes the number of competitors, platform of each competitor's product, quality of

content (rated low, medium, high). Approximate revenue of competitors, if available. The Idea Lead should also score the overall competitive landscape on a scale from 1-10 (10 being most competitive) and attempt to determine whether the competitors are profitable or not.

6. Refine new product idea: Based on the acquired background research, the Idea Lead will provide a recommendation on product platform (i.e. tabloid, glossy, web), distribution frequency and mode, page count, and suggested content.
7. Submit the New Product Idea Submission form: Once all required data inputs are gathered, the Idea Lead must fill out the relevant form, and provide to the Management Lead.

b. Management Review: The Management Lead will then review the New Product Idea Submission form, including any assumptions made and evaluate for feasibility and profitability as follows:
   1. Review New Product Idea Submission Form: The Management Lead will review the submission to determine if all required data inputs are correctly identified, and accurate in nature.
   2. Input target market assumptions: The Management Lead will input the target market assumptions including households, distribution frequency and mode, and suggested page count into the New Product Evaluation Model Template (Excel).
   3. Set success metrics: Using the New Product Evaluation Model, the Management Lead will set profitability and market share targets for the new product.
   4. Go or no-go decision: Using the New Product Evaluation Model, the Management Lead will make a data-driven decision on whether the new product idea would be a profitable venture. The Management Lead will review the required revenue to become profitable and work with the Marketing Lead to determine if target advertising revenue can be achieved prior to initial product release.

c. Trial Phase: If a new product is released, the Management Lead will monitor actual revenue and readership data against defined success metrics to determine whether a new product's production should continue.
   1. Breakeven: The Management Lead will establish a breakeven point, at which point the new product must have recovered all costs associated with its initial release and production. If the product has not hit breakeven by the defined checkpoint, production should be suspended. The Management Lead will use the New Product Evaluation Model Template to graphically track the product's monthly progress.
   2. Profitability: The Management Lead will define a checkpoint at which the product must achieve its target profitability. The Management Lead will use the New Product Evaluation Model Template to graphically track the product's monthly profitability.
   3. Market share: The Management Lead will also track the new product's actual market share against its target. The Management Lead will use the New Product Evaluation Model Template to graphically track the product's monthly profitability.

NEW PRODUCT IDEA FORM TEMPLATE

We created a template form for pitching new product ideas. The form includes all the data and research requirements so that CBMC staff members know the exact expectations of a new product pitch. It guides the user through the data collection and research steps with specific questions to answer. We further integrated this form into an online application. This form serves to simplify the idea submission process and collect the necessary data. To create a New Product Evaluation Model, we researched CBMC’s reader base with regard to households in for a specific area. This data was included in a spreadsheet that can be easily sorted by household size, income, education, home value, etc.

Through several interviews with CBMC management, we defined key cost and revenue metrics and built them into an Excel Template; we included a control panel that would allow for customization the assumptions change. This allows the user to model and project costs across various product platforms based on key criteria such as product size, page count, distribution frequency and mode, content, and target households as shown in Figure 1.

CBMC New Product Model

Data Inputs (Monthly)

	Product		
Inputs (Projections)	**Tabloid**	**Glossy**	**Website**
Target Households / Web Views	<Enter>	<Enter>	<Enter>
Distro / Update Frequency (Month)	<Enter>	<Enter>	<Enter>
Pages	<Enter>	<Enter>	<Enter>
Width (inches)	<Enter>	<Enter>	
Height (inches)	<Enter>	<Enter>	
Stories per page	<Enter>	<Enter>	<Enter>
Direct Mail	<yes or no>	<yes or no>	
Target Markup (%)	< % >	< % >	
Category Annual Ad Market	<Enter>	<Enter>	<Enter>
Target Market Share	< % >	< % >	< % >

Data Outputs

Estimated Cost Output			
Outputs	**Tabloid**	**Glossy**	**Website**
Startup Cost			
Cost Per Page			
Monthly Cost			
Reve ıue Target Output			
Target Ad Revenue / Pg based on Profit			
Target Ad Revenue / Pg based on Mkt Share			
Min Annual Ad Market			
Test: Profit Tgt matches Mkt Share Tgt?			
Test: Actual Ad Market > Min Ad Mkt?	yes	yes	yes
Assumptions Make Sense?	no	no	no

Figure 1: New Product Development Model Excel, Tab 1

We also built in an assumption checker to ensure that the markup of the target matches market share target, and the category advertising revenue can support the new product based on our projected market share target. If the assumptions do not align, the user can adjust the inputs until they do. Per Example in Figure 2, if the markup-based revenue target was calculated much below the market share-based revenue target, the user would need to increase distribution type inputs (i.e. page count/distribution frequency) in order to create additional ad revenue space to match the market share target.

		Product
Data Inputs (Monthly)	**Inputs (Projections)**	**Glossy**
	Target Households / Web Views	200000
	Distro / Update Frequency (Month)	2
	Pages	15
	Width (inches)	6
	Height (inches)	8
	Stories per page	2
	Direct Mail	no
	Target Markup (%)	60%
	Category Annual Ad Market	$ 5,300,000
	Target Market Share	45%
Data Outputs	**Estimated Cost Output**	
	Outputs	**Glossy**
	Startup Cost	$ -
	Cost Per Page	$ 1,295
	Monthly Cost	$ 38,856
	Revenue Target Output	
	Target Ad Revenue / Pg based on Profit	$ 2,072
	Target Ad Revenue / Pg based on Mkt Share	$ 6,625
	Min Annual Ad Market	$ 1,657,840
	Test: Profit Tgt matches Mkt Share Tgt?	no
	Test: Actual Ad Market > Min Ad Mkt?	yes
	Assumptions Make Sense?	no

Figure 3 shows the changes made to the inputs (page count increased to forty-three) that result in the “yes” feedback, indicating that the ad revenue target now matches the market share target.

		Product
Data Inputs (Monthly)	**Inputs (Projections)**	**Glossy**
	Target Households / Web Views	200000
	Distro / Update Frequency (Month)	2
	Pages	43
	Width (inches)	6
	Height (inches)	8
	Stories per page	2
	Direct Mail	no
	Target Markup (%)	60%
	Category Annual Ad Market	$ 5,300,000
	Target Market Share	45%
Data Outputs	**Estimated Cost Output**	
	Outputs	**Glossy**
	Startup Cost	$ -
	Cost Per Page	$ 1,295
	Monthly Cost	$ 111,386
	Revenue Target Output	
	Target Ad Revenue / Pg based on Profit	$ 2,072
	Target Ad Revenue / Pg based on Mkt Share	$ 2,311
	Min Annual Ad Market	$ 4,752,473
	Test: Profit Tgt matches Mkt Share Tgt?	yes
	Test: Actual Ad Market > Min Ad Mkt?	yes
	Assumptions Make Sense?	yes

Along with CBMC management we established key success metrics that management found critical to a new product release. The first metric was to ensure that the product achieved a baseline advertising commitment prior to release, which was dependent on the cost of the product. The second was to achieve break-even at six months. We used a formula based approach to determine target revenue for the first six months to ensure break-even. The formula dynamically adjusts profit markup to display the required revenue for the initial release, and for each of the first six months to break-even, as depicted in Figure 4 and Figure 5.

Glossy			Inputs			
Month	Market Share Target	Revenue Target	Actual Revenue	Actual Market Share	Cum Profit Target	Cum Profit Actual
1	3%	$ 76,856	$ 85,000	19%	$ (34,530)	$ (26,386)
2	5%	$ 86,881	$ 95,000	22%	$ (59,035)	$ (42,772)
3	8%	$ 100,247	$ 110,000	25%	$ (70,173)	$ (44,158)
4	10%	$ 113,614	$ 117,000	26%	$ (67,946)	$ (38,544)
5	13%	$ 130,322	$ 128,000	29%	$ (49,010)	$ (21,930)
6	15%	$ 160,396	$ 190,000	43%	$ -	$ 56,683
7	18%	$ 169,307	$ 199,000	45%	$ 57,921	$ 144,297
8	20%	$ 173,762	$ 250,000	57%	$ 120,297	$ 282,911
9	23%	$ 175,990			$ 184,901	
10	25%	$ 177,104			$ 250,619	
11	28%	$ 177,661			$ 316,893	
12	30%	$ 177,939			$ 383,447	
13	33%	$ 178,079			$ 450,139	
14	35%	$ 178,148			$ 516,901	
15	38%	$ 178,183			$ 583,698	
16	40%	$ 178,200			$ 650,512	
17	43%	$ 178,209			$ 717,335	
18	45%	$ 178,213			$ 784,162	

FIGURE 5: NEW PRODUCT DEVELOPMENT MODEL EXCEL, TAB 3A (MONTHLY REVENUE, TARGETS V. ACTUALS)

CONCLUSION

We developed a work-flow process to establish steps to generate a new product idea, from initial research to final management evaluation. We created an online Product Idea Submission form that can be used by employees and management along with an Excel evaluation template to assess the viability of proposed project. We validated our process by obtaining a set of real market data on the home improvement media market from CBMC. Overall, our process streamlined and standardize the idea generation and evaluation process, while ensuring resources are spent on ideas that meet CBMC's success targets.

ACKNOWLEDGEMENTS

We would like to recognize the dedicated work by Lindsay Boyd, Amruta Kulkarni, Sean Pierce, Taekyoo Shin, and Brett Thomason who participated in the data gathering, analysis and earlier drafts of the research for this case under the supervision of the authors.

REFERENCES

Excel electronic file will be available upon request

DECISION SCIENCES INSTITUTE
Paraconsistent Logic in Company Evaluation

Adriana Kroenke
Universidade Regional de Blumenau
E-mail: akroenke@furb.br

Jorge Eduardo Scarpin
Concordia College
E-mail: jscarpin@cord.edu

Luana Sara Bizatto
Universidade Regional de Blumenau
E-mail: luanasarabizatto@hotmail.com

Nelson Hein
Universidade Regional de Blumenau
E-mail: hein@furb.br

ABSTRACT

The aim of this research is to evaluate the economic and financial viability of companies in the steel and metallurgy sector listed on the Brazilian stock exchange B3 in the period 2011 to 2018, using the perception of research accountants synthesized and measured using paraconsistent logic.

KEYWORDS: Paraconsistent Logic; Economic-financial indicators; Steel and Metallurgy.

INTRODUCTION

According to Falconi (2017), management tools in steelmaking need to be improved. One of the challenges that emerge in this subsector is the implementation of large investment plans, in which Falconi (2017) cites as some of the solutions: increase in Ebitda, productivity gains through standardization of operational processes, improvement of product quality, reduction of budget deviations and optimization of project development and implementation (deadlines, cost and quality).

However, the evaluation of a company's economic-financial information may suggest its potential to receive future investments (Aguiar, Corrar & Batistella, 2004). However, the success or failure of a particular project may differ between stakeholders, i.e. their views on performance may involve subjectivity. It is not always the decision-making about performance that consists of prioritizing the maximization of company value, due to personal impressions, feelings and opinions, that is, choices depend on subjectivity (Fernandes & Mazzioni, 2015).

Aiming at reliable scores of company performance, the integration of indicators is essential (Gasparetto, 2004). This synthesis can be done in several ways. However, the perception and interpretation of the accounting information by expert accountants depend on their training, vision and their theoretical-practical influences in which the use of Paraconsistent Logic allows to perform a form of integrated analysis between the indicators. The isolated interpretation of the quotients also has its importance (Souza, 2007). However, the generalization of a block of information in a single vector is far superior.

According to Bortoluzzi (2009), the use of only one indicator is not enough to explain performance, but the use of many indicators tends to confuse users. Therefore, it is vital that each sector of activity analyzes what the most relevant indicators to analyze the company's position (Soares, 2006; Fuenmayor & Angulo, 2008) are. Thus, one has as a research question: what is the degree of economic-financial viability attributed by research accountants of companies in the sector of Metallurgy and Siderurgy, synthesized and measured by the paraconsistent logic annotated?

The objective of the research is to evaluate the degree of the feasibility of companies in the Metallurgy and Metallurgy sector analyzed by accountants researchers from three university institutions of the South of Brazil, observing their economic-financial indicators.

The research is justified primarily due to the small number of investigations that use the economic-financial indicators in the evaluation of companies in the sector of Iron and Metallurgy. Also, it should be noted that there is a global crisis of overcapacity in the steel industry due to the growth of Chinese steel production, which has hit developed countries (The Economist, 2016), and a reduction in production in recent decades.

In addition to the market difficulties encountered in the steel and metallurgy sector, limitations are still perceivable in Brazil, such as: government economic policy based on a valued currency, lack of a specific industrial policy for the subsector, possibly tending to hamper performance improvement, reinforcing this research in the Metallurgy and Steel Industry (Poso, 2015).

THEORETICAL REFERENCE

Economic-financial indicators

Historically, the use of indicators is essential to companies and also to the scientific context (Antunes & Martins, 2007), which facilitate the understanding of complex phenomena as they are tools for synthesizing reality (Campos & Melo, 2008).

Primarily economic-financial indicators are used to measure aspects of companies through quantitative relationships, to indicate the economic and financial situation (Matos et al., 2003; Fuenmayor & Angulo, 2008; Rodríguez & Gutiérrez, 2011), which are related to activities, internal controls, organizations, and business management (Fuenmayor & Angulo, 2008, García Molina & Artola, 2010, Rodríguez & Gutiérrez, 2011).

The disclosure of economic-financial indicators occurs through the relationship between accounts or groups of financial statement accounts, where the indicators are generally divided into five categories: liquidity, activity, indebtedness, profitability (profitability) and stock analysis (García, Tabuenca, Martí, & Romero, 2007, Azeredo, Souza & Gomes Machado, 2009, Bortoluzzi, 2009, Gitman, 2010, Rodríguez & Gutiérrez, 2011).

Generally, debt, activity, and liquidity indicators are associated with data that may aid in knowledge about a company's operations (Ehrhardt & Brigham, 2012). According to Gitman (2010), indicators of indebtedness, activity, and liquidity contribute to risk measurement, while profitability indicators consist of measurement of return, while indicators of stock analysis can capture risk and return.

Despite the importance of economic-financial indicators, it is estimated that until the mid-1970s, business management used to rely only on one indicator of its decisions. However, since 1913, criticisms have emerged regarding the exclusive use of liquidity to understand the financial situation of the company. Alexander Wall was one of those responsible for the questioning of the restricted use of liquidity in business analysis. In this way, the use of indicators in isolation is not very useful to the company at present due to the changes in the productive systems, which are more complicated, but also a single financial indicator is not able to create value (Antunes & Martins, 2007).

In order to analyze the performance of companies, restricting themselves to liquidity measures alone may be inadequate, because according to Gitman (2010), current assets and current liabilities may not be able to make a real description of the company. However, Molina Y Artola (2010) argue that the comparison between indicators of liquidity, activity, indebtedness, and profitability are essential for analysis purposes, since according to the authors, the level of indebtedness is affected by the profitability of the company, as well as current liquidity is weakened by the rotation of inventories.

According to the specific objective of each indicator, Rodríguez, and Gutiérrez (2011) report that it is up to the company to select the indicators that best evidence its performance and its positions in the business, when considering its sectors of operations and operations.

Economic-financial positioning

Among the benefits in the literature regarding the company's position in the business, Prahalad & Hamel (1990) emphasize that a prominent position contributes to companies guaranteeing a superior economic performance. Other authors such as Notteboom & Winkelmans (2001) and Frémont (2009) argue that a favorable positioning in the business tends to provide higher profitability to the companies.

Winter (2002) argues that the industry environment in which firms operate is an influence on their economic performance. In order for a company to differentiate itself from the others in its industry, like Crook, Bratton, Street & Ketchen (2006) and Teece et al. (1997), it should possibly have a well-planned strategy. According to these arguments, Porter (1996) clarifies that the achievement of strategic positioning is related to the choice of activities that are different from the rivals so that there is a resistance to competition by the company.

Reed & DeFillippi (1990) argue that for firms to have an advantage over others, firms should develop internal competence. Specifically, Hofer and Schendel (1978) argue that competitive advantage can be achieved by deploying resources in these companies. Winter (2002) points out that as there is increased competition, increased turbulence, and reduced profits, among alternatives, managers may seek to compete with resources. According to the author, the use of internal resources can be used even to achieve impossible results (Winter, 2002).

On the other hand, Baron and Kenny (1986) argue that the accumulation of capabilities without a strategy can result in a resource extravaganza, that is, the resource effects on the performance of the business can be changed, which makes it necessary to mediate the resources that are in the organization. In this way, the strategic deployment of resources with the addition of new capabilities to the existing ones can provide the competitive advantage (Wernerfelt, 1984; Barney, 1986, 1991, Grant, 1991, Peteraf, 1993, Acar & Zehir, 2010).

Especially Tampoe & Macmillan (2000) argue that the best business strategies are the ability of the company to meet the needs of its customers. Storage, which integrates any logistics system, plays a crucial role in promoting a customer service at low cost (Stock & Lambert, 2001). Thus, according to Chopra & Meindl (2003), inventories can contribute to the adequacy and balance between supply needs and customer demand. However, Bowersox & Closs (2001) argue that the reduction of some resources committed in inventories may also be viable since it leads to cost reduction and increase profitability and profitability of a company.

Siderurgy and Metallurgy

Given the importance of the economic-financial positioning, for the steel and metallurgy subsector technology is increasingly advanced, as this segment requires highly efficient logistics, depending on the efficiency of the transport operations in the whole chain from the ore deposit to the end user of steel (Siderurgia Brasil, 2011, Libertá, 2015). According to Siderurgia Brasil (2011), material handling logistics has always been a fundamental factor in guaranteeing good results at Siderurgia. Specifically, transport in this sub-sector needs to be executed with maximum efficiency and at the lowest possible cost, which consequently leads to a greater need for modern equipment as well as more advanced management systems (Siderurgia Brasil, 2011).

Siderurgy and Metallurgy are subdivided into three segments: iron and steel artifacts, copper and steel artifacts, in which the last segment (Siderurgia) presents five companies, and iron and steel artifacts total four companies, and copper artifacts with a company (BM&FBovespa, 2016). Metallurgy is related to the processes of extraction, casting, manufacturing, and treatment of metals and alloys, such as: copper, aluminum, nickel, tin, zinc, and steel (MME, 2011).

The metallurgical sector has considerable importance for the Brazilian economic scenario since it has its relevance to other activities in the country such as civil construction, capital goods and the automobile industry (MME, 2011, 2013, 2016). As this sector is present in the value chain of several organizations, this segment can serve as a buffer against the impacts of the crisis, in order to reduce the adverse effects in large companies, and thus assuming importance to economic stability (Milk, 2014).

On the steel sector, this is known as the steel metallurgy sector (MME, 2011), or it can be said that it is the branch of metallurgy which is focused on the treatment of steel and cast iron (Silva, 2014). Among the steelmakers, CNM (2012) reports that the companies Gerdau, Usiminas, and Sid Nacional are ranked among the best positions in the world ranking.

In general, the steel industry is recognized as complicated, because it is necessary for the intensive use of energy, enormous capital contribution, and still the maturation of the investment is usually slow. However, in Brazil, the production of products and by-products is still simple, which in turn generates less value, even though there is an increase in the share of nobler products in national steel production (CNM, 2012).

According to information from the Brazilian Steel Industry (2015), the Brazilian steel industry's main objective is to participate in the world steel trade, aiming at the sustainable development of the country. Recently, China is the largest producer and consumer in the steel sector (World Steel Association, 2013), but it should be considered that Brazil currently has the largest steel industrial park in South America, and is also considered the largest producer in America Latina. Especially in 2014, Brazil ranked sixth as a net exporter of steel, and as the ninth largest producer of steel in the world, and in 2015 reached the eighth position in the world ranking as an exporter in the metallurgical sector (MME, 2016).

Paraconsistent logic

Western logical thinking was formed by the laws of Classical Logic, that is, by the law of non-contradiction. This classic logic has been created and sustained by rigid laws and is based on binary reasoning, in which its statements may be true or false, but cannot be partially true and partially false (Carvalho & Abe, 2011).

Often difficulties are encountered in establishing limits to be made affirmative or negative about the quality of things. The boundaries between "False" and "True" are uncertain, ambiguous, indefinite and contradictory, for many personal statements cannot only be classified as "true" or "false," or "yes" or "no." Thus, Non-Classical Logic seeks to give more satisfactory answers to problems related to situations in which the laws that are required by Classical Logic do not fit (Carvalho & Abe, 2011).

Paraconsistent Logic (LP), as Non-Classical Logic, is based on inconsistent and non-trivial theories, that is, it can manipulate systems that are both inconsistent and nontrivial (without the risk of trivialization). According to Carvalho & Abe (2011), Paraconsistent Logic invalidates the principle of Non-Contradiction and admits the treatment of different signals.

The Polish J. Lukasiewick and the Russian philosopher N. A. Vasilév are the forerunners of the Paraconsistent Logic around 1910. The publication of works of these authors consisted in the possibility of a logic that could consider the contradictions; their studies were independent of each other despite being performed at the same time. However, none of them had a broad view of non-classical logic, unlike today, because they were still restricted to traditional Aristotelian logic (Carvalho & Abe, 2011).

In 1948, Jäskowski formalized a paraconsistent propositional calculus which is termed as a discursive propositional calculus. At the same time (1954), however, the independent logician Newton Carneiro Affonso da Costa developed several paraconsistent systems (Carvalho & Abe, 2011).

There is a difference between the works of Jáskowski and Da Costa, while Jáskowski has elaborated a paraconsistent propositional calculation called Discursive (or Discursive) Propositional Calculus, Da Costa has deepened several paraconsistent logics that cover all the standard logical levels (Carvalho & Abe, 2011). According to the development of Paraconsistent Logic with Da Costa, studies related to this Logic were carried out. Among some of the researches that applied the Paraconsistent Logic, Torres (2004) developed an intelligent system for the control of autonomous mobile robots, and Falquete (2004) used in case-based reasoning systems for the treatment of inconsistencies. It is also possible to cite the work of Carvalho, Brunstein, & Abe (2004) in the decision making, for the evaluation of the project of implantation of a factory, through the use of Paraconsistent Logic.

Technological evolution has allowed a better application of the Paraconsistent Logic because instead of ignoring specific problems of inconsistencies in refuting such as false or confirming them as true, it allows the modeling of human behavior. Therefore, Paraconsistent Logic differs from quantitative reasoning methodologies by providing interpretations for different information (Carvalho, 2002).

The Peruvian philosopher Francisco Miró Quesada created the term Paraconsistent Logic in which he was first employed in a conference held at the University of Campinas in 1976 at the Third Latin American Symposium on Mathematical Logic (Carvalho, 2002).

According to Carvalho (2002), the calculation or logic underlies the theory T, that is,

$$T = \{A : T \vdash_c A\}$$

A represents the syntactic consequence of T, and in this condition A must belong to T, or it is also said that A is the "theorem" of T. According to Carvalho & Abe (2011), for a given theory T to be paraconsistent it needs to attend to two conditions: to be inconsistent and non-trivial.

A theory T is inconsistent when it has at least one "theorem" A of a set F of all sentences (formulas) whose negation ¬A is also "theorem" of T. For example, if a given sentence p "John is mortal "is considered in a theory T, it will also be considered its negation, that is, the -p meaning "John is not mortal. It makes theory T inconsistent. Otherwise, the theory is said to be consistent (Carvalho & Abe, 2011).

Some examples of situations of inconsistencies can be mentioned. According to Silva Filho (1999), in a condominium meeting, for the decision on a reform in the building, different opinions must be considered, since opinions are not always unanimous. If decisions were always unanimous, the trustee's decision would be facilitated, but some decision-makers wish to reform, and in contrast to others, not generating contradictions.

Another example cited by the authors is an administrator who is the head of a team that has the decision to promote one of its employees. To this end, various information is collected which come from various sources such as direct management, co-workers, personnel department, and others. The information coming from various sources can be simultaneously totally favorable and contrary. Usually, there are two paths to this situation, or the administrator considers that the information is insufficient, or that it is contradictory, and it is necessary to search for new information in his decision making about the employee.

According to Bishop & Cazarini (2006), specifically in the area of medicine, when there is a medical committee to decide which treatment is more feasible for a particular patient, there are often differences between opinions in which doctors end up proposing different treatments. In these particular cases, paraconsistency work can contribute favorably to making better decisions.

The presence of inconsistency can also be seen in the area of law during the trial of the defendant, according to Bishop & Cazarini (2006), in which there is evidence of contradictory evidence from lawyers, both defense and prosecution. About the decision that the jury must take on blame or innocence of the defendant, there is the presence of the paraconsistent, which, when considered, may facilitate the finding of the most appropriate final result.

Paraconsistency should also be considered as frequently present in companies. The authors, Bispo & Cazarini (2006), also cite especially the situation when a research is carried out in relation to a specific type of innovation, in which it is recurrent that such innovative product or technology can simultaneously present good receptivity by potential clients and also considerable aversion by other customers which are also considered potential. In this research, inconsistency can occur at the moment when the specialists answer the questionnaires, in which it can be visualized if the economic-financial position of the company is allocated in a region of inconsistency.

METHODS

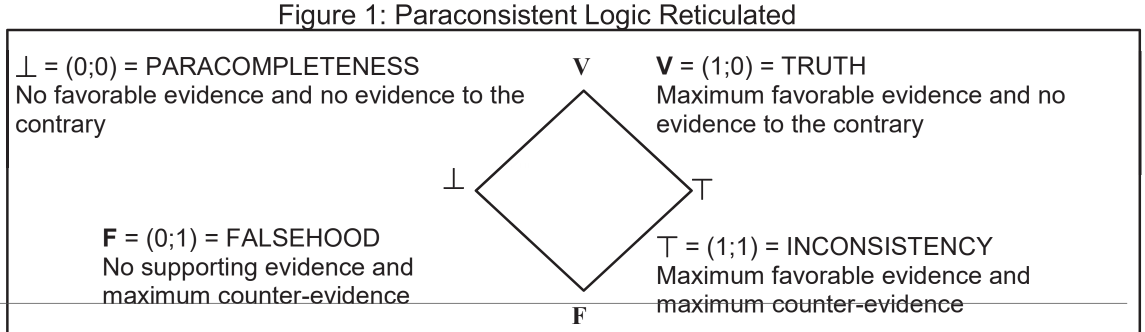
This study evaluated nine steel and metallurgy companies listed on BM&FBovespa using six specialists: two researchers and two finance assistants from three Brazilian universities.

In order to verify the influence of the indicators on the economic and financial positioning of the companies of Siderurgia and Metallurgy, the period from 2011 to 2018 was delimited for the study. For this purpose, the paraconsistent logic was used to determine the accounting position of the companies under analysis within the grid of annotations.

Technically, the Paraconsistent Logic consists of establishing propositions and then parameterizing them so that the factors of more significant influence in the decisions can be isolated, and, with the collaboration of specialists, notes are made for these factors. The objective of these procedures is to consider the concepts of uncertainty, inconsistency, and paracomplete so that the nuances of reality are better reflected and understood (Carvalho & Abe, 2011).

The annotations consist of the attribution of degrees of belief (μ_1) and degrees of disbelief (μ_2) for each elementary proposition p; these values range from 0 to 1 and are independent. According to Carvalho and Abe (2011), these notes are related to the concepts used in the LPA, in the case: inconsistency, paracomplete (indetermination), notions of truth and notions of falsity. The Hasse diagram represents the lattice in which the reticulate of the annotations can be called, where the pair (0; 0) represented by $\perp$ is the minimum of the lattice, and the maximum consists of the pair (1; 1) which is represented by $\top$. (Carvalho, 2002). Figure 1 illustrates this representation:

Figure 1: Paraconsistent Logic Reticulated

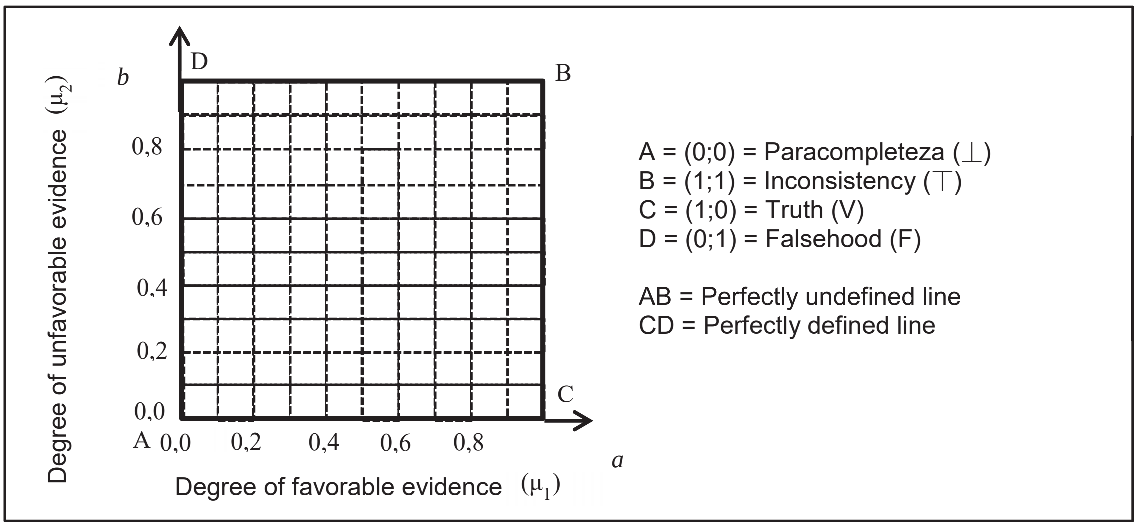


Thus, the pair (1; 0), which is associated with p, represents total belief and at the same time no disbelief in p; (Carvalho, 2002), or according to Carvalho & Abe (2011), this situation represents maximum favorable evidence and no contrary evidence. Even the pair (0,1) reveals the opposite: no belief and total disbelief in p, constituting a logical state of falsity (representation of no favorable evidence and together presents maximum contrary evidence).

The pair (1,1) means total belief and disbelief in p, translating a logical state called inconsistency (maximum favorable evidence and maximum contrary evidence). Finally, we have the pair (0,0), which indicates the total absence of belief and disbelief in p, that is, it consists of a logical state of indetermination or also called a state of paracompleteness (no favorable evidence and no evidence contrary) (Carvalho, 2002).

According to Carvalho (2002) and Carvalho & Abe (2011), the study of LPA can also occur in a Unitary Square in the Cartesian Plane (QUPC), because this QUPC is a way to represent the grid of annotations containing records of Degrees of Evidence Favorable (Belief) μ_1 moreover, Degrees of Unfavorable Evidence (Disbelief) μ_2. Figure 2 represents the Cartesian plane unit square (QUPC):

Figure 2: Unit chart of notes



Logic states in the reticulate associated with the four points are also called final states in which they belong to the sides AC, AD, CB, and DB of the QUPC. Through these sides, interpretations can be performed; for example, on the AC side, we have that the degree of belief varies between 0 and 1. However, the degree of disbelief is null (μ_2 = 0).

As a given point starts from A and advances to C, it passes from a situation of total indetermination (absence of all information), and in this way, one reaches a situation considered perfectly defined, that is, of total belief assuming a state of truth. Conversely, the AD side of the square has a zero belief degree (μ_1 = 0), and this time the degree of disbelief(μ_2) it varies from 0 (minimum) starting from A to C which represents 1 (maximum), and it is a perfectly defined situation (i.e., total disbelief or even the logical value is called falsehood) (Carvalho, 2002).

For the CB side, the degree of belief is equal to the maximum (μ_1 = 0) and is kept constant. However, there may be variation in the degree of disbelief (μ_2), which starts at minimum (zero) in C, up to the maximum value (1) in B. Thus it passes from a logical state of truth (situation perfectly defined) in C, to another logical value of inconsistency in which one has a belief and total disbelief at the same time (Carvalho, 2002). Finally, the DB side has the degree of disbelief kept to the maximum (μ_1 = 1), and the degree of belief varies from the minimum (zero) to the maximum (1). Therefore, there is a passage from a falsehood situation in D to another situation of inconsistency in B (Carvalho, 2002).

It should be noted that in the QUPC, there will never be a total belief or total disbelief, that is, a situation of absolute truth or absolute falsehood will never be seen, nor is there a situation of total inconsistency or complete misinformation. Thus, for the internal points we have 0< μ_1 < 1 e 0 < μ_2 < 1 (Carvalho, 2002).

The Variables

Because some companies do not present share information for each year and also negative equity, we opted for the exclusion of some economic-financial indicators. The economic-financial indicators are presented in Table 1.

Table 1: Construction of economic and financial		
VARIABLE	FORMULA	AUTHORS
LG	$\frac{\text{Current Assets + Long-Term Assets}}{\text{Current Liabilities + Long-Term Liabilities}}$	Souza (2007); Matarazzo (2010)
CR	$\frac{\text{Current Assets}}{\text{Current Liabilities}}$	Souza (2007); Matarazzo (2010); Gitman, (2010), Ehrhardt & Brigham (2012)
ATR	$\frac{\text{Current Assets - Inventories}}{\text{Current Liabilities}}$	Souza (2007); Matarazzo (2010); Gitman (2010); Ehrhardt & Brigham (2012)
PC/CT	$\frac{\text{Current Liabilities}}{\text{Total Liabilities}}$	Matarazzo (2010); Silva (2012)
ITR	$\frac{\text{Inventories}}{\text{Cost of Goods Sold}}$	Souza (2007); Silva (2012)
RTR	$\frac{\text{Accounts Receivable}}{Net\ Sales}$	Souza (2007); Gitman (2010); Ehrhardt & Brigham (2012)
AT	$\frac{\text{Net Sales}}{\text{Average Total Assets}}$	Matarazzo (2010); Silva (2012)
ROA	$\frac{\text{Net Income}}{\text{Average Total Assets}}$	Souza (2007); Gitman (2010); Ehrhardt & Brigham (2012)
GPR	$\frac{\text{Gross Profit}}{\text{Net Sales}}$	Souza (2007); Gitman (2010); Ehrhardt & Brigham (2012)
EPS	$\frac{\text{Net Income}}{\text{Number of Shares Outstanding}}$	Souza (2007); Gitman (2010)

When we use tertiles, the standard indicators for the framing of each of the companies were determined in sections, in which the first, second and third tertiles were considered for the sections of the paraconsistent logic, respectively. The three sections used in the paraconsistent logic were: "S" _ "1", which translates the conditions in which the indicator is unfavorable to the researcher and/or auxiliaries; "S" _ "2", which translates the conditions in which the indicator is indifferent to the researcher and/or auxiliaries; and finally, "S" _ "3", which translates the conditions in which the indicator is favorable to the researcher and/or auxiliaries (Carvalho & Abe, 2011). Figure 1 depicts the paraconsistent logic tertiles, which were also represented in the questionnaires.

Figure 3: Representation of tertiles used as sections of Paraconsistent Logic for questionnaires

Panel A: Standard indicators for indicators that the higher, the better (general liquidity, current liquidity, dry liquidity, immediate liquidity, asset turnover, asset profitability, gross margin, earnings per share, dividends per share and price/earnings)

	Low	Medium	High	
Maximum	1º Tercil **(S_3)**	2º Tercil **(S_2)**	3º Tercil **(S_1)**	Minimum

Panel B: Standard indicators for indicators that the lower, the better (the composition of indebtedness, the average term of inventory renewal and an average term of receipt of sales)

	Low	Medium	High	
Minimum	1º Tercil **(S_1)**	2º Tercil **(S_2)**	3º Tercil **(S_3)**	Maximum

(S_1) Translates conditions where the factor is **unfavorable**

(***a***) Degree of Favorable Evidence	(***b***) Degree of Unfavorable Evidence

(**S_2**) Translates conditions into the factor is **indifferent**

(***a***) Degree of Favorable Evidence	(***b***) Degree of Unfavorable Evidence

(S_3) Condition translates into factor is **favorable**

(***a***) Degree of Favorable Evidence	(***b***) Degree of Unfavorable Evidence

For the determination of the standard value of the distribution, statistical measures such as arithmetic average, fashion, median, quartiles, decile, or even percentile are usually used (Silva, 2012). The choice specifically for tertiles is due to the necessity of the establishment of sections of the Paraconsistent Annotated Logic.

The business environment in the sections for each indicator performed by tertiles depends on whether the indicators are of the higher, the better, or the lower, the better. According to this logic, indicators belonging to Panel A located in the 1st tertile are characterized as unfavorable, and those that are allocated in the third tertile are considered favorable. About Panel B indicators that are in the 1st tertile, they are in favor of the companies, while the companies in the third tercile are unfavorable. The companies that are part of the 2 rd tercile, both in Panel A and in Panel B present such indicators as reasonable.

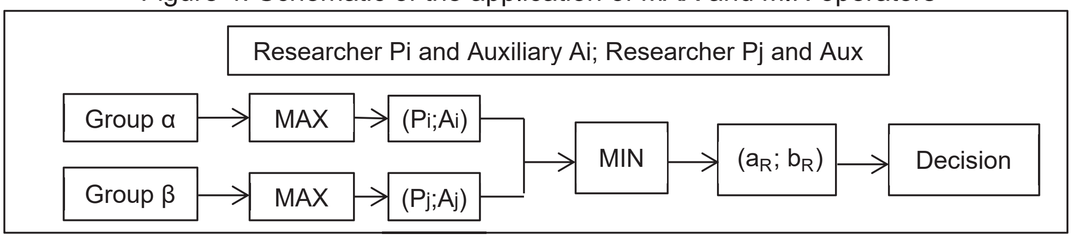
The business environment was made using three (3) teams of specialists. Each team was composed of two (2) professors/researchers belonging to the Postgraduate Programs in Accounting (PPGCCs) of the following higher education institutions (IES): Federal University of Paraná (UFPR), Federal University of Santa Catarina (UFSC)) and Regional Universities of Blumenau (FURB). Each researcher selected an assistant from among his masters and doctoral advisors. Thus, each of the 3 HEIs involved participated with a team of 4 members. It can symbolically establish the set {P1, P2, A1, A2}, where P1 and P2 are the professors/researchers and A1, and A2 are auxiliary.

The following set of ordered pairs performed the evaluations of each institution involved: {(P1, A1), (P1, A2), (P2, A2), (P2, A2)}. This type of composition intended to seek some underlying pattern in the evaluations.

The evaluators depended on the indication of degrees of indication of the factors in their three possibilities: favorable, indifferent and unfavorable (S1, S2, S3), associating in each of them a degree of favorable belief and another one of unfavorable belief, with values between 0 zero) and 1 (one).

The quartile of the pertinence of each company in each variable was previously determined, remaining the framework of the sections with the calculated indicators and measurement of the grades by the respondents. Then the MAX and MIN operators of the Annotated Paraconsistent Logic were applied, according to Figure 4, where Group α consists of Researcher Pi and auxiliary Ai and Group β adds the Researcher Pj and auxiliary Aj.

Figure 4: Schematic of the application of MAX and MIN operators



The maximization operator is applied from the degree of certainty within the set of annotations for each group separately, that is, it chooses both the highest degree of favorable evidence and the lowest degree of unfavorable evidence. Thus, the application of MÁX is made among the opinions of experts who are not all determinants, that is, it suffices the favorable opinion of only one expert so that the result is considered satisfactory (Carvalho & Abe, 2011). In contrast, the MIN operator consists of minimizing the degree of certainty within a set of annotations, with the choice of the least degree of favorable evidence and a higher degree of unfavorable evidence.

This minimization operator is applied among the maximum obtained from Groups A and B. Finally, the barycentre of each company was calculated, and this graph was calculated along with the other companies under study.

In particular, a proper evaluation, without the bias of inconsistency or for completeness, has an excellent adherence to the straight line segment that links the ordered pairs (0,1) to (1,0), that is, the values evaluated and framed in the lattice thus having almost no inconsistency (1,1) or paracompleteness (0,0). To generate the degree of viability (GV) of the companies from their position in the lattice, given by the ordered pair (x, y), will be projected on the line $x + y - 1 = 0$. To do so, we will initially calculate the distance from the point (x, y) to the line by equation $d(P,r) = \frac{|ax_0 + by_0 - c|}{\sqrt{a^2 + b^2}}$, which in this case will be given by $d(P,r) = \frac{|x-y-1|}{\sqrt{2}}$. To obtain the location of the projected point, we will calculate the Euclidean distance between (x, y) and the point (0,1) that will be called d +, since it will present the gross distance of the point (x, y) viability in V = (1.0). Similarly, the gross distance between (x, y) and the non-viability scenario (0,1) will be calculated. Using the Pythagorean Theorem, we will calculate the distance d and d, which will be the distances of the point projected on the line, to the point (0,1) and (1,0), that is, the feasibility and inviability. In mathematical terms, it will be: $d^{via} = \sqrt{(d^+)^2 + (d(P,r))^2}$ e $d^{inv} = \sqrt{(d^-)^2 + (d(P,r))^2}$. Finally, the degree of feasibility to be calculated: $GV_i = \frac{d_i^{inv}}{d_i^{via} + d_i^{inv}}$. Of course, each pair of evaluators will have a new scenario that will be synthesized by the average.

RESULTS ANALYSIS

The lattices for each year of the period from 2011 to 2018, obtained through the paraconsistent logic, are presented below. For esthetic reasons, only the results referring to the year 2018 will be presented, participating that the other years follow the same construction. Tables 01, 02 and 03 bring the results of the year 2015, from the framework made by the specialists of UFSC.

In Table 02 are the combinations made by the two groups of FURB expert reviewers, i.e., professors (P) and students (A) where P = {A1, A2} and A = {B1, B2}.

Table 2: Paraconsistent evaluation; year 2018; FURB specialists					
FURB - 2018 (A1 e B1)			FURB - 2018 (A1 e B2)		
Ferbasa	0.695	0.338	Ferbasa	0.588	0.442
Fibam	0.502	0.534	Fibam	0.412	0.635
Gerdau	0.471	0.501	Gerdau	0.484	0.495
Gerdau Met	0.417	0.558	Gerdau Met	0.415	0.561
Mangels Indl	0.506	0.522	Mangels Indl	0.468	0.555
Panatlantica	0.598	0.448	Panatlantica	0.619	0.410
Paranapanema	0.503	0.502	Paranapanema	0.419	0.588
Sid Nacional	0.582	0.408	Sid Nacional	0.582	0.435
Tekno	0.454	0.550	Tekno	0.408	0.603
Usiminas	0.399	0.604	Usiminas	0.375	0.658
FURB - 2018 (A2 e B1)			FURB - 2018 (A2 e B2)		
Ferbasa	0.668	0.355	Ferbasa	0.596	0.408
Fibam	0.448	0.568	Fibam	0.388	0.635
Gerdau	0.528	0.466	Gerdau	0.522	0.482
Gerdau Met	0.415	0.562	Gerdau Met	0.411	0.578
Mangels Indl	0.545	0.492	Mangels Indl	0.515	0.497
Panatlantica	0.642	0.395	Panatlantica	0.623	0.404
Paranapanema	0.448	0.581	Paranapanema	0.408	0.592
Sid Nacional	0.558	0.428	Sid Nacional	0.573	0.412
Tekno	0.441	0.578	Tekno	0.403	0.597
Usiminas	0.363	0.646	Usiminas	0.362	0.665

Following are the combinations made by the two groups of expert evaluators of UFPR, i.e., professors (P) and students (A) where P = {E1, E2} and A = {F1, F2}.

Table 3: Paraconsistent evaluation; year 2015; specialists from					
UFPR - 2018 (C1 e D1)			UFPR - 2018 (C1 e D2)		
Ferbasa	0.604	0.227	Ferbasa	0.760	0.265
Fibam	0.562	0.412	Fibam	0.527	0.415
Gerdau	0.500	0.458	Gerdau	0.515	0.431
Gerdau Met	0.515	0.438	Gerdau Met	0.492	0.485
Mangels Indl	0.585	0.385	Mangels Indl	0.627	0.404
Panatlantica	0.619	0.304	Panatlantica	0.650	0.296
Paranapanema	0.573	0.354	Paranapanema	0.525	0.400
Sid Nacional	0.623	0.219	Sid Nacional	0.619	0.265

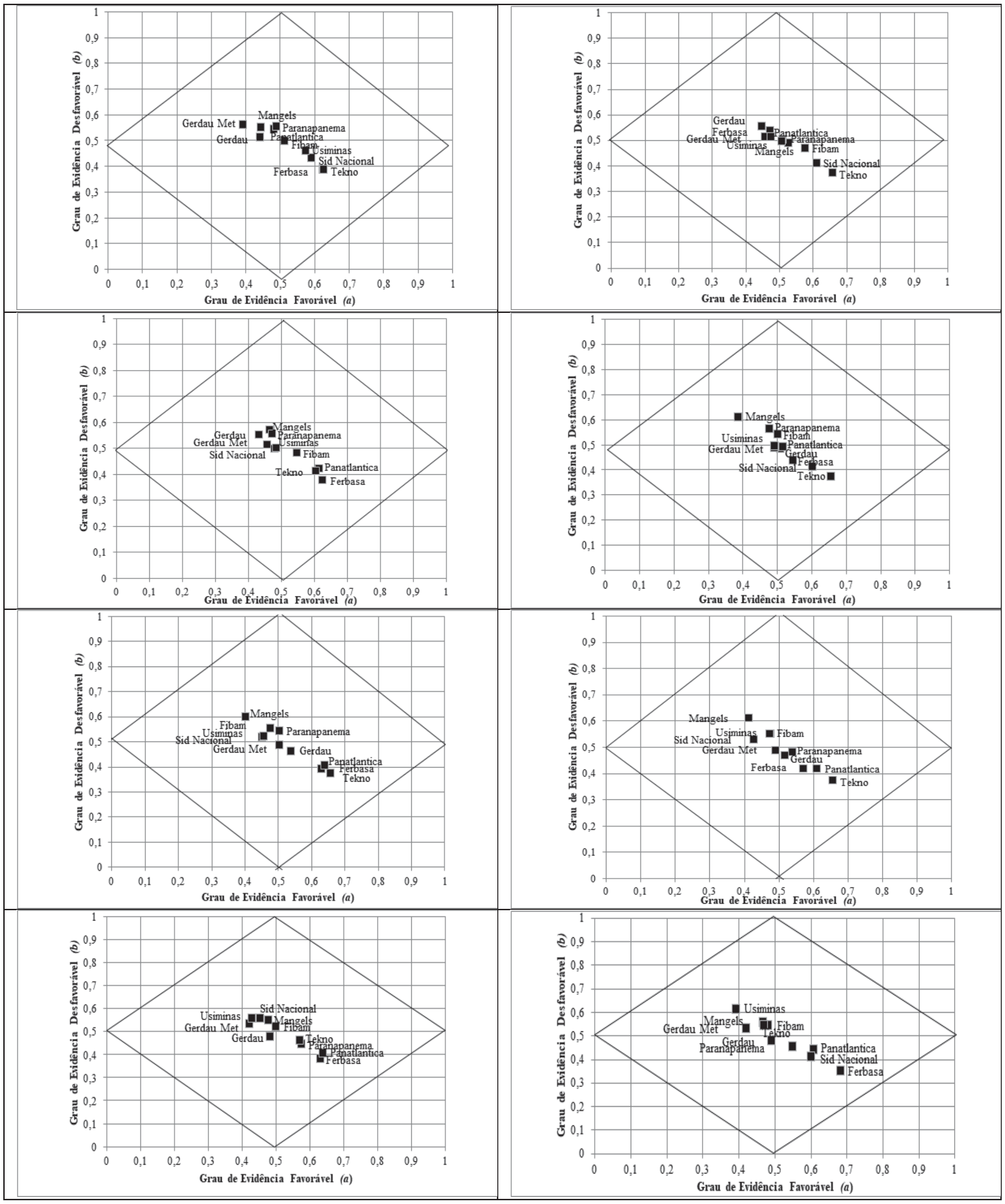
Tekno	0.581	0.273	Tekno	0.612	0.315
Usiminas	0.535	0.415	Usiminas	0.550	0.415
UFPR - 2018 (C2 e D1)			UFPR - 2018 (C2 e D2)		
Ferbasa	0.538	0.385	Ferbasa	0.635	0.381
Fibam	0.588	0.396	Fibam	0.555	0.555
Gerdau	0.535	0.396	Gerdau	0.558	0.369
Gerdau Met	0.519	0.385	Gerdau Met	0.492	0.432
Mangels Indl	0.638	0.323	Mangels Indl	0.665	0.331
Panatlantica	0.627	0.312	Panatlantica	0.685	0.305
Paranapanema	0.581	0.369	Paranapanema	0.535	0.381
Sid Nacional	0.619	0.300	Sid Nacional	0.708	0.296
Tekno	0.538	0.358	Tekno	0.558	0.358
Usiminas	0.550	0.381	Usiminas	0.570	0.382

In Table 04 are the combinations made by the two groups of UFSC expert reviewers, i.e., professors (P) and students (A) where P = {E1, E2}, and A = {F1, F2}.

Table 4: Paraconsistent evaluation; year 2018; specialists from					
UFSC - 2018 (E1 e F1)			UFSC - 2018 (E1 e F2)		
Ferbasa	0.493	0.401	Ferbasa	0.492	0.419
Fibam	0.276	0.599	Fibam	0.348	0.604
Gerdau	0.484	0.434	Gerdau	0.440	0.431
Gerdau Met	0.336	0.523	Gerdau Met	0.278	0.565
Mangels Indl	0.410	0.417	Mangels Indl	0.484	0.438
Panatlantica	0.579	0.338	Panatlantica	0.548	0.376
Paranapanema	0.282	0.600	Paranapanema	0.264	0.657
Sid Nacional	0.482	0.355	Sid Nacional	0.581	0.376
Tekno	0.407	0.494	Tekno	0.386	0.517
Usiminas	0.235	0.645	Usiminas	0.218	0.710
UFSC - 2018 (E2 e F1)			UFSC - 2018 (E2 e F2)		
Ferbasa	0.452	0.429	Ferbasa	0.439	0.444
Fibam	0.247	0.631	Fibam	0.304	0.608
Gerdau	0.456	0.393	Gerdau	0.382	0.439
Gerdau Met	0.345	0.472	Gerdau Met	0.260	0.570
Mangels Indl	0.374	0.443	Mangels Indl	0.406	0.466
Panatlantica	0.554	0.339	Panatlantica	0.518	0.377
Paranapanema	0.269	0.588	Paranapanema	0.236	0.675
Sid Nacional	0.399	0.398	Sid Nacional	0.445	0.418
Tekno	0.349	0.524	Tekno	0.321	0.548
Usiminas	0.196	0.655	Usiminas	0.172	0.706

After 12 analyzes for each year, the position of each company was reduced to only one ordered pair (x, y), bringing the favorable degree of evidence (x) and its unfavorable degree of evidence (y). The values were synthesized using the mean of the ordinates and abscissa. The final result for the period 2011-2018 is in Figure 5.

Figure 5: Corporate positioning of the companies in the lattices - period 2011/18



In Figure 5, the accounting position of companies between the years 2011 to 2018 and should be understood by the matched charts: 2011/12, 2013/14, 2015/16, and 2017/18. In the first graph, the companies Tekno and Ferbasa presented the best positions. This result is consistent with the conclusions obtained in previous years, which were verified in the study of Kroenke (2009), for the period from 2004 to 2008, that the two companies also presented the first positions. The company Usiminas is in the third position, that is, it is ranked among the best. In contrast, the company Mangels is among the last place.

In the second chart, although the Ferbasa company is among the best positions in 2011, it loses much of its position when we compared to other companies in the subsector in 2012. The global crisis interfered with Ferbasa's performance, negatively affecting its costs and reducing the net income. These consequences may justify its change of position about Steel and Metallurgy. It should also be highlighted the loss of position of Usiminas, perhaps not as significant as Ferbasa, but it is noticeable since it had a distance from the viability region when compared to its position in 2011.

About Paranapanema, an improvement in its economic-financial position can be observed. According to the results of the fiscal year 2012, the company has higher gross revenue, sales volume, net income, net margin, and shareholders' equity when compared to 2008. It should be noted that Mangels also improved economic performance in 2012.

The third graph represents the positioning of the companies in the grid for the year 2013, in which Ferbasa resumes its position, as well as the company Panatlantica takes on new positions, being among the best along with Tekno and Ferbasa. On the other hand, the company Sid Nacional, which was ranked among the best in 2011 and 2012, loses an essential position in Siderurgia and Metallurgy. Mangels, which improved its positioning, returned to be among the worst positions, as happened in 2011.

It can be verified that the company Panatlantica, which was in unfavorable positions in the years 2011 and 2012, presents a significant advance in its position for the year of 2013.

In 2011, it can be noted that the company Sid Nacional, has a great improvement in its economic-financial positioning. However, it is noticed that Fibam has a considerable loss in its positioning when compared to the years of 2011 and 2012. Again it is recorded that Mangels is considerably in the last position, approaching more of the region of inviability than the other companies Steel and Metallurgy. The fifth graph shows the reticulation with the positioning of companies in the year 2015.

In 2015, Tekno, Ferbasa, and Panatlantica continued to present the best positions since 2013. However, as in 2013, it is noticeable that Sid Nacional again had an unfavorable position in Siderurgy and Metallurgy for the year 2015. In almost all previous years of the research, Mangels is in the last position. Specifically this year, Mangels was unable to continue its growth level due to the international crisis and also the European crisis, which was identified in 2014. It can also be seen that Fibam is in the last position. We can justify it by the worsening of the European crisis, in which in Fibam caused a decrease in sales, and consequently with a subsequent reduction in production in the year 2015.

In the sixth graph, specifically about Paranapanema, it is noticeable that this one has an improvement in its economic-financial positioning as of 2013. In contrast, except for 2014, the National Sid became among the worst positions of Siderurgia and Metallurgy. Mangels is again in the worst position. In the year 2016, due to its bank debts, Mangels decided to implement an asset sales process. The seventh graph represents the positioning of the companies in the network for the year 2017.

In the period from 2011 to 2016, Tekno was found to be in the top positions. However, in the year 2017, it can be observed that its positioning is reduced, losing four positions. This result of the company can be a consequence of the economic crisis of the country, which took place in the middle of 2017. This year, Tekno presented a considerable drop in supply to the domestic appliance and automotive sectors.

As already verified, Usiminas lost the position after the year of 2011, progressively lost positioning over the years, and in the year 2017 took the last position. This change of position in almost all the period is due to Usiminas being impacted by both the global crisis and the national economic crisis. Similarly, the National Sid also lost position again.

In the eighth and last graph, in general, it is noticeable that the company Mangels is in the last position in almost every year. Possibly, its positioning is due to the financial difficulties that already had in the global crisis, as well as to the steel crisis, which generated significant losses for the company in 2014 and 2015. In 2018, it was evidenced that Tekno has a significant loss in its economic and financial position due to the country's economic crisis. On the other hand, Sid Nacional achieved a favorable position, as it had also presented in the years 2011 and 2012.

Finally, the degree of general viability of the companies analyzed is analyzed. Table IV shows the GV of earned by each HEI in the period studied and its position in the viability ranking.

Table 5: Viability of the analyzed companies and their position in the ranking						
Empresa	**GV-UFSC**	**Posição**	**GV-UFPR**	**Posição**	**GV-FURB**	**Posição**
Ferbasa	0.541901	1ª	0.691138	1ª	0.586226	1ª
Fibam	0.327728	8ª	0.596294	6ª	0.415433	10ª
Gerdau	0.498995	5ª	0.552244	10ª	0.533942	3ª
Gerdau met	0.459769	6ª	0.562208	8ª	0.477668	6ª
Mangels Indl	0.445731	7ª	0.617089	5ª	0.455264	7ª
Panatlantica	0.510043	3ª	0.636549	4ª	0.546707	2ª
Paranapanema	0.311868	9ª	0.590777	7ª	0.434327	8ª
Sid Nacional	0.500048	4ª	0.658402	3ª	0.488654	5ª
Tekno	0.523108	2ª	0.667193	2ª	0.524567	4ª
Usiminas	0.299522	10ª	0.557875	9ª	0.417981	9ª

It should be noted that Kendall's correlation between GV attributed by UFSC and FURB is 90.3% (significant at the 1% level). Between the UFSC and UFPR, it is 59.9% without showing any significance. Similarly, the relationship between UFPR and FURB, the correlation is 57.7%, also without significance.

Between the UFSC and FURB evaluation, Kendall's ordinal correlation coefficient registered 77.8% among the positions in the ranking, being significant at the 1% level. The ordinal correlation between the UFPR and UFSC rankings registered 60.0%, significant at the 5% level.

FINAL CONSIDERATIONS

The objective of this article was to verify the influence of the economic and financial indicators on the economic and financial position of the Metallurgy and Metallurgy companies listed on the BM&FBovespa from 2011 to 2018. Through the application of the Paraconsistent Logic, it was possible to determine the economic, financial institutions. The methodology was characterized as descriptive, with documentary analysis as well as quantitative approach.

As a main result, current liquidity hurt the economic and financial position of Siderurgia and Metallurgy, as opposed to influences on other indicators such as dry liquidity, profitability on assets, asset turnover and average term for renewal of inventories, which had positive influences. In this way, the current liquidity of these companies is possibly weakened by the renewal of inventories, since the non-consideration of inventories through dry liquidity positively influenced the companies' positioning.

Therefore, it should be noted that steel and metallurgy may increasingly require more efficient logistics, which also depends on the efficiency of transportation operations throughout the chain. Thus, according to the importance of the improvement in the subsector, it is emphasized that there may be more significant needs for modern equipment, as well as the implementation of more advanced management systems.

These results reinforce the importance of new strategies for this subsector, to improve both the financial health and the economic and financial position of these companies. Moreover, given the consistency of the Paraconsistent Logic to the determination of the companies' positions, which are not allocated in regions neither for fullness nor for inconsistency, this evidence shows that Paraconsistent Logic may apply to the accounting area since few studies using this methodology in Accounting.

Among suggestions for future research, it is suggested that the same study be replicated in other sectors and different businesses, because according to literature, the consideration of the indicators to the study is different, depending on the sector in which it is being analyzed. Another suggestion is to increase the number of indicators used in the study wherever possible.

REFERÊNCIAS

Acar, A. Z., & Zehir, C. (2010). The harmonized effects of generic strategies and business capabilities on business performance. *Journal of Business Economics and Management*, *11*(4), 689-711.

Aguiar, A. B., Corrar, L. J., & Batistella, F. D. (2004). Adoção de práticas de governança corporativa e o comportamento das ações na Bovespa: evidências empíricas. *Revista de Administra&ccdeil; ão da Universidade de São Paulo*, *39*(4).

Antunes, M. T. P., & Martins, E. (2007). Capital intelectual: seu entendimento e seus impactos no desempenho de grandes empresas brasileiras. *Revista de Administração e Contabilidade Unisinos*, *4(01)*, 05-21.

Assaf Neto, A. (1996). Alexandre & MARTINS, Eliseu. *Administração Financeira. São Paulo, Atlas*, *1*.

Assaf Neto, A. (2003). Finanças corporativas e valor. *São Paulo: Atlas*, *3*.

Azeredo, A. J., Souza, M. A. D., & Gomes Machado, D. (2009). Desempenho econômico-financeiro de indústrias calçadistas brasileiras: uma análise do período de 2000 a 2006. *Revista de Contabilidade e Organizações*, *3*(6), 117-142.

Barney, J. B. (1986). Organizational culture: can it be a source of sustained competitive advantage?. *Academy of management review*, *11*(3), 656-665.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, *17*(1), 99-120.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

BM&FBovespa. Disponível em: <http://www.bmfbovespa.com.br/pt_br/produtos/listados-a-vista-e-derivativos/renda-variavel/empresas-listadas.htm>. Acesso em: 17 mar. 2017.

Bortoluzzi, S. C. (2009). *Avaliação de desempenho econômico-financeiro da empresa Marel Indústria de Móveis SA: a contribuição da metodologia multicritério de apoio à decisão construtivista (MCDA-C).* 2009. Dissertação de Mestrado. Universidade Federal de Santa Catarina, Florianópolis.

Bowesox, D., & Closs, D. (2001). Logística empresarial: o processo de integração da cadeia de suprimentos. *São Paulo, Atlas*.

Campos, L. M. de Souza, Melo, D. A. de (2008). Indicadores de desempenho dos Sistemas de Gestão Ambiental (SGA): uma pesquisa teórica. *Production, 18(3)*, 540-555.

Carvalho, F. R. de, Brunstein, I., & Abe, J. M. (2004). Paraconsistent Annotated Logic in Viability Analysis: an Approach to Product Launching. In D. M. Dubois (Ed.), *AIP Conference Proceedings*. Sacramento, California.

Carvalho, F. R. De; Abe, J. M. (2011). *Tomadas de decisão com ferramentas da lógica paraconsistente anotada: método paraconsistente de decisão*. Blucher: São Paulo.

Chopra, S., & Meindl, P. (2003). Gerenciamento da Cadeia de Suprimentos: Estratégia, Planejamento, e Operação. In *Gerenciamento da Cadeia de Suprimentos: Estratégia, Planejamento, e Operação*. Pearson Prentice Hall.

Confederação Nacional dos Metalúrgicos. Disponível em: < http://www.cnmcut.org.br/conteudo/ferbasa-se-recupera-da-crise-de-2009-com-lucro-liquido-de-36-milhoes>. Acesso em: 13 jul. 2017.

Crook, T. R., Bratton, V. K., Street, V. L., & Ketchen Jr, D. J. (2006). Has strategic management shed the normal science straightjacket?: Revisiting Bettis'(1991) critiques. *Journal of Managerial Issues*, 409-423.

Dill, R. et al. (2014). Corporate profitability analysis: a novel application for paraconsistent logic. *Applied Mathematical Sciences, 8*(26), 1271-1288.

Ehrhardt, M. C.; Brigham, E. F. (2012). *Administração financeira: teoria e prática*. Cengage Learning: São Paulo.

FALCONI. Disponível em: <https://www.falconi.com/segmentos/siderurgia/>. Acesso em 8 dez. 2017.

Fávero, L. P., Belfiore, P., Silva, F. L. da, & Chan, B. L. (2009). *Analise de Dados - Modelagem Multivariada para Tomada de Decisoes* (1st ed.). São Paulo: Elsevier.

Fernandes, F. C., & Mazzioni, S. (2015). A correlação entre a remuneração dos executivos e o desempenho de empresas brasileiras do setor financeiro. *Contabilidade Vista & Revista*, *26*(2), 41-64.

Frémont, A. (2009). Shipping lines and logistics. *Transport Reviews*, *29*(4), 537-554.

Fuenmayor, C.; Angulo, L. (2008). *Índices econômicos financieros en la toma de decisiones para empresas del sector farmacéutico del municipio "Jesus Enrique Lossada" del estado zulia"*. 2008. Trabalho de conclusão de curso. Universidad Rafael Urdaneta, Maracaibo.

García Molina, P. S., & Artola, M. A. (2010). El mito de la creación de valor en las PyMEs.

García Tabuenca, A., Crespo Espert, J. L., Pablo Martí, F., & Crecente Romero, F. J. (2007). *Perfil de los Empresarios y Resultados de sus Empresas*. Instituto Universitario de Análisis Económico y Social.

Gasparetto, V. (2004). O papel da Contabilidade no provimento de Informações para a Avaliação do Desempenho Empresarial. *Revista Contemporânea de Contabilidade*, *1*(2), 109-122.

Gitman, L. J. (2003). *Principios de administración financiera*. Pearson Educación.

Gitman, L. J. (2010). *Princípios de administração financeira*. Pearson: São Paulo.

Grant, R. M. (1991). The resource-based theory of competitive advantage: implications for strategy formulation. *California management review*, *33*(3), 114-135.

Hofer, C. W., & Schendel, D. (1978). *Strategy formulation: Analytical concepts*. West Publ..

Hwang, C. L., & Yoon, K. (1981). Multiple criteria decision making. *Lecture Notes in Economics and Mathematical Systems*, *186*, 58-191.

Indústria Aço Brasil. Disponível em: < http://www.acobrasil.org.br/site2015/historia.asp>. Acesso em 8 dez. 2017.

Leite, R. R. (2014). *Mecanismo para superação das barreiras na implantação da produção mais limpa em pequenas e médias empresas: estudo de múltiplos casos no setor metalúrgico*. Dissertação de Mestrado. Universidade Nove de Julho, São Paulo.

Libertá. Disponível em: <http://www.libertaconsultoria.com.br/atuacao.php?servico=siderurgia-e-metalurgia>. Acesso em 8 dez. 2017.

Matarazzo, D. C. (2010). *Análise financeira de balanços: abordagem gerencial*. Atlas: São Paulo.

Matos, R., Cardoso, A., Duarte, P., Ashley, R., Molinari, A., & Schulz, A. (2003). Performance indicators for wastewater services-towards a manual of best practice. *Water science and technology: water supply*, *3*(1-2), 365-371.

Ministério de Minas e Energia (MME). Disponível em: < http://www.mme.gov.br/documents/10584/1865251/Anuario_Setor_Metalurgico_2011_base_2010.pdf/a27e2c66-9d93-43b1-ae4c-4b105786f0fc>. Acesso em 16 abr. 2017.

Ministério de Minas e Energia (MME). Disponível em: < http://www.mme.gov.br/documents/10584/1865251/Anuario_Setor_Metalurgico_2013_base_2012_parte_1.pdf/ac9ca050-9ab7-4e63-8c04-ca13903bfc9e>. Acesso em: 16 abr. 2017.

Ministério de Minas e Energia (MME). Disponível em: <http://www.mme.gov.br/documents/1138775/1732813/Annu%C3%A1rio+Estat%C3%ADstico+do+Setor+Metal%C3%BArgico+2015.pdf/3cd2fe18-4daa-4e51-8899-53f0cba47573>. Acesso em: 16 abr. 2017

Ministério de Minas e Energia (MME). Disponível em: <http://www.mme.gov.br/documents/1138775/1732813/ANU%C3%81RIO+METAL%C3%9ARGICO+2016_vers%C3%A3o+3.pdf/9595d304-7072-4313-a190-b5b97a10e7a5>. Acesso em: 16 abr. 2017.

Notteboom, TE, & Winkelmans, W. (2001). Mudanças estruturais na logística: como as autoridades portuárias enfrentarão o desafio ?. *Política e Gestão Marítima* , *28* (1), 71-89.

Porter, M. E. (1996). What is strategy. *Published November*.

Poso, A. T. (2015). *A siderurgia brasileira e mundial: o desenvolvimento desigual recente*. 2015. Tese de Doutorado. Universidade de São Paulo, São Paulo.

Prahalad, C. K., & Hamel, G. (1990). The core competence of the corporation. *Boston (Ma)*, *1990*, 235-256.

Reed, R., & DeFillippi, R. J. (1990). Causal ambiguity, barriers to imitation, and sustainable competitive advantage. *Academy of management review*, *15*(1), 88-102.

Rodríguez, G. B. R.; Gutiérrez, A. C. R. (2011). *Estrategias en la organización administrativa, análisis e interpretación a los estados financieros del almacén "detalles" de la ciudad de loja, en los periodos económicos 2008-2009*. Trabalho de Conclusão de Curso. Universidad Nacional de Loja, Loja.

Siderurgia Brasil. Disponível em: http://www.siderurgiabrasil.com.br/portal >. Acesso em 8 dez. 2017.

Silva, J. P. da. (2012). *Análise financeira das empresas.* Atlas: São Paulo.

Soares, M. A. (2006) *Análise de indicadores para avaliação de desempenho econômico-financeiro de operadoras de planos de saúde brasileiras: uma aplicação da análise fatorial*. 2006. Tese de Doutorado. Universidade de São Paulo: São Paulo.

Souza, S. L. de. (2007). *Indicadores de sucesso e seus reflexos na gestão financeira das sociedades anônimas metal-mecânicas de capital aberto*. 2007. Dissertação de Mestrado. Universidade do Estado de Santa Catarina: Joinville.

Stock, J. R.; Lambert, D. M. (2001). *Strategic logistics management*. New York: McGraw Hill.

Tampoe, M., & Macmillan, H. (2000). Strategic management, process, content and implementation.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy*, *15*(6), 285-305.

The Economist. Disponível em: < http://www.economist.com/topics/iron-and-steel-production>. Acesso em: 17 mar. 2017.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic management journal*, *5*(2), 171-180.

Winter, B. (2002). *A case study determining the relevance of motor body repairs focusing on niche markets outside the insurance industry, to establish a position of competitive advantage* (Doctoral dissertation, University of Natal, Durban).

DECISION SCIENCES INSTITUTE

Drug Recommender System Using Pain Scale for Over-the-Counter Medicines: A Conceptual Framework

Mauli Dalal
Virginia Commonwealth University
Email: dalalmb@vcu.edu

H. Roland Weistroffer
Virginia Commonwealth University
Email: hrweistr@vcu.edu

ABSTRACT

Overdose and misuse of over-the-counter (OTC) painkiller medicines and non-steroidal anti-inflammatory drugs (NSAIDs) is one issue in healthcare that has been addressed insufficiently. We propose a conceptual framework for a recommender system that enables users to self-evaluate the intensity of pain experienced, interprets the user inputs, and recommends relevant drugs.

KEYWORDS: Pain scale, pain management, pain perception, recommender system, OTC

INTRODUCTION

Pain is associated with actual or potential tissue damage, and is often subjective, as the tolerance level may vary from one individual to another. Pain is categorized as being acute or chronic. Acute pain results from an injury, disease, or inflammation and usually comes on suddenly and may result in anxiety or emotional distress. Chronic pain lasts beyond the normal healing time from an injury, surgical procedure, or disease. To alleviate pain, patients are prescribed medications. Unresolved pain can result in reduced mobility and function, increased depression, feeling of loneliness, sleep disturbance, heightened anxiety, and reduced life satisfaction (Hong, Cagle, Van Dussen, Carrion, & Culler, 2016). Acute pain can be treated using over-the-counter (OTC) medicines and non-steroidal anti-inflammatory drugs (NSAIDs).

This study proposes a conceptual framework for a recommender system that enables users to self-evaluate the intensity of pain experienced, interprets the intensity of pain together with other user inputs, and recommends appropriate drugs. We limit the scope of study to managing acute pain for end-users.

The paper is structured as follows: in the following section, we provide background about pain scales, pain management using OTC medicines and NSAIDs, and medical recommender systems. Next, the proposed conceptual framework is described with its modules explained. Lastly, evaluation of the framework, limitations of the study, and directions for future research are addressed.

LITERATURE REVIEW

The International Association for the Study of Pain defines pain as an unpleasant and emotional experience associated with actual or potential tissue damage or described in terms of damage. Pain is subjective as the tolerance level varies from one individual to another. Several factors define pain and its effects. Haefeli and Elfering (2006) in their article about pain assessment for patients with back issues introduced different aspects such as pain severity, chronicity, and pain experience that define pain and its effects. Pain severity encompasses pain-related interference with activities and the intensity of pain. Chronicity encapsulates various definitions for persistent back pain – lasting for several weeks to more than a year. Pain experience comprises pain amount (how much a patient is in pain) and pain affect (degree of emotional arousal or changes in action readiness due to pain sensation(s)).

Assessment of Pain Perception

Various instruments are in use for assessing the perception of pain from a patient's perspective. These provide effective and established approaches to assessing pain, which is a subjective measure. Healthcare practitioners then interpret the rating statements by patients as no pain, mild pain, moderate pain, or extreme pain. This section presents an overview of pain scales used by practitioners.

Table-1: Widely Used Scales to Assess Acute Pain

SCALE	DESCRIPTION
Visual Analog Scale (VAS) / Graphic Rating Scale (GRS)	Consists of a linear scale (a 100 mm long line) with end points *no pain* and *worst imaginable pain*. Respondents are asked to indicate on the line the point that best represents the intensity of pain experienced (Ferreira-Valente et al. 2011). The distance from the *no pain* point indicates the pain intensity.
Numeric Rating Scale (NRS)	Most commonly used scale, with an 11-point rating from *no pain* (=0) to *worst imaginable pain* (=10). Only whole numbers are used for assessing pain. NRS provides less-subtle distinctions between different pain levels than VAS/GRS.
Verbal Rating Scale	Uses a 5-point scale consisting of the expressions *no pain*, *mild pain*, *moderate pain*, *intense pain*, and *maximum pain* that describe increasing levels of pain intensity. Respondents select the phrase that best characterizes their pain (Ferreira-Valente, Pais-Ribeiro, & Jensen, 2011).
Facial Pain Scale (FPS)	Patients are asked to select a facial expression that best captures the level of pain experienced, choosing from 6 different faces that represent increasing levels of pain intensity, where the left-most face represents *no pain*, and the right-most face represents *very much pain* (Ferreira-Valente, Pais-Ribeiro, & Jensen, 2011).
Pain Drawing	Patients are asked to mark areas of pain on the outline of human figure. At times, patients are asked to indicate different types of pain experienced using different symbols.

Haefeli and Elfering (2006) identify three instruments for assessing the intensity of pain – visual analog scale (graph rating scale), numeric rating scale, verbal rating scale, and pain drawing. Breivik et al. (2008) identify and compare different pain scales in use for assessing intensity of pain. For assessing acute pain, they identify visual analog scale (VAS), numeric rating scale (NRS), and verbal categorical rating scale (VRS). These scales best capture the subjective feeling of patient about the intensity of pain. They recommend using these scales for worst, least, or average pain over the last 24 hours or during the week. After simultaneous recordings of pain intensity on VAS, NRS, and VRS scales in a large number of patients, they found the VAS and NRS scales to be superior over the VRS scale.

Ferreira-Valente, Pais-Ribeiro, & Jensen (2011) compared the relative validity of four scales – VAS, NRS, VRS, and FPS-R – so as to detect differences in pain stimulation as well as differences between men and women in response to experimentally induced pain within a sample of Portuguese subjects**.** Glowacki (2015) discussed the improved outcomes due to effective pain management in patients with acute pain, highlights the dimensions of pain management, and reports increase in patient satisfaction with respect to effective pain management. Baratloo et al. (2016), in their study of using caffeine for managing pain, determined that caffeine is frequently used as an adjuvant therapeutic agent for curing migrant headaches.

Table 1 summarizes the scales used to assess intensity of acute pain.

Over-the-Counter (OTC) Medicines and NSAIDs

In situations where patients choose to take medicines without seeing a physician, they can purchase OTC medications to help them alleviate pain. Furthermore, pain may also require specialized diagnostic procedures like central and peripheral neural blockage or monitored drug infusions. Pain relief medications, also termed as analgesics or painkillers, are regulated by the Food and Drug Administration (FDA). These medicines reduce the sensitivity to pain by acting on peripheral and central nervous system of the body. Based upon the sensitivity of pain and its cause, suitable medications are prescribed.

<u>Classification of pain relievers</u>

The FDA classifies analgesics as OTC medications and prescription medications. The former is useful to relieve minor aches associated with fever, cold, flu, toothache, cramps, etc. OTC medicines can be further classified into the following groups:

- Acetaminophen – used as pain relievers, cough suppressants and cold medications.
- Non-steroidal anti-inflammatory drugs (NSAIDs) – relieve fever, minor aches and pains. Common medicines available include aspirin, naproxen, ibuprofen as well as those to treat cold, sinus pressure, allergies. Pain is relieved by reducing the production of prostaglandins that cause pain.

Lessenger and Feinberg (2008) in their clinical review define OTC medicines as pharmaceuticals that do not require a prescription and are sold on shelves of markets, stores, and pharmacies. The article identifies commonly used prescription drugs that have potential to being abused as well as OTC medications that may be abused. They state that prolonged use of high levels of OTC analgesics such as aspirin and acetaminophen has been associated with dysphoric mood states.

Wilcox, Cryer, and Triadafilopoulos (2005) acknowledge the issue of overdosing and misuse of OTC and NSAID pain medicines. They conducted a comparative study with two surveys – Roper Starch survey and National Consumers League – to determine if users of medicines are ignorant of the potential side effects. The types of pain identified from both surveys are arthritis/joint pain, back pain, headache, sports/exercise related pain, and others. The frequency of using these medicines are categorized as daily, several times per week, several times per month, less than several times per month, and as needed. Joint pain or arthritis is factored as the most common cause for taking medicines. The attitudes towards side effects indicate that 46% of respondents from the Roper survey believed OTCs to be safer than prescription medicines, and 37% experienced side effects. The Analysis for National Consumers League survey indicated 49% of respondents were somewhat or very concerned about side effects, and 36% believe OTCs are safer than prescription medicines. The authors insist an educational intervention for patients as well as physicians is needed.

Blondell, Azadfard, and Wisniewski (2013) summarized the medications used for treating acute pain in adults. They state that acetaminophen is the first-line medicine for treating mild to moderate acute pain. Ibuprofen and naproxen are the preferred NSAIDs for treating mild to moderate acute pain, due to their effectiveness, side effects, cost, and over-the-counter availability. Cycloygenase-2 selective NSAIDs, which have similar effectiveness to non-selective NSAIDs, are second-line medications for treating mild to moderate pain because of their greater costs. The authors also identify the safety and effectiveness of acetaminophen, aspirin, and non-selective NSAIDs from Cochrane reviews with respect to amount needed to produce 50% pain reduction in 4-6 hours.

Pardutz and Schoenen (2010), in their review of clinical and experimental data for using NSAIDS to manage acute migraine pain, identify aspirin, acetaminophen and other NSAIDs as the most frequently used drugs. The randomized control trials reviewed show evidence that acetaminophens show less efficacy as compared to other NSAIDs. The authors also emphasize that there is no evidence that any particular NSAID is more or less effective than others, with notable exception of acetaminophen, and they recommend to select a drug with rapid GI absorption and the most favorable efficacy/side effect profile. Gelfand and Goadsby (2014) in their review summarize what is known about predicting the response related to medication withdrawal.

Pombo, Araújo, and Viana (2014), in their comprehensive literature review, focused on identifying computer technologies that facilitate knowledge discovery from clinical data produced by either patients or healthcare professionals. The strengths and weaknesses for rule-based algorithms, nonstandard set theory, artificial neural networks (ANNs), and statistical learning algorithms with examples are identified from the extant literature. The issues identified from the research include lack of integration with mobile devices, use of web-based interfaces, and scarcity of systems that allow for patients to enter data.

Westergaard, Munksgaard, Bendtsen, and Jensen (2016) present their perspective about headaches resulting from overuse of medications.

Medical Recommender Systems

Sodsee and Komkhao (2013) proposed a medical recommender system combined with evidence-based medicine. Their proposed *evidence-based medical recommender system* collects data related to medical history of patients, case reports and expert opinions, and experimental studies that are analyzed with intelligent algorithms. Some of the intelligent algorithms of relevance identified are neural networks, fuzzy theories, support vector machines, and data mining techniques. The emphasis of this approach is on telemedicine, i.e. allowing physicians to access the system anywhere with Internet connectivity, and recommendations from outcomes by intelligent algorithms delivered to physicians. Furthermore, the physicians can provide consultation to their patients.

Lalloo, Jibb, Rivera, Agarwal, and Stinson (2014) in their study reviewed the pain self-management applications that are available for download from application stores for smartphones. The criteria for identifying the applications for review included goal of app (treatment or mitigation of pain) and primary user (person in pain). The assessment of the applications was done based on the researcher-developed list of important self-management functionalities. The limitations identified include a lack of integrated features for personalization, lack of involvement of health care professionals in the development and evaluation of the app, lack of foundation in current research or established theories, and lack of feasibility or effectiveness testing. The authors also stress the need of research for developing comprehensive, evidence-based, and clinically informed self-management applications for smartphones that can be used by for people living with pain.

Heidarabadi et al. (2017) reviewed articles dealing with information systems for pain management of patients with spinal cord injuries. The 18 articles reviewed demonstrate the positive role of such systems in increasing accuracy and improving physician’s decision-making. Although the study is specific to chronic pain from spinal cord injury and identifies advantages for physicians in that context, some of its disadvantages are also applicable to the context of our study. The authors state that the greatest disadvantages of the systems reviewed are a lack of integration with mobile devices, a lack of web-based interfaces, and a lack of features that allow data entry by patients. Our proposed framework addresses this issue by allowing for the users to rate their pain and input the pain sensations experienced.

Ozsoy, Özyer, Polat, and Alhajj (2018) use a *Pareto dominance and collaborative filtering* method adapted from recommender systems to address the drug-repositioning problem. The selected method identifies drugs most similar to a target drug as well as neighbor drugs in order to predict new indication of the target drug.

We are not aware of any published studies dealing with recommender systems for selecting self-treatment with OTC pain relief medicine.

CONCEPTUAL FRAMEWORK

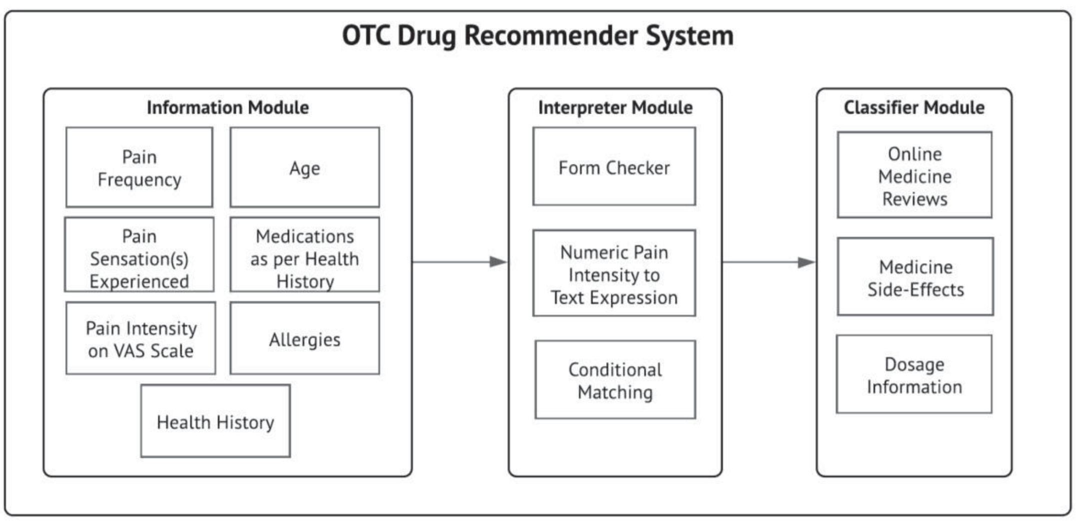
As shown in Figure 1, our proposed framework consists of three modules – *information module*, *interpreter module*, and *classifier module*.

The ***information module*** requires users to provide information related to their age, existing allergies, medical conditions (health history), other medications taken, the type of pain sensation, how often the pain occurs (pain frequency), as well as rate the intensity of pain experienced using VAS scale. We consider age, medications as per health history, allergies, pain intensity on VAS scale, and health history to be the mandatory fields.

Cues related to different types of pain sensation with explanation are presented to the users, so as to clarify the terminology related to pain sensation, and thus reducing errors in selecting incorrect options, which do not apply to their situations. Some examples of pain sensations include throbbing, palpitation, aching, shooting, pins and needles, pinching, crawling, itching, numbness, tightness, burning, radiating, nagging, phantom pain, etc. It is also important to have information about all medications taken by users for any allergies they may have. This information is required for recommending appropriate pain medicines given the interpretation of pain scale, and ensuring that the recommended medicine does not react with their ongoing treatment. Pain frequency requires for users to input how frequent they experience these pain sensations, and possibly provide description about the duration of the sensation. For example, a patient may experience a shooting sensation in a foot, the first thing every morning. If the users cannot find the specific pain sensation from the terms listed by the recommender system, they will be presented with an 'other' option. Selecting this option will then allow them to describe the sensation as experienced in their own words.

Pain intensity is entered on the VAS scale and users are provided instructions about using this scale to express the intensity of pain, to ensure that they do not assess the pain incorrectly.

Figure 1: Proposed Conceptual Framework



The ***interpreter module*** consists of a form checker, a convertor for numeric scale to text expression, and conditional matching. The form checker does a validation check to ensure that the required information (mandatory fields) is entered by the user. Age, medications as per health history, allergies, pain intensity on VAS scale, and health history are the mandatory fields. If any of these mandatory fields are incomplete, users are asked to provide the missing information and complete the form. The numeric level of pain is read and interpreted as a text expression. Collins, Moore, & Mcquay (1997) state “Our results indicate that if a patient records a baseline VAS score in excess of 30 mm they would probably have recorded at least moderate pain on a 4-point categorical scale; in excess of 54 mm and they would probably have recorded severe pain.” Using the interpretation of the 100-mm VAS scale as per the study by Jensen, Chen, & Brugger (2003), a rating of 0 to 4 mm can be considered no pain, 5 to 44 mm can be considered mild pain, 45 to 74 mm can be considered moderate pain, and 75 to 100 mm can be considered severe pain.

Conditional matching takes into consideration the data from the information module and the diagnosis of the user to determine the most effective medicine for the user. Conditional matching within the interpreter module is used to consider different scenarios based on allergies and medications taken by users, and the indicated pain sensation, frequency, and intensity, to identify which medicine is the most effective. A simple example for conditional matching is, say user1 who experiences pain due to cancer, versus user2 who experiences pain due to inflammation. Their respective subjective assessment of pain is likely to differ. That is, the severe pain experienced due to inflammation will be differentiated by conditional matching from the severe pain that is experienced by the cancer patient. Some users may accidentally provide information related to allergies, medicines as per health history within the health history field. This issue can also be considered as a conditional matching issue. In this context, conditional matching may distinguish between phrases used by users to explain their allergies from medicines taken as per health history.

The **classifier module** directs the users to online reviews of recommended medicines, their side-effects, and dosage information. It is important to note that the performance of the classifier module is dependent upon the information that is provided by users and the interpretation of pain by the interpreter module.

EVALUATION APPROACH

The proposed framework requires basic information related to users’ health history, allergies (if any), medicines taken, pain sensation, its frequency as well as their age. This information, in addition to the patient’s personal information is collected by healthcare practitioners before designing a treatment plan for the patient. In our framework, users are not required to provide personal information, but based upon the information they provide they will be informed about relevant over-the-counter (OTC) medicine(s) that can be taken by them, as well as reviews that describe possible side-effects and appropriate dosage information. This framework guides the development of recommender systems that can address the issue of overuse and abuse of OTC medicines.

The evaluation of information and interpreter modules is done simultaneously so as to ensure that the system can make correct interpretations about pain intensity, guide users to provide details in mandatory fields, and use conditional matching to recommend medicines to users.

To evaluate the performance of the classifier module, data is obtained from drug review websites (drugs.com, askapatient.com, and drugratingz.com); especially those websites with user reviews, information related to side effects, and dosages for the said medicine. Upon collecting reviews and cleaning data, text-mining software, such as QDA Miner and Enterprise Miner SAS can be used for analysis.

By synthesizing the data collected from these sources, the performance of the classifier module can be evaluated based upon its recall, precision, and accuracy.

Table 2 presents the conceptual definitions for the evaluation criteria.

Table-2 Evaluation Criteria

CRITERIA	DESCRIPTION
Recall	% of total relevant results correctly classified by the classifier
Precision	Proportion of outcomes correctly classified by the classifier
Accuracy	Proportion of negative outcomes classified correctly by the classifier

The performance of the overall framework with respect to information provided by users, the intensity of pain experienced, allergies, and age, and the interpretation of pain scale, and medicine recommendations must be evaluated with a healthcare practitioner. Furthermore, the authenticity of reviews provided by users on the websites, i.e. accuracy of side effects stated for the said medicine and appropriate dosages, can be validated only by healthcare practitioners. Hence, the recommendations presented by the classifier module should first be validated with healthcare practitioners. Also, the overall prediction as per our proposed framework for recommender system must be evaluated by healthcare practitioners.

LIMITATIONS AND FUTURE DIRECTIONS

The proposed framework needs to be evaluated in two ways. First, by simulation with QDA miner and Enterprise Miner SAS using data such as the required information in the information module and the classifier module. Second, the outcome must be validated by healthcare practitioners. The performance of the classifier module in terms of dosage information, side effects of medicine, and online medicine reviews should also be evaluated and validated by healthcare practitioners, to avoid providing users with incorrect dosage information and side effects for the medicines. In our proposed conceptual framework, we have not included healthcare practitioners, which is a limitation that should be addressed in future.

Our framework does not consider updating reviews of medicines and classifying comments for newly added comments. This is an issue that can be addressed in future research. Additionally, our framework does not incorporate a 'ask for help' feature where users can consult healthcare practitioners about the drug recommended by the system, or instances wherein they are unclear about side-effects of the recommended medicines.

Lastly, security and privacy concerns in extant healthcare systems are excluded from the scope of our framework. The users are not required to register with their email information for using our framework, and their medical information provided in the information module is not stored within our framework. Hence, future research can incorporate additional features that allow users to create their profile(s), update their information by offering enhanced security identification protocol. The security identification protocol can then be capable of addressing security and privacy concerns that are widely encountered in healthcare systems.

CONTRIBUTION AND CONCLUSIONS

Our conceptual framework presents a new perspective for recommender systems that may lead to the development of practical applications that allow users experiencing pain to self-manage that pain and find appropriate treatment in an efficient and effective manner. Furthermore, our framework may also lead to new ideas for more general research into medical recommender systems based on user needs and user provided input.

REFERENCES

Baratloo, A., Rouhipour, A., Forouzanfar, M. M., Safari, S., Amiri, M., & Negida, A. (2016). The Role of Caffeine in Pain Management: A Brief Literature Review. *Anesthesiology and Pain Medicine*, *6*(3), 33193. https://doi.org/10.5812/aapm.33193

Blondell, R. D., Azadfard, M., & Wisniewski, A. M. (2013). Pharmacologic Therapy for Acute Pain. *American Family Physician*, 766–772. Retrieved from http://www.goodrx.com

Breivik, H., Borchgrevink, P. C., Allen, S. M., Rosseland, L. A., Romundstad, L., Breivik Hals, E. K., … Stubhaug, A. (2008). Assessment of pain. *British Journal of Anaesthesia*, *101*(1), 17–24. https://doi.org/10.1093/bja/aen103

Collins, S. L., Moore, R. A., & Mcquay, H. J. (1997). The visual analogue pain intensity scale: what is moderate pain in millimetres? *PAIN*. Retrieved from https://pdf.sciencedirectassets.com/271050/1-s2.0-S0304395900X00255/1-s2.0-S0304395997000055/main.pdf?x-amz-security-token=AgoJb3JpZ2luX2VjEMP%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJIMEYCIQCU5nDwnbrKe137FG1nBVPbZDrd0dBvqi5q02czETQVHAIhAPXtQmpRGOTq

Ferreira-Valente, M. A., Pais-Ribeiro, J. L., & Jensen, M. P. (2011). Validity of four pain intensity rating scales. *Pain*, *152*(10), 2399–2404. https://doi.org/10.1016/j.pain.2011.07.005

Gelfand, A. A., & Goadsby, P. J. (2014, July 5). Medication Overuse in Children and Adolescents. *Current Pain and Headache Reports*. Springer US. https://doi.org/10.1007/s11916-014-0428-1

Glowacki, D. (2015). Effective Pain Management and Improvements in Patients' Outcomes and Satisfaction. *Critical Care Nurse*, *35*(3), 33–41. https://doi.org/10.4037/ccn2015440

Haefeli, M., & Elfering, A. (2006). Pain assessment. *European Spine Journal*, *15 Suppl 1*(Suppl 1), S17-24. https://doi.org/10.1007/s00586-005-1044-x

Heidarabadi, N. N., Safdari, R., Saeidi, M. G., Kolivand, P., Tahmasbian, S., & Hakemi, L. (2017). A review of developed information systems for pain management in patients with spinal cord injury. *Acta Medica Mediterranea*, *33*(2), 237–244. https://doi.org/10.19193/0393-6384_2017_2_035

Hong, S., Cagle, J. G., Van Dussen, D. J., Carrion, I. V., & Culler, K. L. (2016). Willingness to use pain medication to treat pain. *Pain Medicine (United States)*, *17*(1), 74–84. https://doi.org/10.1111/pme.12854

Jensen, M. P., Chen, C., & Brugger, A. M. (2003). Interpretation of Visual Analog Scale Ratings and Change Scores: A Reanalysis of Two Clinical Trials of Postoperative Pain. *The Journal of Pain*, *4*(7), 407–414. https://doi.org/10.1016/S1526-5900(03)00716-8

Lalloo, C., Jibb, L. A., Rivera, J., Agarwal, A., & Stinson, J. N. (2014). "There's a Pain App for That." *The Clinical Journal of Pain*, *31*(6), 557–563. https://doi.org/10.1097/ajp.0000000000000171

Lessenger, J. E., & Feinberg, S. D. (2008). Abuse of Prescription and Over-the-Counter Medications. *The Journal of the American Board of Family Medicine*, *21*(1), 45–54. https://doi.org/10.3122/jabfm.2008.01.070071

Ozsoy, M. G., Özyer, T., Polat, F., & Alhajj, R. (2018). Realizing drug repositioning by adapting a recommendation system to handle the process. *BMC Bioinformatics*, *19*(1), 136. https://doi.org/10.1186/s12859-018-2142-1

Pardutz, A., & Schoenen, J. (2010, June 17). NSAIDs in the acute treatment of migraine: A review of clinical and experimental data. *Pharmaceuticals*. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ph3061966

Pombo, N., Araújo, P., & Viana, J. (2014). Knowledge discovery in clinical decision support systems for pain management: A systematic review. *Artificial Intelligence in Medicine*. Elsevier B.V. https://doi.org/10.1016/j.artmed.2013.11.005

Sodsee, S., & Komkhao, M. (2013). Evidence-based Medical Recommender Systems : A Review. *International Journal of Information Processing and Management*, *4*(September), 114–120. Retrieved from https://pdfs.semanticscholar.org/55f2/95e4e2963f976b727f72371602ff7a91fd4b.pdf

Westergaard, M. L., Munksgaard, S. B., Bendtsen, L., & Jensen, R. H. (2016). Medication-overuse headache: a perspective review. *Therapeutic Advances in Drug Safety*, *7*(4), 147–158. https://doi.org/10.1177/2042098616653390

Wilcox, C. M., Cryer, B., & Triadafilopoulos, G. (2005). Patterns of use and public perception of over-the-counter pain relievers: Focus on nonsteroidal antiinflammatory drugs. *Journal of Rheumatology*, *32*(11), 2218–2224. Retrieved from http://www.jrheum.org/content/32/11/2218http://www.jrheum.org/alerts1.SignupforTOCsandotheralertshttp://jrheum.com/faqwww.jrheum.orgwww.jrheum.orgDownloadedfrom

DECISION SCIENCES INSTITUTE

Decision Analysis in Distributed Air Sampling Cloud Network in Epidemic and Pandemic Prediction

Joseph Bango
University of Bridgeport
Email: jbango@my.bridgeport.edu

Ruba Deeb
University of Bridgeport
Email: rubadeeb@bridgeport.edu

ABSTRACT

Proposed here is a new concept for a network of specialized continuous ambient air sampling systems, which employ a novel non-destructive ionization and separation method, coupled to a near real-time genomic sequencer. The network of preferably viral samplers would constitute a plurality of distributed nodes throughout the world, with viral concentration, identification, and mutation data uploaded to the cloud for epidemic/pandemic predictive modeling. The proposed model offers the ability to migrate from outbreak surveillance, to outbreak forecast. In addition, the capability for continuous pathogen data of predominately viral genomic sequencing offers an enhanced capability to help track antigenic drift and antigenic shift. While optimized for viral capture and analysis, any airborne pathogen or spore can be accepted using the technology. An effective monitoring system necessitates a predictive modeling system that is dynamic, which responds and adapts to similar epidemiological infection trends. Applications include world health monitoring, pandemic prediction, and detection of real-time bioterror pathogen deployment.

KEYWORDS: Epidemic, Pandemic, Air Sampling, Cloud, Electrospray, Genomic Sequencing

INTRODUCTION

Early efforts to predict the spread of disease have been traced back to the 14th century. Most notable in historical epidemics is the Black Death, which engulfed Europe from 1347 to 1351, killing an estimated 200 million people. Attempts to predict the spread of disease have, until recently, relied essentially on weather predictions coupled with the emergence of known vectors of a given pathogen. New software developments have allowed computer models to make use of predictive analysis of disease trends to help isolate the most likely path for disease expansion. Data from hospitals and clinics, combined with key word search of social media and air travel data add additional variables to the prediction process.

Mathematical models for epidemic spread were first employed in the early 20th century, which separated susceptible individuals into groups, which relate to various stages of the infectious disease. Such stages may include exposure to known vectors, infectious but not symptomatic, symptomatic, and death. Transition from one group or compartment to another are predicted using differential equations (Fuppa, 2017). These deterministic models necessitate prior history of the disease to yield useful prediction results (Brauer & Castillo-Chavez 2012). The most common predictive model is known as the SIR approach, which takes into consideration those individuals who are susceptible 'S' to a specific disease, such as infants or the elderly, the quantity of those already infectious represented as 'I', and the quantity of those in an exposed population that have recovered as ‘R’, who have now developed immunity. The model, however, is not applicable in cases where patients do not develop immunity, as in the case of influenza, and populations are continually at risk for redeveloping the disease, hence the need for continual vaccine development based on antigenic shift and antigenic drift. Antigenic shift is due to genetic mutation of a human flu virus after human exposure to an animal viral species that also can affect humans. Antigenic drift is caused by random events that alter the viral structure, which require new immune system adaptation for antibodies to recognize the new antigen sequence. Because of these changes in many airborne infectious diseases, defining who has been exposed and the time of exposure that defines the onset of the disease becomes ever more imperative if the transition rate to a given population is to be predicted with any accuracy. Clearly, a means by which additional near real-time genetic sequence information about airborne viral exposure to susceptible populations, and the exposure magnitude to a target population, would yield improved disease onset extrapolation, and help mitigate potential epidemic or pandemic events. Continuous airborne sampling and genetic sequencing of viral species coupled to a predictive model would offer a vast improvement in epidemiological prediction of evolving influenza or even potential bioterror agent exposure to large populations, mitigating the level of risk across those susceptible to a given pathogen threat.

LITERATURE REVIEW

The literature review pertaining to this study did not reveal any results on continuous air sampling of potential viral infectious agents where real time or near real time genetic sequencing has been proposed or attempted. Present methods of genetic sequencing rely on incubating samples in the lab in preparation for genetic amplification (PCR or Polymerase Chain Reaction) and subsequent sequencing, which can take upwards of a day or more to complete. A faster approach as proposed herein, could have a significant benefit in reducing the spread of a disease, and even help in narrowing the identification of patient zero, the source of a communicable disease. The review covering the decision process for both naturally occurring infectious diseases and for deliberate events relating to terrorist biowarfare pathogen release left the identification of a given pathogen in the hands of a clinic or hospital. The time constant by which patent admission, examination, sample collection, culturing, and genetic or other identification of the infectious agent, ranged from hours to days after exposure. In each instance, the exposure to susceptible populations cannot be ascertained with any degree of accuracy. Epidemiological data management systems (epiDMS) are increasingly being utilized to help predict the next epidemic spread. However, to be effective, real time data is necessary. In epiDMS, “the data, models, and underlying model parameters dynamically evolve over time. This necessitates continuous analyses and interpretations of the incoming data and adaptation of the networks and models. Therefore, simulation ensembles may need to be continuously revised and refined as the situation on the ground changes” (Liu et al, 2016). With epiDMS, healthcare providers provide the input data across the region of suspected infection. The modeling software does not by itself collect any pathogen data nor does it develop dynamic sequencing of a particular airborne pathogen. “Once the query is executed and the relevant simulations are identified, epiDMS then organizes the results in the form of a navigable hierarchy, based on the temporal dynamics of the disease: scenarios that result in similar patterns are grouped under the same branch, while simulations that show key differences in disease development are placed under different branches of the navigation hierarchy. The user can then navigate on this hierarchy using drill-down and roll-up operations and filter sets of simulations for further analysis” (Liu et al, 2016).

Many references refer to surveillance systems, but do not articulate exactly what those surveillance systems are, other than relying on hospitals and clinics to define what is and is not a verified outbreak of the target infectious agent. Because of the limited amount of available data in predictive models, some scholars have suggested the use of *Partially Observable Markov Decision Processes* (POMDPs) to anticipate a disease outbreak (Izadi & Buckeridge 2007).

The same situation of limited real time input data exists for bioterror predictive analysis. The assumption in this case is that disease will spread rapidly, in order to become as much a tool mass disruption as it would be a tool of mass destruction. Accordingly, inputs from medical care facilities, work absenteeism, and social media are currently the primary data sources for existing mathematical models. The goal of the Department of Homeland Security's *'BioWatch'* was to have an automated system that continuously samples the air in an effort to add a new near real time dimension to the aforementioned input variables, however all attempts at an automated BioWatch have met with failure. The equipment required for an airborne sample capture and subsequent processing were the key factors in DHS's BioWatch failure. At present, BioWatch consists of rooftop 'witness plate' air samples that are cultured and sequenced daily and brought to a laboratory for information on new and emerging disease. The process is cumbersome, hardly real time, and fraught with errors.

Other predictive models that look at animal farm infection, which could in some instances serve as vectors for human disease as in swine and bird flu, employ "an epidemic model to observed, spatially-explicit, infection data at weekly intervals throughout two historical outbreaks of foot-and-mouth disease, UK in 2001 and Miyazaki, Japan in 2010 (Probert, Jewell, et al, 2018).

Improved Epidemic Prediction With Live Agent Sequencing and Cloud Interface

Recent research by our group suggests that an improvement in epidemic prediction can be achieved using a novel means of distributed nodes capable of ambient live agent pathogen assessment, specifically genomic sequencing, with DNA/RNA data uploaded to the Cloud via the world-wide web.

In the following *Table 1*, some of the features of the proposed distributed pathogen air sampling nodes coupled to the cloud is summarized. The proposed system air sampling nodes in mass production are targeted to cost less than $5,000 U.S., while the DHS BioWatch program remains very costly because of the sheer number of highly trained personnel to prepare, distribute, and later collect and process the culture plates necessary. Distributed automated nodes also can be employed to sample air pathogens at animal farms which may be located great distances from urban centers and thus can add many hours or days to a BioWatch-like witness plate sampling approach to be conducted. One especially useful added advantage to an automated air pathogen system connected to 'the internet of things', is that selected nodes can be designed to automatically process mosquitos to look for the presence of such insect borne diseases such as Zika, Malaria, and EEE or Eastern Equine Encephalitis, to name a few. Early identification via near real-time genetic sequencing of insect borne disease jump starts the epidemic prediction process because no healthcare workers are required to identify the presence of a known disease in a given geographic area. Those at risk can be simply warned to taken appropriate precaution thus averting any epidemic.

HYPOTHESIS

We hypothesize that automated airborne live agent sampling nodes coupled to the cloud can be situated in multiple public spaces in greater numbers than a DHS BioWatch concept as presently employed by the U.S. government, because of low cost and small size of the approach. The epiDMS is a software package that relies exclusively on the data input of healthcare workers who have taken patient samples and identified a given airborne pathogen through laboratory testing. The proposed nodal analysis does not rely exclusively on healthcare worker input, but complements such data inputs to arrive at a more accurate prediction of potential disease to spread in a given geographical populace. Furthermore, a cloud-based nodal analysis of pathogens in public spaces such as airports, train and bus stations, allows predictions to be made regarding cities at risk based on travel destinations. Appearance of a disease in a new city after human-host incubation can provide data, which can be used to truncate the source via traveller bookings at an airport.

Table 1 –
Pros / Cons of Traditional Predictive Modeling vs. Pathogen Sampling & Cloud Interface

Comparison of Epidemic Prediction System				
TECHNOLOGY	Public Space Monitoring	Detection Under 1 Hour	Stand Alone	Cost per node
Sampling Cloud Nodes	YES	YES	YES	YES
DHS BioWatch	YES	NO	NO	NO
epiDMS	NO	NO	NO	NO

<u>The Importance of Detecting Viral Pathogens in Near Real Time in Public Areas</u>

Computational software promise to help stem the spread of disease that can decimate civilian populations, and incapacitate military forces. “Forecasts would be extremely helpful to public health officials in containing infectious diseases, but it is really difficult. The science of forecasting is a work in progress. It’s akin to trying to solve a jigsaw puzzle with some of the pieces missing and a vague sketch of what the finished image should look like” (Bushey, 2016). Because of the time constant between when an individual becomes exposed and infected to an airborne pathogen such as influenza, and when that individual becomes symptomatic may be many days (1-4 day incubation period), vector transmission to other persons will likely have occurred. Only a small fraction, generally the very young or the very old, in the early stages, will end up in a hospital or clinic where positive identification of the virus has been obtained and is available as a statistic. The ability to monitor the level of viral load in a local population outside of the hospital or clinic setting, coupled with an ability to know precisely what the virus genome sequence, would advance the state-of-the-art in epidemic predictive capability substantially. We believe we may be able to achieve this goal through the merging of several well-established technologies, and the combination of new developments in gene sequencing to be described.

Influenza, as well as many bacterial agents, can survive for hours as an airborne species. Because droplets from a cough or sneeze encapsulate many virons, the probability for disease transmission in congested public areas is very high. In general, particles that are extremely small (< 5 microns) can remain airborne for many hours and the smaller the particulate, the longer the residence time. As virons move through the air, evaporation of water (90% water, 0.5-5% mucous) causes the droplet to become smaller and thus more mobile. Residence time is further enhanced by humidity and the lack of direct sunlight. The research lab of Dr. Lydia Bourouiba at MIT has investigated the fluid dynamics of disease transmission, and has determined experimentally that cough ejecta and sneeze cloud duration of typically approximately 250 milliseconds, can literally shoot across a room in a matter of seconds, 200 times farther than previously thought (Bourouiba, 2019).

Given the high airborne viral load that exists in congested public spaces, it would seem natural to seek out such locations to place a viral air sampler that can provide real time indication of ambient viral species. The goal of such a sampler would be to capture trace species indicative of viral agents, concentrate such species to a level of at least 200 nanograms or better, and subject the viral species to continuous genomic sequencing. The viral concentration and sequence data ideally could then be uploaded to the cloud, for deterministic computer modeling (where successive operations are based on the previous state) that merges the data with other variables previously described, such as PAHO (Pan American Health Organization), weather, and social media input variables. The added advantage of the sequencing functionality is that emergent strains can be revealed in an almost near real-time fashion even before a patient-vector becomes symptomatic. An added benefit is that many patients who are symptomatic may never seek professional medical attention, and thus never end up as a PAHO data point. A distribution of automated viral air samplers worldwide, seeded in locations such as airports and train stations, could provide data never previously attainable in epidemiology.

METHODS

The proposed system is comprised of several key parts necessary to sample ambient air laden with target pathogens. The entrance to the air sampler utilizes a common aerosol impaction filtration method. High velocity air is channeled around a 90-degree bend, causing heavier fractions to impact and embed in the end channel. Soft media such as low vapor pressure grease allow heavy fraction to be permanently entrained. As lighter fractions negotiate the change in direction, subsequently smaller and smaller fractions can be derived through the stratification based on mass through the airflow. At a suitable point downstream, a cross current soft ionization source is employed to nondestructively capture bioaerosols from the airflow, and remove them from the air stream via electrostatic forces. To employ electrostatic separation, the bioaerosols are nondestructively ionized using a process known as ‘electrospray’. Electrospray is a process by which a conductive fluid, in this case an aqueous solution, is utilized to produce a nanospray of desorbing droplets which attract and capture polar or polarizable species from the air, of which virons are a part. The ability to non-destructively ionize a virus has been well reported in the literature (Bothner & Siuzdak, 2004), where electrospray air capture of polar and polarizable species was accidentally discovered by John Fenn et al at Yale in 1984. As a charged species, final filtration is accomplished using an electrostatic filter, similar to those found on magnetic sector mass spectrometers. In other words, once we have eliminated any dust or particulates using a prefilter (i.e., inertial aerosol separation), pathogens such as bacteria or viruses, can be grouped based on their difference in mass. A charged species passing through an insulated flight tube coupled to a sensitive electrometer circuit can detect the ‘image charge’ of the pathogen species passing through. The charge can provide information about the physical size and mass of the species, thus helping to sort viruses verses bacteria. This approach is known as CDMS or Charge Detection Mass Spectrometry (Pierson, et al., 2016).

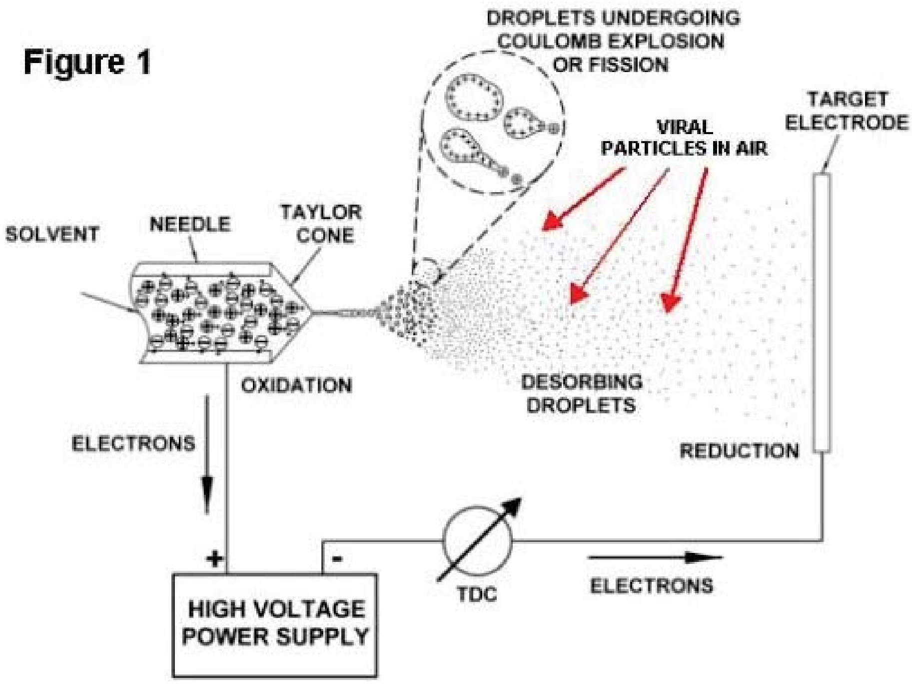
Once the virons have been selected, they are accumulated on a ‘witness plate’, in preparation for automated sequencing. The sequencer of choice is a ‘nano-pore- device, manufactured by Oxford Nanopore, and recently demonstrated aboard the International Space Station or ISS. The Nanopore system employs “bespoke, proprietary pore-forming proteins to create pores in membranes. Pore-forming proteins are common in nature. For example, the protein α-hemolysin and similar protein pores are found naturally in cell membranes, where they act as channels for ions or molecules to be transported in and out of cells. A protein nanopore α-hemolysin is a heptameric protein pore with an inner diameter of 1 nm, about 100,000 times smaller than that of a human hair. This diameter is the same scale as many single molecules, including DNA" (Oxford Nanopore, 2018). In the Nanopore DNA sequencing, the system may process the sample until a minimum of ten-fold read coverage over specified regions of interest has been seen, until a specific mutation has been observed in a sample or until enough sequence data has been collected to reliably assemble a sample against a reference (Oxford Nanopore, 2018). As DNA/RNA moves through the pore, changes in electric potential are translated into genetic sequence data. In discussion with James Brayer (Bango-Brayer email 2016) U.S. representative for Oxford Nanopore, it is feasible to adapt the sequencer to the proposed aerosol viral collector-concentrator. The process then is capture, concentrate, extract, genetic library prep, and continuous flow sequencing. The present system sensitivity is 200 nano grams of genetic material, but lower levels are anticipated. Importantly, no PCR or Polymerase Chain Reaction, an amplification step to duplicate genetic material in a concentration sufficient for identification, is required. Protein nanopores such as alpha-hemolysion or MspA have been demonstrated to work well. A longer-term solution with improved performance may be a graphene nanopore. Regardless of the genetic sequencer employed, the electrospray capture and pathogen selection concept is adaptable to any subsequent downstream analytical tool. The electrospray capture technology has been proven and applied to biowarfare agent capture under prior Navy, Marine Corps, and National Academy of Science support. In addition, the technology was featured in a NASA Tech Briefs edition in 2011-12 (Bango et al) after application on a spacecraft air filtration system contract that included Dr. John Fenn and former astronaut Dr. Buzz Aldrin (Bango & Dziekan 2013).

RESULTS

The lead author and his colleagues at Yale under the late John Fenn (2002 Nobel Prize Chemistry) discovered that electrospray ionization, now traditionally used for mass spectrometry, has the ability to "getter" or capture polar or polarizable trace species from the air, of which bio-species are a part. As is well known, since 9/11, there has been a flurry of development of many new methods and devices to identify potentially toxic agents of all kinds. But what seems to be missing is the fact that one needs to capture enough of the target species in order to achieve detection. Dr. Fenn once stated that the sensors should not be called detectors, but "identifiers", and that a "detector" is really composed of a collection device and an identifier. If we collect more analyte from the surrounding air-diluent, we need a less sensitive identifier and vice-versa. It is even better if the collection device captures and concentrates only the class of materials of interest while ignoring the background. Electrospray is unique in that it allows for non-destructive capture of all airborne pathogens, including spores, bacteria, and viruses out of the ambient air. For the purposes of tracking the spread of influenza, the present effort is directed at viral particles. Electrospray was shown to capture tobacco mosaic virus as far back as 2003 (Fuerstenau et al, 2003).

Subsequent experiments in our lab have revealed that ambient bacteria and viruses can be softly ionized by electrospray and carried into the partial pressure region of a vacuum system, detected using charge detection, and deposited on a 'witness plate' made 'sticky' using a thin layer of ionic liquid. Ionic liquids are room temperature salts, which have essentially zero vapor pressure. (Under present NASA support, contract 80NNSC19C0329, a system using this capture approach is being designed for possible microorganism life detection on the Saturn moon Enceladus).

Figure 1



Basic electrospray configuration illustrating an analyte dissolved in a solution and transformed into a jet of discrete charged droplets through the formation of a 'Taylor Cone'. The evaporation of the solvent species leaves behind just the analyte as a multiply charged ion or particle. This is a soft ionization process, which leaves fragile biospecies unharmed. (Source: Bango et al)

Viral Sequencing

The input to the pathogen sample pre-sorts the size of input media using well-known aerosol techniques. In the preferred embodiment of an ambient pathogen sensor, ionized viral species are introduced into a partial pressure region, and impacted onto a nanopore plate coated with a thin layer of ionic liquid previously described. The method of viral identification is to use the ionic liquid as a nanopore electrolyte that channels DNA and RNA to develop genetic sequence data, which is then uploaded to the cloud. As a result, viral sequence data can be collected and disseminated from multiple sources in near real time. Ideally, the sources of air sampling include all locations of human congestion, such as key travel points. These include airports, train and bus stations, shopping malls, public buildings, as well as hospitals and clinics. Livestock and avian breeding facilities can be included. The resulting influx of genetic data, combined with existing hospital and clinic data, plus travel and social media inputs, will enhance epidemiological predictive models for the spread of viral based disease.

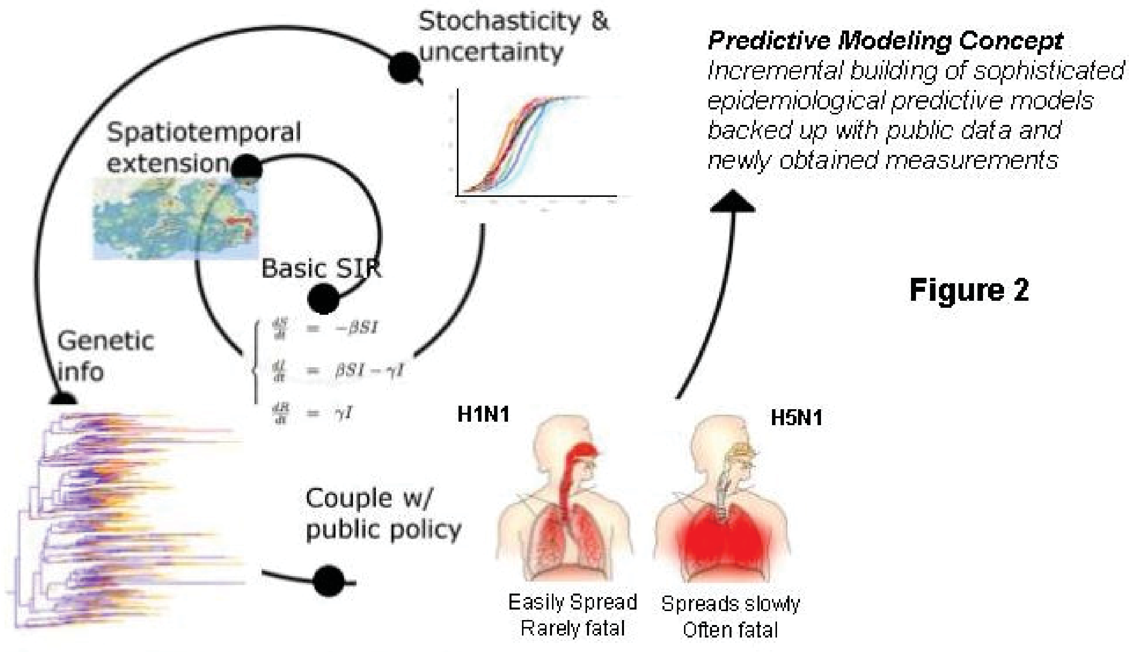


Figure 2

Source: Yannis Pantazis, University of Crete, Department of Mathematics

<u>Decision Inputs</u>

The ability to share near real time viral genetic sequence data from multiple air samplers situated around the globe via the cloud, with comparisons to archived and other real time epidemiological data sources, would offer significant saving of life due to improved decisions based on deterministic computer modeling. The decision to curtail travel, impose quarantines, or to assist in precise vaccine fabrication and distribution all benefit from knowing precisely what pathogen is present where and when, and how presented to what population at risk.

Referring to Figure 3, additional information to supplement the decision algorithm is virtually unlimited. Ideally, verified infection attributable to a specific genetic sequence of an airborne pathogen from clinics and hospitals, coupled with current information about specific disease data from the Center for Disease Control, Homeland Security, the World Health Organization, and more serve to help track the spread of infection. Official healthcare and department of defense sites in the U.S. and with respective agencies in other participating countries, is provided for Cloud access as a genetic sequence. Information sourced from social media will be less precise, subject to word or phase identification of such general information such as individual reporting of personal or family illness. While social media is obviously not a reliable data input as is sequencing data in a laboratory, a critical mass of postings relating to disease may help improve epidemic model predictive results by affecting the sensitivity threshold of when certain predetermined interventional action should be taken in a suspect disease outbreak. Such intervention may be to concentrate specific regions to increased sample acquisition.

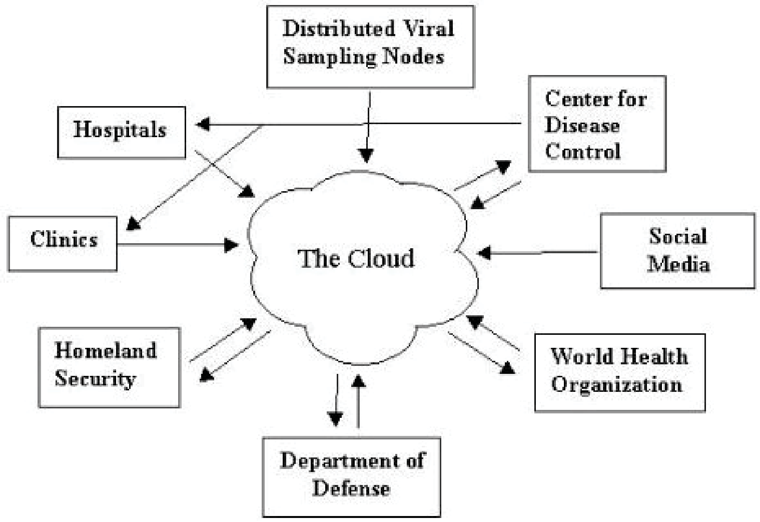


Figure 3

Schematic representation of preferred data flow to a worldwide cloud network utilizing distributed viral sampling/sequencing nodes.
The equivalent to U.S. CDC and DHS in each participating country are not shown but would participate in the sharing of data.

Source: J. Bango

Sensitivity in Epidemic Decision Making & Risk Management

A significant problem in any epidemiological detection system concept or methodology, is ascertaining when to signal an alert regarding the possible beginnings of an epidemic. How many cases signify the start of an epidemic, and will a public alert make the situation worse are questions that come to mind? An airborne pathogen sampler situated in a public location which acquires and sequences a new strain of influenza for example, is not a condition sufficient to trigger an alert. Identical strains must be conclusively identified by healthcare professionals in a clinical setting where the infectious capability and debilitating results in humans has been clearly and definitively verified. While data input for identical strains may be derived from multiple distributed nodes, i.e., airborne samplers located in multiple airports for example, there will be a threshold at which public notification and/or travel restrictions need to be implemented. The combination of new strains, clinical cases, and defined spread of such strains radiating across a known geographic area, each serve as discrete decision inputs in a Boolean truth table. The results of the decision matrix in such a truth table determine when proactive action to stem the infection spread must transpire. However, if public notification of what is believed to be the onset of an epidemic occurs, there is a strong potential for psychosomatic responses which may have contracted the pathogen but are not yet symptomatic, either due to a delay in incubation or who may be affected but are only carriers. Such individuals seeking treatment believing they are sick but yet test positive to an immunoassay test could skew the trigger point in a decision tree to restrict travel or initiate inoculation protocols or other interventional activity. There could also be false positives or other variables where faulty inputs may skew the decision matrix.

A 30-year-old sensitivity conjecture involving how to measure the result in a given Boolean function, determines the result of an output bit based on the input string. Until recently, variable factors that determine whether a given input bit is asserted high or low, has not been quantified mathematically. Dr. Hao Huang of Emory University recently published an elegant proof, which solves this problem in computer science. “This "sensitivity" conjecture has stumped many of the most prominent computer scientists over the years, yet the new proof is so simple that one researcher summed it up in a single tweet: "This conjecture has stood as one of the most frustrating and embarrassing open problems in all of combinatorics and theoretical computer science," wrote Scott Aaronson of the University of Texas, Austin, in a blog post.

The conjecture concerns Boolean functions. Every computer circuit is some combination of Boolean functions, making them "the bricks and mortar of whatever you're doing in computer science" (Quanta Magazine, 2019; Clark, 2019). Furthermore, scientists at the French National Center for Scientific Research, described the sensitivity dilemma as “filling out a series of yes/no questions on a bank loan application.” that provides a score that determines qualifications. This process is a Boolean function where answers to questions are the input bits, and the decision is the output bit.

It is conceivable that an erroneous answer to a single question can change a decision outcome. If your application gets denied, you might wonder whether you could have changed the outcome by lying on a single question — perhaps, by claiming that you earn more than $50,000 when you really don’t. If that lie would have flipped the outcome, computer scientists say that the Boolean function is “sensitive” to the value of that particular bit. If, say, there are seven different lies you could have told that would have each separately flipped the outcome, then for your loan profile, the sensitivity of the Boolean function is seven.

Computer scientists define the overall sensitivity of the Boolean function as the biggest sensitivity value when looking at all the different possible profiles. In some sense, this measure calculates how many of the questions are truly important in the most borderline cases — the applications that could most easily have swung the other way if they’d been ever so slightly different” (Klarreich, 2019). Please refer to Figure 4.

Taking the sensitivity factor into consideration in epidemiological decision-making, should make the decision tree more accurate and thus effective in stemming the spread of disease. We must determine, similarly, what questions are essential in those borderline cases, such as defining at what point a patient becomes symptomatic and whether a carrier who is not symptomatic is assigned the same set bit value. After all, a symptomatic patient might be coughing and sneezing, and therefore expelling particulates that substantially increase the likelihood of airborne transmission to another uninfected individual, whereas a carrier may infect only by direct contact. The two carriers are not equal and thus should be assigned different sensitivity values in the decision logic network.

In developing new computer code for an improved epidemiological early warning system, the sensitivity factor must be taken into consideration as it would for a hardware Boolean switch subject to bit-flip errors. The addition of data concerning ambient pathogens sampled that have been positively identified coupled with objective patient infection rates by the same pathogen supplied by healthcare facilities treating such patients, will yield rapid assessment of infectious virulence and thus the ability to effect rapid countermeasures to stem the spread of the disease. Objective data may be added to subjective data from sources such as social media for possible input correlations if desired.

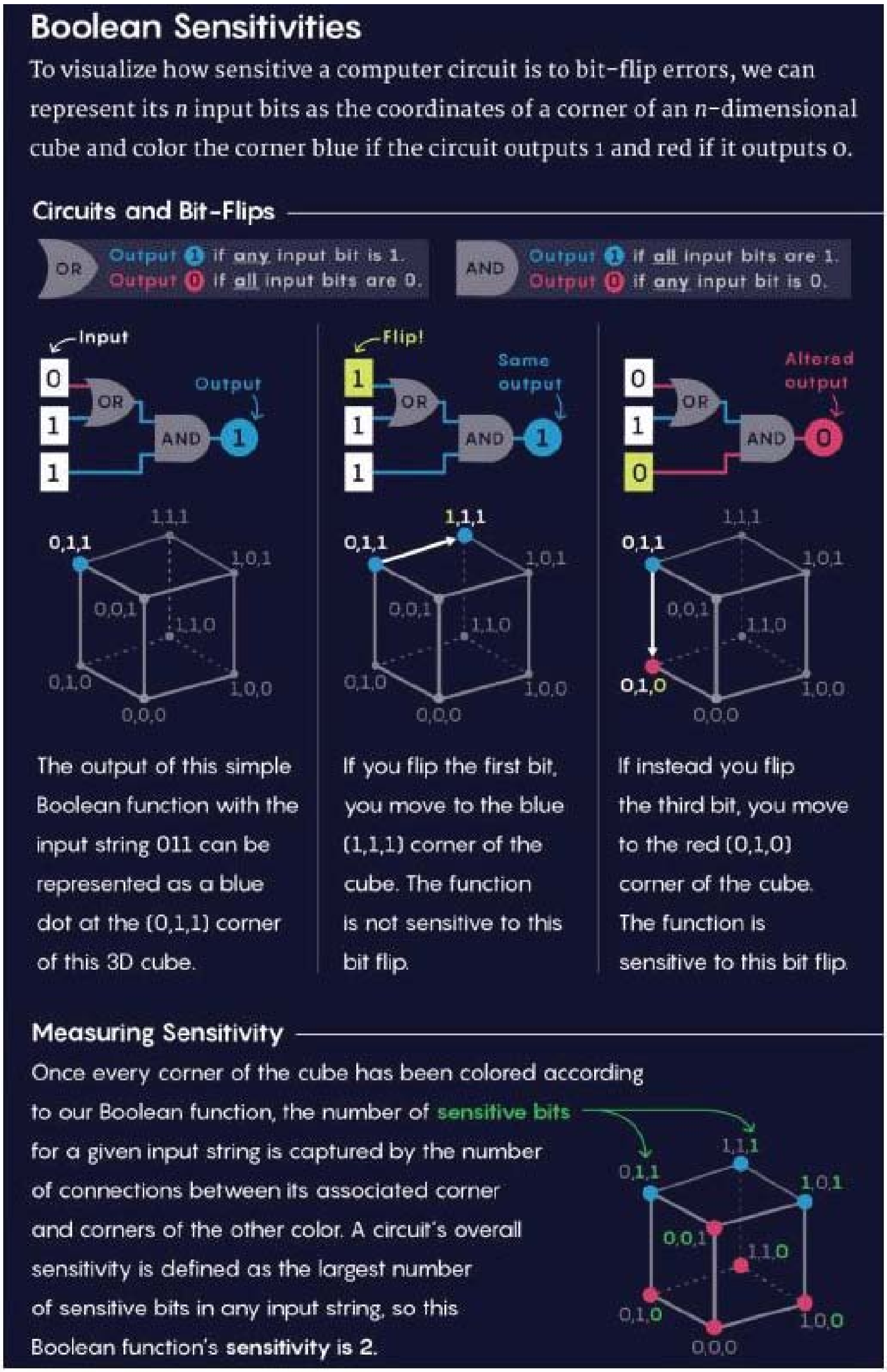


Figure 4

Concept of sensitivity in bit-flipping in a logic circuit influenced by sensitivity criteria

Live Agent Assessment

It has been indicated that in any airborne pathogen monitoring system, there needs to be provision for "live agent assessment". According to the CDC Ebola Response Team Incident Commander, there needs to be a means by which "to collect an aliquot which will NOT be inactivated or further processed – in order to have material readily at hand to assess for viable organism (whether bacterial or viral) in "real time" after a proteomic or nucleic acid assessment of material is completed – the absence of this capability would be a system deficit."

The proposed electrospray system would offer the capability to capture, concentrate, and retain live pathogen agents for subsequent collection and study in the laboratory. The essential system that will allow us to capture, charge, weigh, and collect viral species has already been constructed under NASA support for a future life finder mission (NASA Contract 80NSSC18C0056), shown in Figures 5 and 6. In this system, an electrospray source coupled to a vacuum system with charged species electrostatic acceleration and charge detection has been proven. A variant of the technology is being replicated in a separate system for capture and identification of nebulized viral particles introduced at the electrospray source. Note that the system shown is only an experimental research system. Production air samplers can be reduced to the size not much larger than a typical kitchen microwave oven.

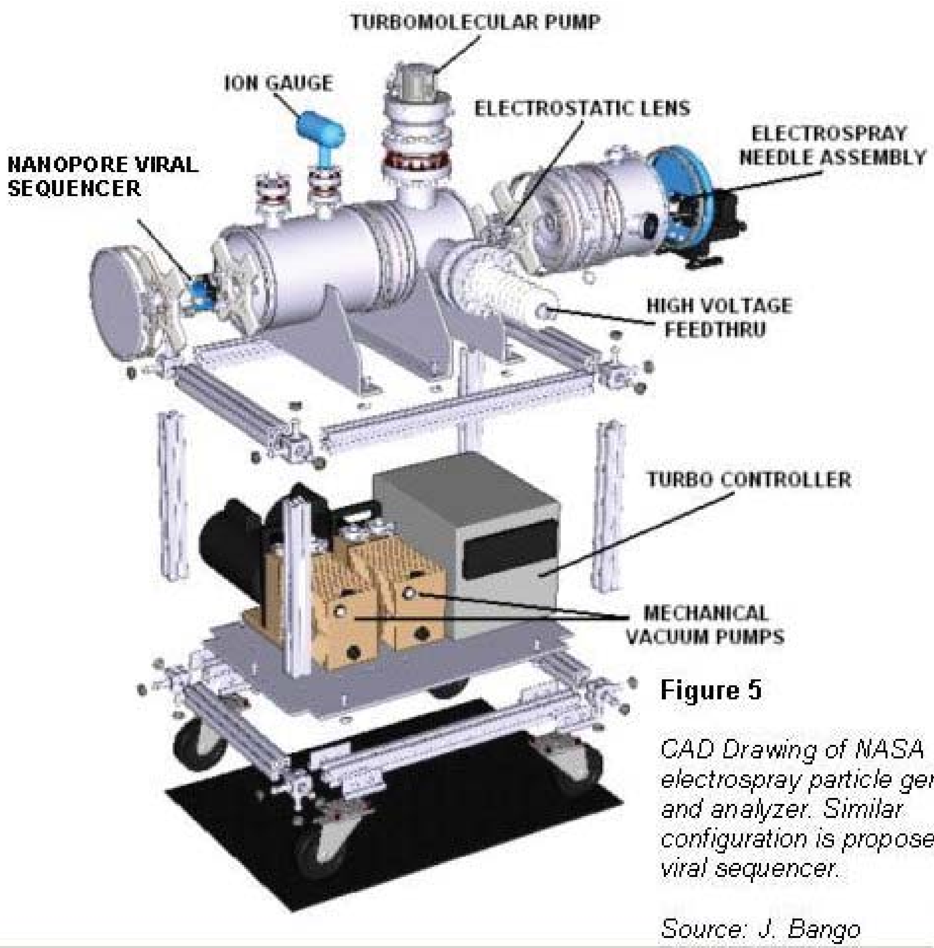


Figure 5

CAD Drawing of NASA electrospray particle generator and analyzer. Similar configuration is proposed for viral sequencer.

Source: J. Bango

Figure 6

Actual electrospray system as delivered to NASA. The viral sequencer will be similar in that the electrospray source at the right will be used to attract, capture, and ionize nebulized viral species. The viral 'ions' will be drawn into the partial pressure region of the system, and electrostatically channeled onto a solid-state nanopore plate upon which ionic liquid is used as the electrolyte. A patent has been applied for.

Source: J. Bango

DISCUSSION AND CONCLUSIONS

The key lacking factor in all present decision processes in infectious disease prediction and intervention, has been a lack of real time and near real time information about the type of viral pathogen, its concentration, and its distribution among target populations.

The development of a decision matrix that takes into consideration the "sensitivity" conjecture will improve any response to the beginning of an epidemic. In the spread of any airborne transmissible disease, the decision of when to implement an interventional protocol cannot be absolute or binary based on inputs such as a finite threshold of reported cases. Supporting information derived from sources such as social media can influence when and where hard genetic sequence data derived from infected patients and airborne samples should be concentrated. The proposed electrospray-nanopore sequencing-cloud combination, deployed worldwide with participating international healthcare agencies, promises to help address the predictive model deficit to yield improved public health interventional management in precluding the next epidemic or pandemic.

REFERENCES

Bango, J., & Dziekan, M. (2013, October 12). Electrospray Collection of Lunar Dust. Retrieved from https://www.techbriefs.com/component/content/article/tb/techbriefs/physical-sciences/12657

Bothner, B., & Siuzdak, G. (2004, March 5). Electrospray Ionization of a Whole Virus: Analyzing Mass ... Retrieved from https://masspec.scripps.edu/publications/public_pdf/100_art.pdf

Brauer, F., & Castillo-Chávez Carlos. (2012).*Mathematical models in population biology and epidemiology*. New York: Springer.

Bravata, D. M. (2002, June). Bioterrorism Preparedness and Response: Use of Information Technologies and Decision Support Systems: Summary. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK11887/

Bravata, D. M., Mcdonald, K. M., Smith, W. M., Rydzak, C., Szeto, H., Buckeridge, D. L., … Owens, D. K. (2004). Systematic Review: Surveillance Systems for Early Detection of Bioterrorism-Related Diseases.*Annals of Internal Medicine*,*140*(11), 910. doi: 10.7326/0003-4819-140-11-200406010-00013

Bravata, D. M., Sundaram, V., Mcdonald, K. M., Smith, W. M., Szeto, H., Schleinitz, M. D., & Owens, D. K. (2004). Evaluating Detection and Diagnostic Decision Support Systems for

Bioterrorism Response. *Emerging Infectious Diseases*,*10*(1), 100–108. doi: 10.3201/eid1001.030243

Oxford Nanopore. (2016, September 16).*Oxford Nanopore*.

Bushey, R. (2016, February 13). How DARPA Plans to Efficiently Battle Epidemics. Retrieved from https://medium.com/@R_Bushey/how-darpa-plans-to-efficiently-battle-epidemics-c183d0e79401

Clark, C. (2019, July 9). Emory mathematician to present a proof of the Sensitivity Conjecture. Retrieved from https://news.emory.edu/stories/2019/07/esc_sensitivity_conjecture_huang/campus.html

Foppa, I. M. (2017). W.O. Kermack and A.G. McKendrick: A seminal contribution to the mathematical theory of epidemics (1927).*A Historical Introduction to Mathematical Modeling of Infectious Diseases*, 59–87. doi: 10.1016/b978-0-12-802260-3.00004-3

Fuerstenau, S. (2003, November 1). Whole Virus Mass Analysis by Electrospray Ionization. Retrieved from https://www.jstage.jst.go.jp/article/massspec/51/1/51_1_50/_pdf/-char/en

Guha, S., Pease, L. F., Brorson, K. A., Tarlov, M. J., & Zachariah, M. R. (2011). Evaluation of electrospray differential mobility analysis for virus particle analysis: Potential applications for biomanufacturing. *Journal of Virological Methods*,*178*(1-2), 201–208. doi: 10.1016/j.jviromet.2011.09.012

Hay, S. I., Simba, M., Busolo, M., Noor, A. M., Guyatt, H. L., Ochola, S. A., & Snow, R. W. (2002). Defining and Detecting Malaria Epidemics in the Highlands of Western Kenya. *Emerging Infectious Diseases*,*8*(6), 555–562. doi: 10.3201/eid0806.010310

Izadi , M. T. (n.d.). Decision Theoretic Analysis of Improving Epidemic Detection. Retrieved from https://www.researchgate.net/publication/23162517_Decision_Theoretic_Analysis_of_Improving_Epidemic_Detection

Klarreich, E., & Quanta Magazine. (2019, July 25). Decades-Old Computer Science Conjecture Solved in Two Pages. Retrieved from https://www.quantamagazine.org/mathematician-solves-computer-science-conjecture-in-two-pages-20190725/

Liu, S., Poccia, S., Candan, K. S., Chowell, G., & Sapino, M. L. (2016). epiDMS: Data Management and Analytics for Decision-Making From Epidemic Spread Simulation Ensembles. *Journal of Infectious Diseases*,*214*(suppl 4). doi: 10.1093/infdis/jiw305

(2019, February 16). Lydia Bourouiba. Retrieved from https://cee.mit.edu/bourouiba

Pierson, E. E., Keifer, D. Z., Kukreja, A. A., Wang, J. C.-Y., Zlotnick, A., & Jarrold, M. F. (2016). Charge Detection Mass Spectrometry Identifies Preferred Non-Icosahedral Polymorphs in the Self-Assembly of Woodchuck Hepatitis Virus Capsids. *Journal of Molecular Biology*,*428*(2), 292–300. doi: 10.1016/j.jmb.2015.06.019

Probert, W. J. M., Jewell, C. P., Werkman, M., Fonnesbeck, C. J., Goto, Y., Runge, M. C., … Tildesley, M. J. (n.d.). Real-time decision-making during emergency disease outbreaks. Retrieved from https://doi.org/10.1371/journal.pcbi.1006202

Probert, W. J. M., Jewell, C. P., Werkman, M., Fonnesbeck, C. J., Goto, Y., Runge, M. C., … Tildesley, M. J. (2018). Real-time decision-making during emergency disease outbreaks. *PLOS Computational Biology*,*14*(7). doi: 10.1371/journal.pcbi.1006202

Quanta Magazine. (2019, July 26). Mathematician Solves Computer Science Conjecture in Two Pages. Retrieved from https://cacm.acm.org/careers/238399-mathematician-solves-computer-science-conjecture-in-two-pages/fulltext

Teich, J. M., Wagner, M. M., Mackenzie, C. F., & Schafer, K. O. (2002). The informatics response in disaster, terrorism, and war. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC344563/

(2018, May 21). The nanopore sequencing workflow. Retrieved from https://nanoporetech.com/how-it-works/nanopore-sequencing-workflow

(2018, May 18). Types of nanopores. Retrieved from https://nanoporetech.com/how-it-works/types-of-nanopores

Whiteman, H. (2014, April 10). Droplets from coughs and sneezes travel farther than you think. Retrieved from http://www.medicalnewstoday.com/articles/275309.php

DECISION SCIENCES INSTITUTE

5G MEC Offloading: Two QoE-based Strategies

Yang Lu
Old Dominion University
College of Charleston
Email: y1lu@odu.edu

ABSTRACT

5G MEC provides computing resources as services by 5G communication technologies. But, compared with conventional centralized cloud, MEC is a distributed resource with dynamic changes and conflicts between UEs (user equipment) and MEC servers. How to allocate MEC resources to meet users' requirements and to maintain QoE is an important issue in IT and across business disciplines. Offloading is an appropriate way to distribute MEC resources to achieve high levels of utilization and effectiveness. This study proposes two mechanisms that can work under complex conditions, such as multiple UEs' offloading requests, low latency, low energy consumption, and task consistency.

KEYWORDS: Mobile edge computing (MEC), Offloading, 5G MEC, MEC deployment scheme, Resource allocation, Quality of Experience (QoE)

INTRODUCTION

The development of mobile communication technologies and the adaptation of intelligent devices, various network services and applications are quickly emerging, and end-users are increasingly demanding high network performance, such as broad bandwidth, low latency, availability, reliability, and security. Although the processing power of the CPU (the central processing unit) of the new mobile device continue to strengthen, it still cannot process requests based on massive data in a limited time. In addition, the huge consumption of battery impedes the processing of applications by local servers. These issues affect the performance and the QoE (quality of experience) of the services on the UEs (user equipment). 5G MEC offloading has the potential to solve the problems and to provide high-performance services to end users (Mao, et al., 2017a; Guo, Liu & Zhang, 2018; Mach & Becvar, 2017).

MEC (Mobile edge computing) refers to the deployment of computing and storage resources at the edge of the mobile network, which provide IT services and cloud computing capabilities for mobile networks, and which provide end-users with low (no) latency and with high-performance service solutions. MEC is a critical factor improving the QoE of 5G network. As one of the key technologies in the MEC, task offloading refers to the transmission of part or all of the tasks of a UE with limited computing power to a cloud server through a network. Offloading allows a UE to extend functionality to the cloud and leverages the cloud's powerful computing power to expand its computing capacity, to decrease execution latency, to extend battery life, and to provide a high QoE for end users. The offloading technology consists of three aspects: a decision on computation offloading, the allocation of computing resources, and system implementation. In particular, a decision on computation offloading solves the problem of what resources the UE can offload, how to offload tasks, and how many resources need to be offloaded. Allocation of computing resources focuses on how to appropriately allocate resources that can be offloaded among entities (cloud, MEC, and UEs). System implementation deals with the implementation strategy for achieving offloading resources on the MEC platform (Abbas, et al., 2018; Mao, et al., 2017b; Pan & McElhannon, 2018).

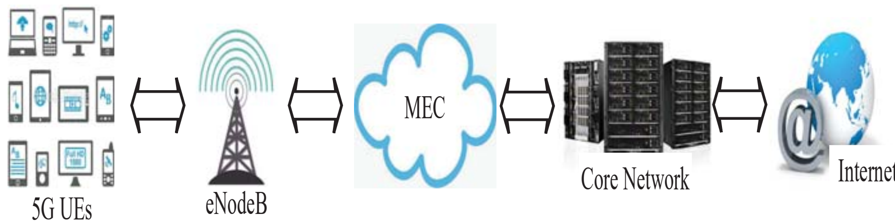
5G MEC offers the advantages of scalability, flexibility, mobility, virtualization, low cost and no terminal limitations, and is widely used in data sensing, medical and health industries, social networking, multimedia searches, and many other fields. It significantly improves the computing processing capability and the service quality in various fields, showing good development prospects and promising social benefits. The MEC offloading can transmit files to the MEC server that is closest to the UE. It not only reduces the network workload, but it solves the problems of energy consumption and transmission costs. Offloading also helps in the development of emerging technologies with zero latency. For example, autonomous vehicles need to sense road conditions, obstacles, and the driving information of surrounding vehicles in real time. Faster transmission and more accurate analysis of the related massive data and calculation can be realized through 5G MEC offloading (Wang, C. et al., 2017a; Liu, et al., 2016; Tao, et al., 2017).

The paper is outlined as follows. Section II depicts an overview of current research on MEC offloading. The two offloading mechanisms are described and explained in Section III. Section IV addresses research challenges and future trends toward 5G MEC offloading. The conclusion is in Section V.

CURRENT RESEARCH

The UE offloads partial of the computing task to the mobile edge cloud server deployed by the network operator near the base station by communicating with the nearby deployed base station (eNodeB). High-speed transmission of fiber enables rapid cross-response and low-latency connections. The MEC provides the UE with a wide range of caching, calculating and services. Upon the deployment of 5G MEC, it can provide session and business continuity, QoE and pricing, and support for MEC local networks. The MEC is also flexible in the specific deployment modes of the 5G network: centralization and distribution. Centralized deployment supports enhanced gateway capabilities with UE, and the distributed deployment allocates services in different locations. This hierarchical placement of resources in the network makes network management more flexible and dynamic (Muhammad, et al., 2018; Shi, et al., 2016; Li, J. et al., 2018; Zheng, et al., 2016). The following figure (Figure 1) illustrates a complete network of 5G MEC that includes 5G UEs (end users), eNodeB (the edge base station), the MEC server, the core network, and the Internet. eNodeB is the connection bridge between the UEs and the MEC servers. The edge computing server is deployed in the wireless access network, which greatly reduces the distance from the UE. Due to the reduced transmission distance, the task migration of the 5G MEC no longer needs to go through the long backhaul link and the core network, thereby reducing the delay overhead. On the other hand, since the computing capability of the edge server is greater than that of the mobile device, the task processing time is greatly shortened.

Figure 1. 5G MEC Typology



In an offloading process, the UE typically consists of a code parser, a system parser, and a decision engine. The execution of an offloading decision is divided into three steps. First, the code parser determines what can be offloaded, and the specific task offloading depends on the application type and the code and data partitioning. Then, the system parser is responsible for monitoring various parameters, such as the available bandwidth, the size of the offloading file, transmission costs, and energy consumption. Finally, the decision engine decides whether to offload, based on the evaluations. The offloading decision is mainly divided into three categories: the low latency offloading mechanism, the energy efficiency offloading mechanism, and the balancing latency and consumption offloading mechanism. For example, shortening the latency of offloading process, decreasing the energy consumption and costs, and mitigating the execution failure of offloading process. Game theory algorithms and the MDP (Markov Decision Process) are popular models used to estimate potential offloading mechanisms (Howard, 1971; Puterman, 2015; Zhou, et al., 2019; Chen, et al., 2013).

The decision to migrate tasks is the result of many factors, such as end-user, networks, mobile devices, servers and applications. End-user needs to consider offloading costs and QoE. Network mainly focuses on Wi-Fi and 5G. Processing capability, memory and storage are factors that affect mobile devices. Whether the edge server has enough resources to fulfill the offloading tasks and the environment in which the application is running appropriately directly affect the decision of offloading. For the application itself, the higher the computational complexity of the application and the smaller the amount of data transferred, the greater the likelihood of offloading of the tasks.

Among the extant offloading strategies, some are energy-oriented decision methods, some are corresponding time and energy decision methods, some are overall mobile terminals as migration targets, and some are converting mobile terminal applications into multiple partitions that are the smallest unit of offloading.

Low Latency Offloading Mechanism

If a task is executed locally, the elapsed time is the time during which the application performs the task. If the task is offloaded to the MEC, the time spent will include three parts: the time at which the file to be offloaded is transferred to the MEC, the time of the task required to receive file from the MEC, and the time of getting the file back from the MEC to the UE. Therefore, the delay caused by offloading the computing tasks to the MEC affects the QoE. Although a model of considering multiple UEs leads to NP-hardness, a low latency offloading strategy is still a practical approach to the allocation of resources in the MEC to multiple UEs. In order to ensure QoE, many studies, aimed at reducing delays, involve different optimization algorithms and application scenarios. Examples include constructing a model of NP-hardness to resolve offloading for multiple UEs, or the Lyapunov optimal dynamic offloading mechanism associated with energy harvesting technologies. Low latency is a key factor that affects offloading energy consumption and user's QoE. Service with low latency is more attractive to end-users, although they may pay extra for the energy consumption and other costs (Armbrust, et al., 2012; Kwak, et al., 2015; Mao, Zhang & Letaief, 2016; Kao, et al., 2017; Wang, C. et al., 2017b).

Energy Efficiency Offloading Mechanism

The energy consumed for offload tasks to the MEC server consists of two aspects. One is the transmission energy that transfers the offloading file to the MEC, and the other is the energy consumed by the data returned to the UEs. One study used the TDMA (time-division multiple access) system to divide time slots. In each time slot, the UE offloads its file to the MEC based on channel quality and local energy consumption. The optimal offloading strategy was proposed for each UE. If the UE has a priority above a given threshold, the UE completely offloads the computing task to the MEC. Conversely, if the UE has a lower priority than the threshold, only some of the tasks are offloaded to satisfy the delay constraint. For UEs that do not meet the application latency constraints, tasks will be implemented locally. Furthermore, the TDMA mechanism has been extended to the OFDMA-based offloading scheme (orthogonal frequency-division multiple access); this can reduce energy consumption by 90% (You & Huang, 2016; You, et al., 2016). Another study (Zhang, et al., 2016) attempted to minimize offloading transmission and radio energy consumption by optimization algorithm with latency constraints. Offloading in small-cell network is also worth of examining from task and transmission perspectives (Yang, et al., 2018; Munoz, P-Iserte & Vidal, 2015). Cloud computing is scalable and adaptable, it is potential to extend the mobile edge computing offloading mechanisms to other cloud-related systems, such as multi-access edge computing and cloud of things (El Haber, Nguyen & Assi, 2019; Nan, et al., 2017; Guo, et al., 2018).

Balancing Offloading Mechanism

When performing complex computing tasks, such as face recognition systems or real-time video systems, both energy consumption and delay can affect the QoE; a focusing only on one point cannot achieve UEs' requirements. So knowing how to achieve a leverage between energy consumption and latency when performing offloading tasks is crucial to build a balanced strategy to more efficiently allocate resources in a virtuous cycle.

The following parameters are considered in the offloading process as a trade-off analysis: the total amount of data to be processed, the computing power of the UE and the MEC, the channel state between the UE and the SCeNB (the intermediate base station connecting the UE and the MEC), the energy consumption of the UE, waiting time, and executing time, and other related costs (Kao, et al., 2017; You & Huang, 2016). Studies seek to construct the offloading algorithms through different combinations to optimize offloading processes. For example, You, et al. (2016) used Gurobi optimizer to analyze the offloading problem in the static environment and extend the condition to dynamic phenomenon.

5G MEC OFFLOADING MECHANISMS

Due to the diversity of offloading factors and metrics, it is not easy to design an appropriate offloading strategy and objective assessment of offloading performance. In this section, two offloading strategies will be introduced and generalized: the optimal offloading and the consistency offloading mechanisms. The optimal strategy analyzes various costs and consumption of the complete offloading process and seeks the optimal mechanism for each of the UEs to accomplish overall high utilization. The consistency strategy solves the issue of VM migration from one MEC to another, in order to maintain both task consistency and a high QoE with low latency. Both of the mechanisms are capable of handling malicious requests from multiple UEs (Ketykó, et al., 2016; Ndikumana, et al., 2017).

Optimal Offloading Mechanism

The first approach is an optimal offloading mechanism that adopts game theory to allocate limited edge resources to multiple UEs in order to benefit both end-users for maximum benefit and providers for the highest resource utilization. The route between the UEs and the edge servers is random. This leads to different transmission energy consumption, channel interference, and overall costs. If the edge center near the base station receives too many requests and uses the wireless channel to schedule resources, it can cause severe mutual interference. Hence, it is important to examine whether a request needs to be offloaded and how to appropriately allocate resources.

Model

Suppose there are *N* (= {1, 2, …, N}) UEs. Each one has a task that requires computing resources and can be selectively sent to close base stations (eNodeBs). eNodeB connected to the edge server can process the UE's request. The provider determines the communication speed between the UE and the eNodeB, such as file transmission channel, the uplink transmission rate, and the downlink transmission rate. There are *M* (= {1, 2, …, M}) wireless channels between the UEs and the eNodeBs.O_n is the offloading strategy of user *n*. $O_n = 0$ means that the UE will run the task locally, $O_n > 0$ means that the UE will offload the task to the edge. O_{1-N} (= $(O_1, O_2, \dots, O_N)$) represents the offloading strategies of all UEs.

If the total cost of MEC is less than or equal to the total cost of local computing, it is defined as effective edge cloud calculation, as in the following equation

$$C_n^E(O_{1-N}) \le C_n^L \tag{1-1}$$

C_n^L represents the total cost of local computing, and $C_n^E(O_{1-N})$ represents the total cost of MEC. For UEs, the optimal approach assists the UE in rationally selecting an offloading strategy with a high QoE. For providers, the optimal approach achieves higher resource utilization and revenues.

The above function can be described as an optimization problem. It is used to calculate the maximum number of UEs that can obtain MEC resources among all of the UEs. The mathematical expressions are shown in equation (1-2). R_O is 1 when A is true, and 0 otherwise.

$$\begin{gathered} max \sum_{n \in N} R_{\{O_n > 0\}} \\ C_n^E(O_{1-N}) \le C_n^L \\ R_O \in \{0, 1, \dots, \mathrm{M}\} \end{gathered} \tag{1-2}$$

This is an NP-Hardness issue, similar to the biggest packaging problem. Assuming another complete system cost is I_n, the target problem is transformed into a combinatorial optimization problem in a multidimensional discrete space environment, i.e., minimizing I_n. As shown in equation (1-3)

$$\min \sum_{n \in N} I_n$$
$$R_O \in \{0, 1, \ldots, M\} \quad (1\text{-}3)$$

The general expression is:

$$\Gamma_{MEC} = (N, \{O_n\}_{n\in N}, \{I_n\}_{n\in N}) \quad (1\text{-}4)$$

On the condition of

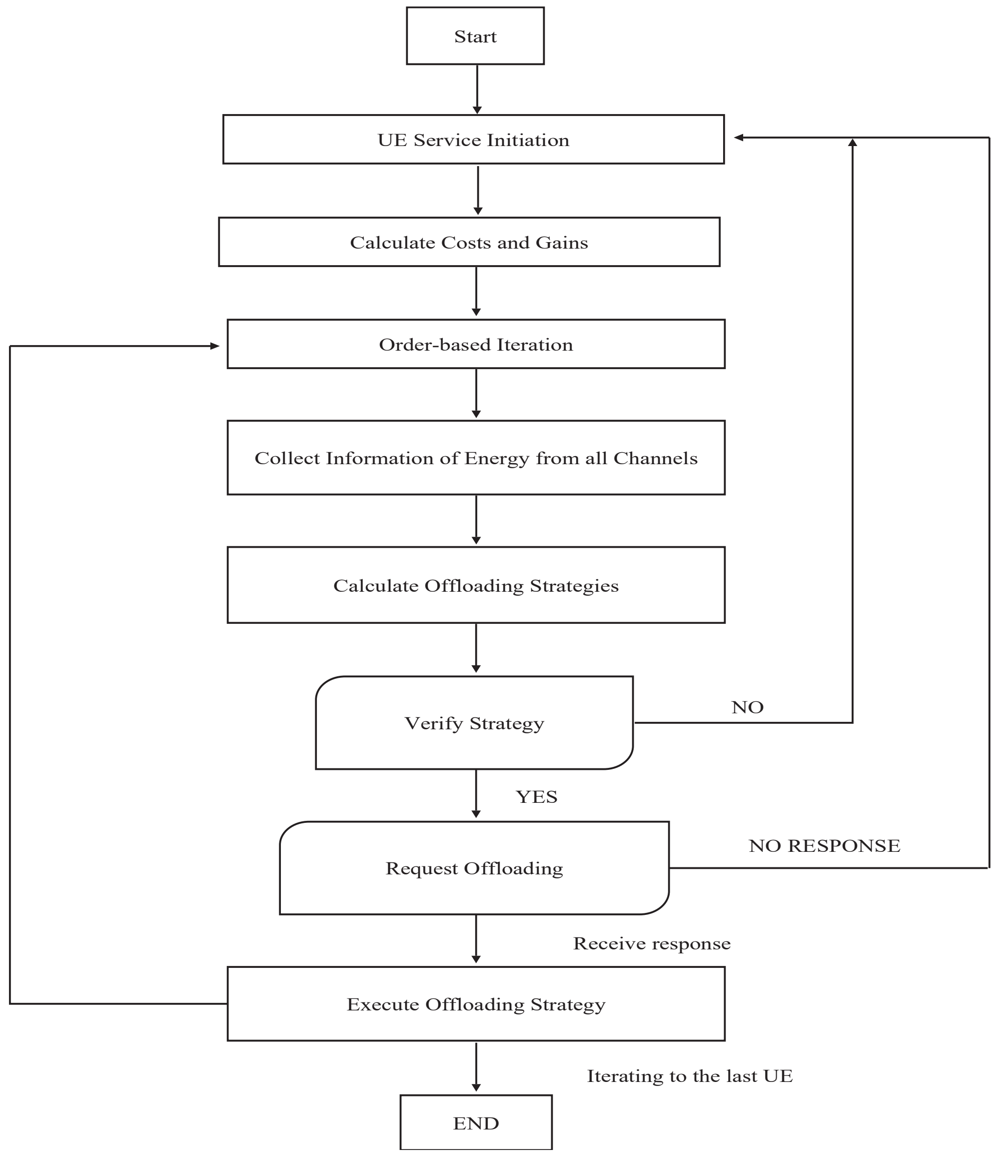
$$I_n(O_n^*, O_{N-n}^*) \leq I_n(O_n, O_{N-n}^*) \quad (1\text{-}5)$$

where O^* (= $(O_1^*, O_2^*, \ldots, O_N^*)$) is the set of all potential UEs' offloading strategies when MEC offloading is under the Nash Equilibrium.

Offloading Process

This section will depict the complete iterative processes in detail (Figure 2). In total, the process consists of ten steps, from beginning to end. The iteration will be terminated until all of the requested tasks are examined and assigned for potential optimal offloading strategies.

Figure 2. The Iteration Process



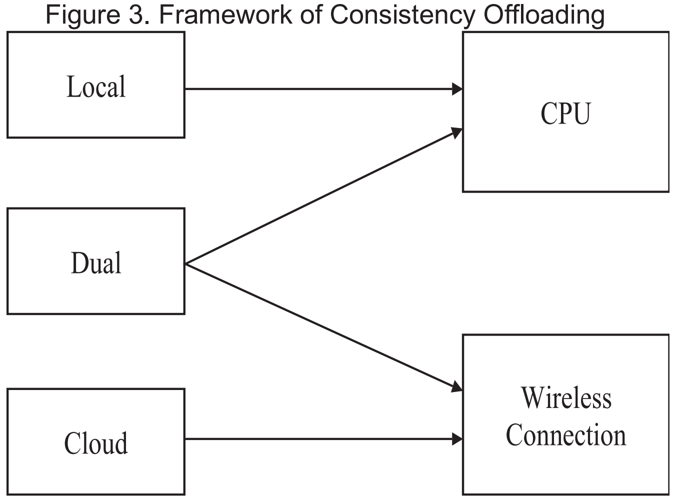
First is the initiation of every UE locally.
Second is the calculation of channel gain and the transmission of each UE based on the random distance between the UEs and the BS.
Third is the ranking of UE based on its power; the high channel power that the UE preferentially selects in the iteration process.
Fourth is the transferring of energy on selected channels and the collecting of energy information from all channels.
Fifth is the calculation of all of the potential offloading strategies and the examination of the related costs and consumption.
Sixth is verifying the strategy. If an offloading strategy is better than the local implementation, continue to the next step; otherwise, it will go back to the initiation for the next UE's decision making.
Seventh, the UE sends an offloading request to the related servers.

Eighth, if the UE receives a response from the servers, it will implement an offloading strategy.
Ninth, if the UE doesn't receive a response from servers, the local implementation will continue. The iteration will go back to the initiation for the next UE.
Tenth, the complete iteration will proceed from the first UE until the last, and then it will end. This approach is effective if all of the users are in a static environment; in that case, the UEs won't change the offloading decisions during the whole process. However, in practice, 5G users move frequently, and tasks relate to multiple MEC servers. In this circumstance, it is necessary to consider the VM migration, in order to maintain the continuity and the consistency of services.

Task Consistency Offloading Mechanism

By processing offloading technology, VM (virtual machine) migration can ensure the continuity and consistency of services. VM migration means that VMs running on the current node will migrate to another more suitable node(s) with higher utilization and lower costs (Wang, et al., 2015). Much research has focused on the offloading strategies from one node to another node. But, in practice, this phenomenon often occurs: VMs need to migrate to multiple nodes in order to achieve the optimal performance of QoE. Thus, the conditions of consistency offloading are complicated and dynamic. In this study, we present a generalized model, which can assist practitioners or scholars to implement offloading mechanisms through various perspectives (Kwak, et al., 2015).

Figure 3. Framework of Consistency Offloading



In 5G MEC offloading, tasks can be divided into three categories: 1) tasks that are executed locally in UE, 2) tasks that are executed in the cloud, and 3) tasks that can be executed either in local UE or in the cloud (Dual). Since it aims at optimizing resource utilization with the tolerated delay, this solution is designed to determine whether the VMs (virtual machine) need to migrate, and how to allocate limited resources after VM migration. It uses MDP (the Markov Decision Process) to evaluate which way is better to maintain task consistency and to guarantee QoE with potential low costs (Wang, et al., 2016; Ding, Fan & Poor, 2019; Jia, Cao & Yang, 2014). According to Figure 3, three factors that directly impact the final offloading strategy: processing costs, processing channel, and processing order.

<u>Processing Costs</u>

Low latency applications, such as IoT (Internet of Things) and the connected terminals, require high reliability and low end-to-end delay (millisecond) communications. To support low latency, virtual machines and data source are offloaded as users move from one MEC to another. The offloading process may have a negative impact on latency. Hence, it is possible to consider a high-speed path with a small delay in the backhaul link, meanwhile, the transmitted file needs to be compressed and the virtual machine recovery process needs to be simplified. In general, transmission, energy, waiting, and executing costs are the major factors that impact the overall utilization in 5G MEC offloading (Min, et al., 2019; Xu, He &Li, 2014; Lu & Xu, 2018; Lyu, et al., 2018; Kherraf, et al., 2019; Zhang, et al., 2017).

Processing Channel

In 5G MEC offloading, there exist two major channels. If tasks are implemented locally, the processing channel is the CPU. If tasks are implemented in the MEC, the wireless transmission channel is the cellular network or the WLAN (wireless local area network). In addition, some task may be allocated to a combined channel, such as CPU and WLAN, CPU and cellular, and cellular and WLAN. Based on the MDP, the particular channel with minimum costs and optimal utilization of the resources will be the offloading strategy (Mach & Becvar, 2017; Mao, et al., 2017b).

Processing Order

In the 5G MEC offloading process, there exist tasks with malicious requirements that come from different UEs. How to effectively fulfill tasks in the MEC, by certain orders, that impact the overall performance of services and resource utilization is the question under consideration. For sequential tasks, the optimal offloading strategy is suitable for processing tasks. For concurrent tasks, load balancing heuristics are used to offload tasks to the MEC. For the partial offloading model, the influence of inter-task dependencies is proposed, and the polynomial time algorithm is used to solve the optimal solution of offloading (Guo, Liu & Zhang, 2018; Kao, et al., 2017; Ding, Fan & Poor, 2019).

The following model is a general MDP that examines the minimum costs and optimal utilization. The goal is to minimize the total costs of a specific task:

$$min \sum \gamma^t C \tag{2-1}$$

where γ^t $(0 \leq \gamma < 1)$ is the discount, and C represents all possible costs during the offloading process. For example, transmission costs, energy consumption, waiting time, service time, etc. As MDP, we have two main algorithms: policy ($\pi(s)$) and value ($V(s)$).

$$\pi(s) = arg \min_a \{\sum_{s'} P(s' \mid s, a)[C(s' \mid s, a) + \gamma V(s')]\} \tag{2-2}$$

$$V(s) = \sum_{s'} P_{\pi(s)}(s, s')[C_{\pi(s)}(s, s') + \gamma V_{s'}] \tag{2-3}$$

There are the state transition *P*, cost function *C*, and the probability ($P_{\pi(s)}(s, s')$) that action $\pi(s)$ in state *s* will lead to state s' for an MDP. We seek the policy ($\pi(s)$) that minimize the discounted costs. Specifically, a represents an action, and *s* and s' represent certain states. The value $V(s)$ is the actual value of action *a* in state s. During the iterative process, these two steps are repeated until there is no change to calculate the strategy with the minimum cost.

CHALLENGES AND FUTURE TRENDS

In 5G networks, according to business needs, MEC can be flexibly layered to maximize resources utilization and to reduce the computational costs and energy consumption with low latency. However, the offloading of 5G MEC still faces problems, such as mobility management, security and interference control.

Mobility Management

In mobility management, in order to complete the offloading of the corresponding tasks, it is necessary to consider low latency and path prediction, in order to achieve high QoE communication.

Low Latency

Low latency applications, such as IoT (Internet of Things) and Vehicular Network, require high reliability and low end-to-end delay (millisecond) communications. To support low latency, virtual machines and data source are offloaded as users move from one MEC zone to another. The offloading process may have a negative impact on application latency. MEC systems require offloading with short time. Hence, it is possible to consider a high-speed path with a small delay in the backhaul link, meanwhile, the transmitted file needs to be compressed and the virtual machine recovery process needs to be simplified (Wang, S. et al., 2017a; Zheng, et al., 2016; Munoz, P-Iserte & Vidal, 2015).

Path Prediction

The key to mobility management is the offloading of virtual machines and file. Traditional MEC offloading only transfer computing tasks to another server when transfer is happening. Inappropriate data offloading will lead to high latency and will increase the MEC network load. The solution is to analyze user tracks when the MEC provides services to users, and to predict the next MEC that the user will arrive at, in order to transfer the data and resources to the new MEC, in advance. But this technology has two main challenges. The first is track prediction. An accurate prediction can achieve seamless switching between MEC servers and can reduce prefetch redundancy. It requires precise modeling and high-complexity machine learning techniques (Li, S. et al., 2018; Xu, Chen & Ren, 2017; Lu, 2019). The second challenge is how to access the data that needs to be delivered in advance. Inaccurate prediction will lead to unnecessary waste. Balancing the amount of data transmitted and the accuracy of the prediction is challenging. In addition, VM migration leads to a heavy burden on the backhaul link and to high latency. It is necessary to implement a technology that can quickly migrate VMs within a short time (e.g., milliseconds). Hence, if VMs can be migrated in advance, it will better solve both heavy load and high latency (Mao, et al., 2017a).

Security Management

Security is a technical challenge in cloud computing offloading. Since 5G MEC is a distributed system, single point is so weak that the attack of a single point may lead to a destruction of the entire system. Multi-tenant mode will cause malicious users to sneak into the network to exploit cloud platform vulnerabilities in order to attack the network. In addition, the open-source software is vulnerable to attacks as well. File that is offloaded to the cloud or to the edge can also be easily attacked or tampered with. Many of the security solutions originally used for cloud computing are no longer suitable for 5G MEC offloading. The security issues are distributed at various levels, including edge node security, network security, data application security, security situation awareness, security management and coordination, and identity management. Security issues in 5G MEC cannot be fully solved by conventional cloud resolution, due to its unique and complex distributed network. Therefore, it is important to design appropriate mechanism to deal with malicious safety issues (Shibin & Kathrine, 2017; Xiao, et al., 2018; Roman, Lopez & Mambo, 2016; Xu, et al., 2019; Chaudhary, Kumar & Zeadally, 2017).

Interference Management

Interference is also one of the key issues that need to be solved during the offloading process. If many applications of UEs are simultaneously offloaded to the MEC server, interference will occur. How to allocate resources with the guaranteed QoE while to mitigating interference is a critical question in 5G MEC offloading. Interference management has multiple implementation methods and is closely related to resource management. The nature of interference is the conflicting use of resources. Unreasonable allocation of network resources is the root of interference. In a distributed 5G MEC network, the large quantity of offloading requests of UEs and the complex network environment can reduce the overall resource utilization. Effective resource allocation is an important means for the management of interference. On the one hand, it can increase network capacity by reasonable use of network resources. On the other hand, it can correct resource allocation strategies and can increase network capacity through interference management (Bu, Yu & Yanikomeroglu, 2015; Li, et al., 2015; Wang, S. et al., 2017b; Huang & Li, 2016).

CONCLUSION

5G MEC is a cloud service platform that runs on the edge network. It can improve both service performance and user experience by deploying business processing and resource scheduling to the cloud service platform. 5G MEC offloads the services and capabilities originally in the cloud center to the edge network. By distributing computing, storage, communication resources at the edge, 5G MEC can effectively decrease network overload, can shorten service latency, can save energy consumption and other related costs, and can guarantee task consistency.
In this study, based on 5G MEC offloading, we present two potential robust and efficient offloading approaches to distribute services and to fulfill UE’s requests: an optimal offloading strategy and a task consistency offloading strategy. It is a good attempt to examine and to allocate cloud resources and offloading tasks from different aspects: low latency, energy efficiency, low costs and consumption, quick executing time, and waiting time.

REFERENCES

Abbas, N., Zhang, Y., Taherkordi, A., & Skeie, T. (2018). Mobile edge computing: A survey. *IEEE Internet of Things Journal*, *5*(1), 450-465.

Armbrust, et al. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50–58.

Bu, S., Yu, F. R., & Yanikomeroglu, H. (2015). Interference-aware energy-efficient resource allocation for OFDMA-based heterogeneous networks with incomplete channel state information. *IEEE Transactions on Vehicular Technology*, *64*(3), 1036-1050.
Chaudhary, R., Kumar, N., & Zeadally, S. (2017). Network service chaining in fog and cloud computing for the 5G environment: data management and security challenges. *IEEE Communications Magazine*, *55*(11), 114-122.
Chen, M., Liew, S. C., Shao, Z., & Kai, C. (2013). Markov approximation for combinatorial network optimization. *IEEE transactions on information theory*, *59*(10), 6301-6327.
Ding, Z., Fan, P., & Poor, H. V. (2019). Impact of non-orthogonal multiple access on the offloading of mobile edge computing. *IEEE Transactions on Communications*, *67*(1), 375-390.
El Haber, E., Nguyen, T. M., & Assi, C. (2019). Joint Optimization of Computational Cost and Devices Energy for Task Offloading in Multi-tier Edge-clouds. *IEEE Transactions on Communications*.
Guo, H., Liu, J., & Zhang, J. (2018). Computation Offloading for Multi-Access Mobile Edge Computing in Ultra-Dense Networks. *IEEE Communications Magazine*, *56*(8), 14-19.
Guo, K., Yang, M., Zhang, Y., & Ji, Y. (2018, March). An Efficient Dynamic Offloading Approach based on Optimization Technique for Mobile Edge Computing. In *2018 6th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)* (pp. 29-36). IEEE.
Howard, R. A. (1971). *Dynamic Probabilistic Systems*. Volume II: Semi-Markov and Decision Processes. Wiley Interscience.
Huang, G., & Li, J. (2016). Interference mitigation for femtocell networks via adaptive frequency reuse. *IEEE Transactions on Vehicular Technology*, *65*(4), 2413-2423.
Jia, M., Cao, J., & Yang, L. (2014, April). Heuristic offloading of concurrent tasks for computation-intensive applications in mobile cloud computing. In *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*(pp. 352-357). IEEE.
Kao, Y. H., Krishnamachari, B., Ra, M. R., & Bai, F. (2017). Hermes: Latency optimal task assignment for resource-constrained mobile computing. *IEEE Transactions on Mobile Computing*, *16*(11), 3056-3069.
Ketykó, I., Kecskés, L., Nemes, C., & Farkas, L. (2016, June). Multi-user computation offloading as multiple knapsack problem for 5G mobile edge computing. In *2016 European Conference on Networks and Communications (EuCNC)* (pp. 225-229). IEEE.
Kherraf, N., Alameddine, H. A., Sharafeddine, S., Assi, C., & Ghrayeb, A. (2019). Optimized Provisioning of Edge Computing Resources with Heterogeneous Workload in IoT Networks. *IEEE Transactions on Network and Service Management*.
Kwak, J., Kim, Y., Lee, J., & Chong, S. (2015). DREAM: Dynamic resource and task allocation for energy minimization in mobile cloud systems. *IEEE Journal on Selected Areas in Communications*, *33*(12), 2510-2523.
Li, C., Zhang, J., Haenggi, M., & Letaief, K. B. (2015). User-centric intercell interference nulling for downlink small cell networks. *IEEE Transactions on Communications*, *63*(4), 1419-1431.
Li, D. X., He, W., & Li, S. (2014). Internet of things in industries: A survey. *IEEE Transactions on industrial informatics*, *10*(4), 2233-2243.
Li, J., Gao, H., Lv, T., & Lu, Y. (2018). Deep reinforcement learning based computation offloading and resource allocation for MEC. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1-6). IEEE.

Li, S., Zhang, N., Lin, S., Kong, L., Katangur, A., Khan, M. K., ... & Zhu, G. (2018). Joint admission control and resource allocation in edge computing for internet of things. *IEEE Network*, *32*(1), 72-79.
Liu, J., Mao, Y., Zhang, J., & Letaief, K. B. (2016, July). Delay-optimal computation task scheduling for mobile-edge computing systems. In *2016 IEEE International Symposium on Information Theory (ISIT)* (pp. 1451-1455). IEEE.
Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, *6*(1), 1-29.
Lu, Y., & Li, D. X. (2018). Internet of Things (IoT) cybersecurity research: a review of current research topics. *IEEE Internet of Things Journal*. Early Access.
Lyu, X., Tian, H., Jiang, L., Vinel, A., Maharjan, S., Gjessing, S., & Zhang, Y. (2018). Selective offloading in mobile edge computing for the green Internet of Things. *IEEE Network*, *32*(1), 54-60.
Mach, P., & Becvar, Z. (2017). Mobile edge computing: A survey on architecture and computation offloading. *IEEE Communications Surveys & Tutorials*, *19*(3), 1628-1656.
Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys & Tutorials*, *19*(4), 2322-2358.
Mao, Y., Zhang, J., & Letaief, K. B. (2016). Dynamic computation offloading for mobile-edge computing with energy harvesting devices. *IEEE Journal on Selected Areas in Communications*, *34*(12), 3590-3605.
Mao, Y., Zhang, J., Song, S. H., & Letaief, K. B. (2017). Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems. *IEEE Transactions on Wireless Communications*, *16*(9), 5994-6009.
Min, M., Xiao, L., Chen, Y., Cheng, P., Wu, D., & Zhuang, W. (2019). Learning-based computation offloading for IoT devices with energy harvesting. *IEEE Transactions on Vehicular Technology*, *68*(2), 1930-1941.
Muhammad, S., Shen, Z., Qi, J., & Zhang, G. (2018, October). Power-Delay Tradeoff in Mobile-Edge Computation Offloading with Heterogeneous Applications. In *2018 IEEE 18th International Conference on Communication Technology (ICCT)* (pp. 227-231). IEEE.
Munoz, O., Pascual-Iserte, A., & Vidal, J. (2015). Optimization of radio and computational resources for energy efficiency in latency-constrained application offloading. *IEEE Transactions on Vehicular Technology*, *64*(10), 4738-4755.
Nan, Y., Li, W., Bao, W., Delicato, F. C., Pires, P. F., Dou, Y., & Zomaya, A. Y. (2017). Adaptive energy-aware computation offloading for cloud of things systems. *IEEE Access*, *5*, 23947-23957.
Ndikumana, A., Ullah, S., LeAnh, T., Tran, N. H., & Hong, C. S. (2017, September). Collaborative cache allocation and computation offloading in mobile edge computing. In *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)* (pp. 366-369). IEEE.
Pan, J., & McElhannon, J. (2018). Future edge cloud and edge computing for internet of things applications. *IEEE Internet of Things Journal*, *5*(1), 439-449.
Puterman, M. L. (2015). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley.
Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, *78*, 680-698.
Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637-646.

Shibin, D., & Kathrine, G. J. W. (2017, February). A comprehensive overview on secure offloading in mobile cloud computing. In *2017 4th International Conference on Electronics and Communication Systems (ICECS)* (pp. 121-124). IEEE.
Tao, X., Ota, K., Dong, M., Qi, H., & Li, K. (2017). Performance guaranteed computation offloading for mobile-edge cloud computing. *IEEE Wireless Communications Letters*, *6*(6), 774-777.
Truong-Huu, T., Tham, C. K., & Niyato, D. (2014, December). To offload or to wait: An opportunistic offloading algorithm for parallel tasks in a mobile cloud. In *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*(pp. 182-189). IEEE.
Wang, C., Liang, C., Yu, F. R., Chen, Q., & Tang, L. (2017a). Computation offloading and resource allocation in wireless cellular networks with mobile edge computing. *IEEE Transactions on Wireless Communications*, *16*(8), 4924-4938.
Wang, C., Yu, F. R., Liang, C., Chen, Q., & Tang, L. (2017b). Joint computation offloading and interference management in wireless cellular networks with mobile edge computing. *IEEE Transactions on Vehicular Technology*, *66*(8), 7432-7445.
Wang, S., Urgaonkar, R., Zafer, M., He, T., Chan, K., & Leung, K. K. (2015, May). Dynamic service migration in mobile edge-clouds. In *2015 IFIP Networking Conference (IFIP Networking)*(pp. 1-9). IEEE.
Wang, S., Zhang, X., Zhang, Y., Wang, L., Yang, J., & Wang, W. (2017a). A survey on mobile edge networks: Convergence of computing, caching and communications. *IEEE Access*, *5*, 6757-6779.
Wang, S., Urgaonkar, R., He, T., Chan, K., Zafer, M., & Leung, K. K. (2017b). Dynamic service placement for mobile micro-clouds with predicted future costs. *IEEE Transactions on Parallel and Distributed Systems*, *28*(4), 1002-1016.
Xiao, L., Wan, X., Dai, C., Du, X., Chen, X., & Guizani, M. (2018). Security in mobile edge caching with reinforcement learning. *IEEE Wireless Communications*, *25*(3), 116-122.
Xu, J., Chen, L., & Ren, S. (2017). Online learning for offloading and autoscaling in energy harvesting mobile edge computing. *IEEE Transactions on Cognitive Communications and Networking*, *3*(3), 361-373.
Xu, X., Xue, Y., Qi, L., Yuan, Y., Zhang, X., Umer, T., & Wan, S. (2019). An edge computing-enabled computation offloading method with privacy preservation for internet of connected vehicles. *Future Generation Computer Systems*, *96*, 89-100.
Yang, L., Zhang, H., Li, M., Guo, J., & Ji, H. (2018). Mobile edge computing empowered energy efficient task offloading in 5G. *IEEE Transactions on Vehicular Technology*, *67*(7), 6398-6409.
You, C., & Huang, K. (2016, December). Multiuser resource allocation for mobile-edge computation offloading. In *2016 IEEE Global Communications Conference (GLOBECOM)* (pp. 1-6). IEEE.
You, C., Huang, K., Chae, H., & Kim, B. H. (2017). Energy-efficient resource allocation for mobile-edge computation offloading. *IEEE Transactions on Wireless Communications*, *16*(3), 1397-1411.
Zhang, H., Guo, J., Yang, L., Li, X., & Ji, H. (2017, May). Computation offloading considering fronthaul and backhaul in small-cell networks integrated with MEC. In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 115-120). IEEE.
Zhang, K., Mao, Y., Leng, S., Zhao, Q., Li, L., Peng, X., ... & Zhang, Y. (2016). Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks. *IEEE access*, *4*, 5896-5907.

Zheng, J., Cai, Y., Wu, Y., & Shen, X. S. (2016, July). Stochastic computation offloading game for mobile cloud computing. In *2016 IEEE/CIC International Conference on Communications in China (ICCC)* (pp. 1-6). IEEE.

Zhou, W., Fang, W., Li, Y., Yuan, B., Li, Y., & Wang, T. (2019). Markov Approximation for Task Offloading and Computation Scaling in Mobile Edge Computing. *Mobile Information Systems*, *2019*.

DECISION SCIENCES INSTITUTE

The Impact of Gamification Achievement on Continuance Intention

Nhat Quang Pham
National Cheng Kung University
Email: RA8077032@mail.ncku.edu.tw

Jeng-chung Victor Chen
National Cheng Kung University
Email: victor@mail.ncku.edu.tw

ABSTRACT

Despite the rise of gamification nowadays, the role of achievement in a gamified information system on users' continuance intention is not well-understood. From the perspective of Self-determination theory, we predicted that gamification achievement affects continuance intention through the mediating roles of identity centrality and lost performance cost. Achievement motivation is considered a moderator to the relationship between gamification achievement and identity centrality. After analyzing 131 samples collected from Reddit.com users, we found that most of our hypotheses are supported. The findings have important implication for Self-determination theory as well as gamification research and practice.

KEYWORDS: Self-determination theory, Gamification, Online platforms, Gamification Achievement, Switching cost.

INTRODUCTION

Gamification is the idea of applying gaming elements in designing a non-gaming system (Koivisto & Hamari, 2019; Robson, Plangger, Kietzmann, McCarthy, & Pitt, 2015). The practice of gamification has been applied in various aspects of daily life such as education, government, and business (Koivisto & Hamari, 2019). In education, gamification has been found to increase students enjoyment and engagement, thus, improving learning outcome (Bonde et al., 2014; De-Marcos, Domínguez, Saenz-de-Navarrete, & Pagés, 2014). In business context, studies showed that gamification has a positive impact on customer engagement (Cechanowicz, Gutwin, Brownell, & Goodfellow, 2013). Furthermore, the application of gaming elements in working environment can also increase employee productivity (Landers, Bauer, & Callan, 2017; Morschheuser & Hamari, 2019). In general, the development of gamification is growing at a rapid pace. In fact, market value of gamification reached 5.5 billion USD in 2018 [https://www.mordorintelligence.com/industry-reports/gamification-market]. Furthermore, gamification is spreading online in recent years.

Literature of gamification has focused on dissecting the components and examining the consequences of gamification. However, there are some research gaps remain. Firstly, although previous studies indicated that reward is an important component of gamification (Hamari & Eranti, 2011; Khaleel, Sahari, Wook, & Ismail, 2016), few studies differentiate between various types of reward in a gamified system. Instead, previous studies tend to group many types of reward systems together under the same umbrella of Gamification. Secondly, to the extent of our knowledge, the role of individual differences in is not yet well-understood in gamification. Most studies conducted in this field assume that a gamified system has the same effect on every user. That is contestable because, from the observation of gaming industry, different people enjoy playing games on many different levels (Chou & Tsai, 2007; Jeng & Teng, 2008). Because of that, we argue that we have no reason to assume the replication of game element in a non-game environment would have the same effect on everyone.

By filling those research gaps in this study, we hope to improve our understanding of the effect of gamification on continuance intention, an important construct in information systems research. Firstly, we focus on only one type of reward, which is permanent reward. Unlike short-term rewards that only benefit users for a specific time period, permanent reward is always retained with users once given. These permanent rewards can be accumulated over time and become users' achievement. In this work, we are interested in how this accumulation of rewards affect continuance intention of users. Secondly, this study investigates the effect of gamification on continuance intention of system users while taking into consideration the uniqueness of each individual. Since this study focuses on the role of achievement, we would like to find out how achievement motivation, a personality trait representing the desire for achievement, could affect the aforementioned relationship. We hope to contribute to the understanding of the effect of gamification in information systems by answering the following research questions:

Q1: What is the effect of gamification achievement on system users' continuance intention?
Q2: What is the effect of achievement motivation on the role of gamification achievement in continuance intention?

To answer those questions, we started from the perspective of Self-determination theory. We then identified the key variables that play an important role in the process in which gamification achievement affects continuance intention. Those are identity centrality, lost performance cost, and achievement motivation. In the next section, we will review relevant literature and theories before constructing hypotheses.

LITERATURE REVIEW

Self-determination Theory

In the domain of human behavior in gaming, Self-determination theory is one of the most popular theoretical foundation (Koivisto & Hamari, 2019). That is because gaming behaviors express the influence of intrinsic motivations, which can be well explained by this theory. Indeed, empirical evidence suggests that people can play games for pleasure and enjoyment from the activity itself without an external reward (Rigby & Ryan, 2011). Due to the similarities between a gamified system and a normal game, we select Self-determination theory as our theoretical framework.

Self-determination theory explains behaviors from the perspective of intrinsic and extrinsic motivation. The theory posits that humans have universal needs for autonomy, competence, and relatedness (Deci & Ryan, 2000; Ryan & Deci, 2000, 2017). The need for autonomy refers to the fact that humans tend to value freedom of choice. Competence refers to the possession of skills or knowledge. Lastly, relatedness is the feeling of being respected, supported and acknowledged by other people (Deci & Ryan, 2000). The theory suggests that when a task can satisfy those needs, the individual is intrinsically motivated to perform the task (Deci & Ryan, 2000). That means even in the absence of external rewards, individuals can still perform the tasks themselves because it brings them pleasure.

In the context of a gamified system, we can expect that the gamification elements can intrinsically motivate individuals to increase the use of the system if they can satisfy one of the 3 needs mentioned above. However, motivation research showed that extrinsic rewards, in some situations, can cause a phenomenon called Motivation crowding out (Frey & Oberholzer-Gee, 1997; Georgellis, Iossa, & Tabvuma, 2010). That is when the introduction of external rewards undermines intrinsic motivation of the individual to carry out the task. In other words, the individual no longer feels enjoyment when engaging in that activity. This phenomenon is also reported in gamification literature (Zhao, Detlor, & Connelly, 2016).

Cognitive evaluation theory, as a subtheory of Self-determination theory, was developed to explain Motivation crowding out effect by focusing on autonomy and competence. According to Deci and Ryan (1985), an external reward can sometimes increase sometimes decrease intrinsic motivation, depending on the characteristics of the reward. If the reward seems to control the behaviors of the individual, lowering the sense of autonomy, then it can eliminate intrinsic motivation. On the other hand, if the reward, instead, signals the competence of the individual, it can raise intrinsic motivation.

Achievement in Gamified Systems

According to Simões, Redondo, and Vilas (2013), 2 categories of gaming elements are game mechanics and game dynamics. In game dynamics, Achievement is found to be an important factor (Khaleel et al., 2016) that can have a positive influence on the intrinsic motivation of players when performing tasks in games (Kumar, 2013).

In gamification literature, achievement can be defined as “goals in an achievement/reward system whose fulfillment is defined through activities and events in other systems” (Hamari & Eranti, 2011). However, we find that definition too broad for the purpose of this study, which is to focus on permanent and accumulative rewards. Therefore, in this study, we use the term “gamification achievement” to refer to the accumulative rewards displayed in a gamified system user’s profile. In reality, not all types of reward can be accumulated into achievements. There are one-time, short-term rewards that only take effect in a limited time window. After that time window ends, the reward is removed and the user can no longer benefit from it. For example, Reddit.com is operating a reward system, in which, premium users can reward other users they like by giving them a badge of honor. A badge can be of different levels (silver, gold, platinum, etc), which signifies different levels of a member’s status. With badges, users can let other people know that their contribution to the community is acknowledged and respected by the others. However, these badges only last for a short period. After that, the rewarded user’s account will turn into a normal account without any badge. This type of short-term reward cannot be accumulated over time, thus, cannot become a long term achievement in a user’s profile.

Identity Centrality

Contemporary literature in self and identity suggests that identity is a fluid, dynamic concept (Oyserman, Elmore, & Smith, 2012; Sussman, 2000). A person can adopt multiple identities within themselves, and at a specific moment, only one of those identities becomes salient to influence emotional and cognitive processes of that individual (Oyserman et al., 2012). However, among many identities an individual can have, there might be identities that are more important than others. These identities are given more weight and are easier to become salient than other identities. Identity centrality refers to the extent to which an identity is important to an individual (Stryker & Serpe, 1994). In this study, we use the term identity centrality to refer to the degree to which the platform member identity is important to a user.

According to Cognitive evaluation theory's viewpoint on competence, an extrinsic reward can make a task meaningful if that reward signifies to the individual about his competence, which is one of the three basic needs of humans, as discussed in the previous section. Because of that, rewards in gaming can raise intrinsic motivation of players. We argue that, in a gamified system, gamification achievement reflects the level of contribution to the community of a user. As is found in online community literature, individuals with high contribution to the online community are more likely to be perceived as high status, highly competent by the community (Alexander Hars, 2002; Stewart, 2005). Therefore, we expect the users with high gamification achievement to have more intrinsic motivation when participating in a gamified system. Self-determination theory also suggests that intrinsic motivation plays an important role in human character development (Deci & Vansteenkiste, 2004). If an activity is perceived as bringing meaningful, the individual will be more likely to incorporate that activity into her self-evaluation and personal needs (Deci, Eghrari, Patrick, & Leone, 1994). For individuals having rich gamification achievement in online platforms, we hypothesize that the importance of a platform member identity is higher than people with low gamification achievement.
H1a: Gamification achievement positively influences identity centrality.

However, the effect that gamification achievement has on identity centrality might not be unambiguous due to another possible mechanism. According to Cognitive evaluation theory, when an individual is a provided reward in a gamified system, it can create a feeling of losing control (Deci & Ryan, 1985). The presence of extrinsic rewards might cause the individual to think that the reason she participates in this platform is to gain those rewards, not because she enjoys doing it. As a result, the individual will perceive her activity to engage in the platform as a violation of her autonomy, thus, undermine her intrinsic motivation. This is consistent with findings in previous studies about gamification in Information Systems. Zhao et al. (2016) found that the introduction of virtual reward in a gamified system reduces individuals' intrinsic motivation to participate in that system. If the virtual rewards in a gamified system signify to the individual about their loss of autonomy, then it is possible that it will reduce the pleasure she has when participating. We can predict that the individual, in that case, will be less likely to internalize the value of membership in that platform, thus, identity centrality will diminish. This effect is the opposite of the aforementioned prediction, making the overall effect of gamification achievement on identity centrality ambiguous. Therefore, we propose a competing hypothesis in addition to H1a.
H1b: Gamification achievement negatively influences identity centrality.

Lost Performance Cost

By signifying their status in the community, highly accomplished community members can increase their influence toward others (Stewart, 2005). According to Self-determination theory, humans have a fundamental need to be recognized and understood by others (Relatedness); therefore, we expect that being able to influence other users and being recognized and respected by the community are important benefits to platform users. Literature in online communities also found that an important reason for people to participate in these platforms is to earn social status among the community members (Lampel & Bhalla, 2007; Y. Wang & Fesenmaier, 2004). In an online platform, giving rewards is a mechanism to give recognition for the contribution of users and signify their status (Alexander Hars, 2002). Therefore, gamification achievement can help increase the utility the platform can provide to its users.

Prior research defines Lost performance cost as the perceived lost of benefits and privileges from switching the use of a system into another (Jones, Mothersbaugh, & Beatty, 2002). Suppose that there is no connection between the 2 systems, all the gamified achievement in the old system will become valueless in the new one. Thus, the users have to restart the whole process to earn a status in the new community from scratch. If it is the case that achievement in a gamified system can provide perceived benefits to users, those perceived benefit will be lost when they switch to using a new system. Furthermore, the more achievement a user has in the old system, the more perceived benefits will be lost from the switch. From that reasoning, we can hypothesize:
H2: Gamification achievement positively influences lost performance cost.

Research in identity found that an identity, when made salient, can influence a person's values and beliefs (Oyserman et al., 2012). Individuals with a salient group identity tend to give a favorable assessment about their group (Turner & Reynolds, 2001). In addition, as explained in the previous sections, a certain identity can be important than others, which means it can influence an individual's values and beliefs stronger than other identities (Stryker & Serpe, 1994). Therefore, if an individual highly values her identity as a member of a group, that identity will become salient more frequently and influences her behaviors stronger than other identities. In the context of a gamified system, we could expect that people who highly value their identity as a member of the platform are more likely to appreciate the utility they receive from it. As a matter of fact, if they switch to use another system, they will have to give up a larger amount of perceived benefit compared to other people. Thus, we can hypothesize:
H3: Identity centrality positively influences lost performance cost.

Continuance Intention

Intention to continue using has been paid much attention in the field of Management Information System because an IS can only offer its full value when the users decide to continue using it (Bhattacherjee, 2001). Aside from the fact that platform providers' revenue is positively correlated with the amount if usage, encouraging users to continue their use of the system is important to platform provider because it increases users' familiarity with the system, thus, improve their quality of system use. That, in turn, helps users to contribute more to the platform and also receive more benefits from it. Literature of IS examined the antecedents of intention to continue using from different perspectives. A stream of research focus on satisfaction and perceived usefulness the system brings to users (Bhattacherjee, 2001; Hsieh, Rai, & Keil, 2008; T. S. H. Teo, S. C. Srivastava, & L. Jiang, 2008).

Another approach focus on the switching cost when a user faces the decision of switching to another system. Studies found that platform can derive users' continuance intention from their resistance to changes, caused by the costs that users perceive when switching the system (Jones et al., 2002; Kim & Perera, 2008). As a subconstruct of switching cost, lost performance cost was also found to positively influence continuance intention (Jones et al., 2002; Kim & Perera, 2008; Van Deventer, 2016). Consistent with the well-established literature, we also hypothesize that people who have high perceived lost benefits when switching the system will be more likely to use the old system.
H4: Lost performance cost positively affects continuance intention.

Moderating Role of Achievement Motivation

Need for Achievement is a personality trait which defines the desire of a person for accomplishment or high standards (Murray, 1938). Individuals with high Need for Achievement tend to be more competitive and more ready to overcome difficulty. Later, D. C. McClelland (1965) used the term achievement motivation to address that concept. Achievement motivation has been found to positively affect how individuals value achievement in working and learning environment (D. McClelland, 1985; D. C. McClelland, 1965; Schoen, 2015; Sikhwari, 2014; M.-T. Wang, Chow, Degol, & Eccles, 2017).

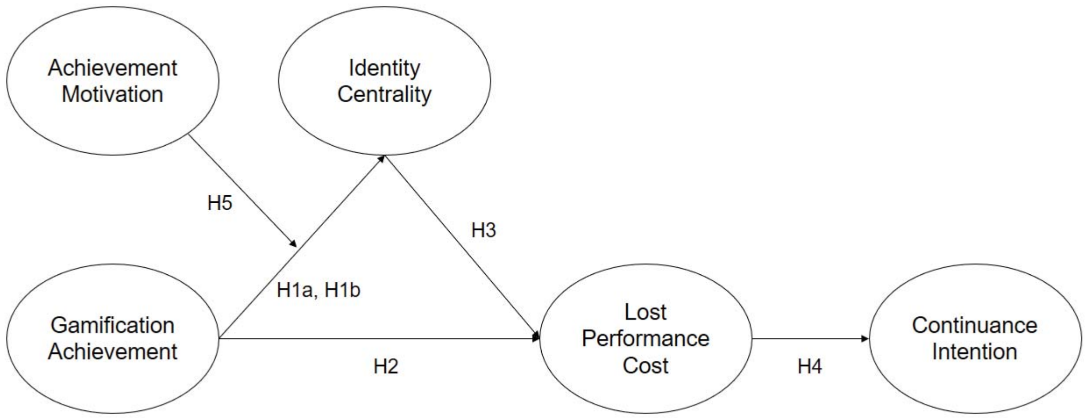
Despite being seemingly similar, the need for competence and achievement motivation are two different concepts. The need for competence refers to an innate basic need of humans meanwhile achievement motivation, coming from dispositional approach, refers to a situational motivation and can vary from people to people. According to previous studies, achievement motivation has been found to moderate the effect of feeling of competence, so that when achievement motivation is high, the effect of feeling of competence is stronger (Hofer & Busch, 2011; Schüler, Sheldon, & Fröhlich, 2010). In a gamified system, we expect the individuals that have high achievement motivation to value badges, scores, status more than people with low achievement motivation because achievements indicate how much they have accomplished compared to other people in the community. As a result, the feeling of competence that gamification achievement will bring to people with high achievement motivation will be greater than people with low achievement motivation.

As stated in the hypothesis development of H1a and H1b, the effect of gamification achievement on identity centrality is ambiguous. That is because we do not know which is stronger out of the two opposing effects, the loss of autonomy or the significance of competence. However, when achievement motivation is high, we can expect that the effect of competence will become stronger, changing the balance of the 2 opposing effects, which make the relationship between gamification achievement and identity centrality become positive. By contrast, when achievement motivation is low, we expect that the effect of competence will become insignificant, which is outweighed by the effect of loss of autonomy, making the relationship between gamification achievement and identity centrality becomes negative. Therefore, we have hypothesis 5:

H5: Achievement motivation moderates the relationship between gamification achievement and identity centrality, such that, when achievement motivation is high, gamification achievement will have a positive effect on identity centrality, but when achievement motivation is low, gamification achievement will have a negative effect on identity centrality.

CONCEPTUAL MODEL

Figure 1: Conceptual Model



METHODOLOGY

Reddit.com is one of the largest knowledge sharing platforms in the world at the moment. Also, this platform applies a gamified system in which users can collect achievement points, called Karma, for their contribution to the platform's discussion. Karma points are defined based on both the quality and quantity of their posts and comments on Reddit. When a user creates a new post or writes a new comment, Karma points will increase. If that post or comment is upvoted by other users, Karma points will also increase. By contrast, if it is downvoted, Karma points will decrease.

Regarding Reddit's immense popularity with diverse user demography and the fact that it applies a gamified system with accumulated achievement, we expect samples collected from this platform to have great generalizability and applicability. Therefore, to test the hypotheses in this study, we distributed a survey to users on Reddit.com. Only the users that own a Reddit account could participate in this study. By the end of the data collection process, we received 131 valid responses for the survey (51% male).

We adopted measurements from previous literature for identity centrality (Cohen et al., 2007), lost performance cost (Jones et al., 2002), continuance intention (T. S. Teo, S. C. Srivastava, & L. Jiang, 2008), and achievement motivation (Lang & Fries, 2006). All of these items are measured on 7-point Likert scale. The specific measurements for these constructs are presented in APPENDIX. For gamification achievement, we operationalized this variable as the number of Karma points a user has. As the number of Karma points has a highly skewed distribution, we had to transform this variable by calculating its logarithm with base 10. This log value is then used in further analyses.

This data set was then analyzed using SEM - Partial least square regression based with the software SmartPLS 3.0.

RESULTS

The data set consists of 131 samples with 51% male. Based on the results of confirmatory factor analysis, we dropped 3 items from the measurement model due to their low factor loadings (lower than 0.7). As can be seen from table 1, the remaining items all have good factor loadings. Reliability statistics indicate that our measurements have good consistency in measuring the research construct. For example, Cronbach's alpha values of all construct are higher than 0.8. Aside from factor loadings, we also take into account Average Variance Extracted (AVE) to assess convergent validity of the constructs. As is showed in the table, all AVE values are higher than 0.7, which suggests that the items can explain well the construct they are intended to measure.

Table 1: Constructs' convergent validity and reliability.

Construct	Item	Mean	Std. Deviation	Factor Loading	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Achievement Motivation	AM1	5.24	1.461	0.856	0.924	1.003	0.945	0.811
	AM3	5.08	1.69	0.922				
	AM4	5.75	1.576	0.908				
	AM5	5.19	1.66	0.914				
	AM2	5.13	1.629	Deleted				
Identity Centrality	IC1	2.48	1.68	0.936	0.856	0.856	0.933	0.874
	IC2	2.1	1.626	0.934				
	IC3	3.85	1.922	Deleted				
Lost Performance Cost	LP1	3.75	2.077	0.876	0.822	0.847	0.893	0.737
	LP2	3.98	1.941	0.909				
	LP3	3.98	1.933	0.785				
	LP4	2.18	1.538	Deleted				
Continuance Intention	CI1	6.05	1.49	0.912	0.896	0.925	0.934	0.825
	CI2	5.51	1.729	0.883				
	CI3	5.88	1.608	0.929				

In this study, we assess discriminant validity using the results of Heterotrait-Monotrait Ratio (HTMT) analysis (Table 2). Since HTMT values of all pairs of constructs in our model are lower than 0.9, we conclude that these constructs have good discriminant validity.

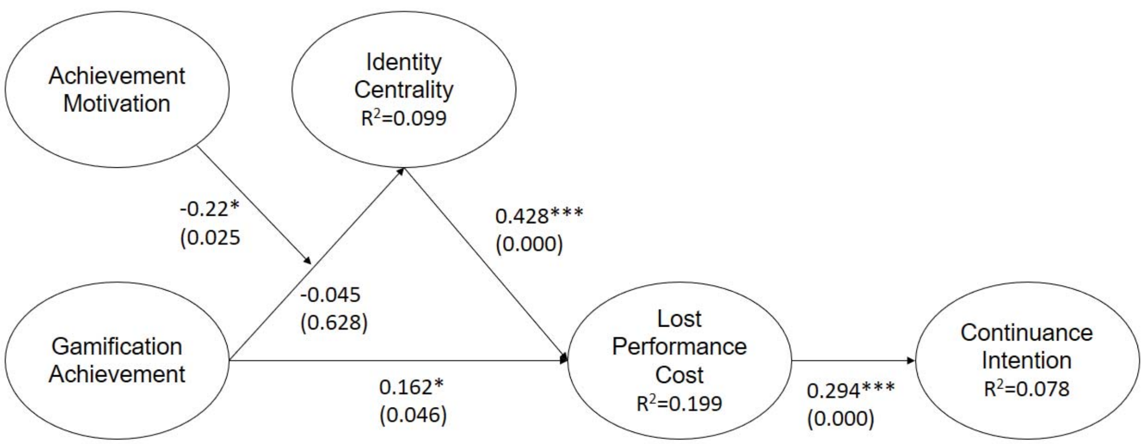
Table 2: The results of Heterotrait-Monotrait Ratio (HTMT) analysis

	Achievement Motivation	Continuance Intention	Identity Centrality	Gamification Achievement
Continuance Intention	0.257			
Identity Centrality	0.234	0.142		

Gamification Achievement	0.146	0.110	0.088	
Lost Performance Cost	0.233	0.319	0.477	0.154

Figure 2 illustrates the results of hypothesis testing. It can be seen from the result that the relationship between gamification achievement and identity centrality (H1a, H1b) is found insignificant (Beta = -0.045, p = 0.628). Next, we can observe significant relationships between gamification achievement and lost performance cost (H2, Beta = 0.162, p = 0.046), identity centrality and lost performance cost (H3, Beta = 0.428, p = 0.000), lost performance cost and continuance intention (H4, Beta = 0.294, p =0.000). In addition, we also found a significant moderating effect achievement motivation has on the relationship between gamification achievement and identity centrality (H5, Beta = -0.22, p = 0.025). This moderating effect will be examined further in the following section.

Figure 2: Results of hypotheses testing

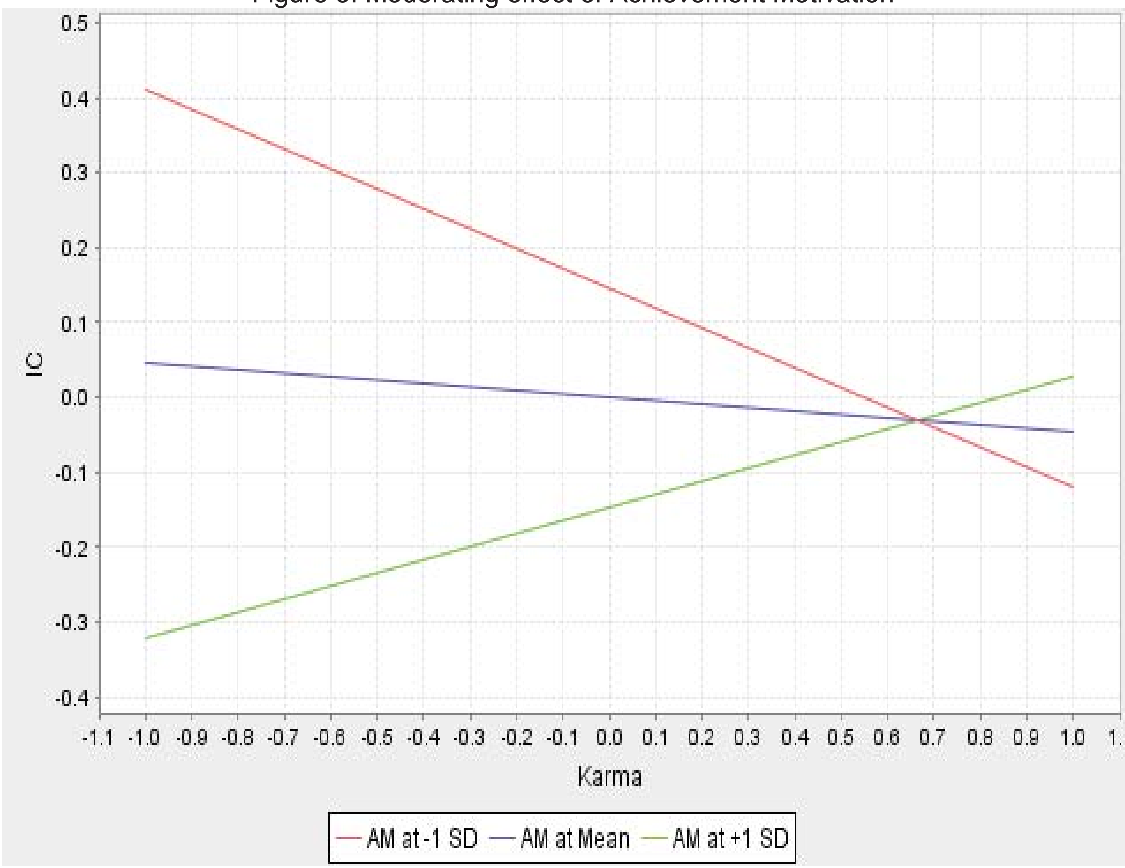


*** p < 0.001, ** p < 0.01, * p < 0.05

Figure 3 describes the direction of the moderating effect predicted in H5. In this graph, “Karma” is the abbreviation of “log of karma points”, which measures gamification achievement. IC is the abbreviation of identity centrality. From the graph, we can see that the relationship between this variable is positive when achievement motivation is high and negative when achievement motivation is low, which is consistent with hypothesis 5. However, we still do not know if these seemingly positive and negative relationships are statistically significant.

Therefore, we conducted another test using SPSS 18 software. Firstly, we used K-means cluster analysis to divide the variable Achievement Motivation into two groups. Then, we split our data set into two data sets based on these two groups. Lastly, a regression analysis is carried out for both groups with gamification achievement (Log karma points) as the independent variable and Identity Centrality as the dependent variable. From the results, when Achievement Motivation is low, Beta = - 0.477, p =0.019; when achievement motivation is high, Beta = 0.099, p =0.313. Therefore, we can conclude that achievement motivation moderates the relationship between gamification achievement and identity centrality, such that, when achievement motivation is low, this relationship is negative, when achievement motivation is high, this relationship is not significant. This result is partially consistent with hypothesis 5.

Figure 3: Moderating effect of Achievement Motivation



From the results of hypotheses testing, we conclude that H2, H3, H4 are supported. Even though the moderating effect of achievement motivation on the relationship between gamification achievement and identity centrality is not completely as predicted, the general direction of the moderating effect is still consistent with H5. Particularly, at low achievement motivation, the causal relationship is negative; at high achievement motivation, the causal relationship changes into a positive direction, even though this causal effect is not statistically significant. Therefore, H5 is partially supported. H1, however, is not supported.

DISCUSSION AND CONCLUSIONS

The results show that achievement in a gamified system increases lost performance cost for users. Even though there is no significant direct effect of gamification achievement on identity centrality, we observed a significant moderating effect of achievement motivation on this relationship. When a high achievement motivation is present, gamification achievement is found to have no impact on identity centrality. Meanwhile, gamification achievement has a negative effect on identity centrality if achievement motivation is low. This is interesting because it shows that self-determination theory can only reliably predict this causal relationship if we take into account individual differences. Next, as predicted, identity centrality has a positive effect on lost performance cost, which in turn, positively influences continuance intention.

In term of contribution to research, firstly, this study adds more to the literature of gamification on information systems. Our results show that gamification achievement influences intention to continue using the system through the mediating roles of identity centrality and lost performance cost. Interestingly, that relationship between gamification achievement and continuance intention is not straightforward. On the one hand, gamification achievement can increase lost performance cost, which, in turn, increases continuance intention. On the other hand, we found that achievement motivation, an individual trait, plays an important role in this relationship. High gamification achievement in the presence of low achievement motivation lessens identity centrality, which negatively affects lost performance cost and continuance intention. However, for individuals with high achievement motivation, the effect of gamification achievement on continuance intention is positive and straightforward. Secondly, this study also contributes to Self-determination theory by providing evidence that validates the theory and by adding individual differences when applying the theory to explain human behaviors. The results supported our prediction about the need for relatedness, which explains why gamification achievement directly increases lost performance cost. About the needs for competence and autonomy, we found that there is no significant direct effect from gamification achievement to identity centrality, probably due to the balance between two opposing effects of the loss of autonomy and the significance of competence. However, this balance changes depending on the level of achievement motivation, such that, at high achievement motivation, this causal effect is negative, at low achievement motivation, it is insignificant. The results of this study also provide evidence for Motivation crowding-out phenomenon. However, we found that the occurrence of this phenomenon varies depending on individual traits, in this case, achievement motivation. This is also a contribution to the literature of Motivation crowding-out. Last but not least, this result adds more to our understanding of Self-determination theory that only the individuals who have high achievement motivation manifest the effect of the need of competence.

For practice, this study provides important implications for gamified information systems. Firstly, platforms can encourage users to stay instead of switching to other platforms by offering permanent rewards for their contribution. In the context of a free internet nowadays, it is tempting for users to switch the use of a platform to another as they have many options at hand. Thus building relationships with users this way might be a low-cost but useful approach. Secondly, platform providers should invest more in users with much gamification achievement to encourage these people to contribute more to the platform. For example, platform providers may give these users physical gift, or extra system use privileges for these users. Compared to normal users, the investment in highly achieved users might be more worthwhile because, according to this study’s results, these users are likely to stick with the platform and contribute more in long-term. Next, the application of a permanent reward system has a more positive effect on people having high achievement motivation. Therefore, platform providers should consider the characteristics of their pool of users in terms of individual personality and culture. With the rise of big data analytics, online platform providers can have much information about their users, thus, a gamified system can be easily customized to match users’ characteristics.

This study has several limitations. Firstly, the data was collected in a single platform, raising the concern for the generalizability of our sample. However, since Reddit.com is one of the largest online forum in the world with diverse demography of users and a wide range of discussion topics, we believe this problem is not too severe. Secondly, we only investigated two direct consequences of gamification achievement, identity centrality and lost performance cost. In fact, the effect of gamification achievement might not be limited to those variables.
We have some suggestions for future research. Firstly, future studies should explore this phenomenon in different platforms with different types of reward. Secondly, scholars can explore more potential consequences of gamification achievement in terms of cognitive and emotional responses from users.

APPENDIX
Measurements

<u>Achievement Motivation</u>
AM1. I like situations, in which I can find out how capable I am.
AM2. When I am confronted with a problem, which I can possibly solve, I am enticed to start working on it immediately.
AM3. I enjoy situations, in which I can make use of my abilities.
AM4. I am appealed by situations allowing me to test my abilities.
AM5. I am attracted by tasks, in which I can test my abilities.

<u>Identity Centrality</u>
IC1. Being a redditor is an important part of who I am.
IC2. Being a redditor is important to my sense of self-esteem.
IC3. I embrace being a redditor.

<u>Lost Performance Cost</u>
LP1. Using Reddit provides me with particular privileges I would not receive elsewhere.
LP2. By continuing to use Reddit, I receive certain benefits that I would not receive if I switched to a new one.
LP3. There are certain benefits I would not retain if I were to switch to another system.
LP4. I would lose preferential treatment if stopped using Reddit.

<u>Continuance Intention</u>
CI1. I intend to use Reddit rather than discontinue it.
CI2. My intention is to use Reddit rather than use any alternative system.
CI3. I will not discontinue the use of Reddit.

REFERENCES

Alexander Hars, S. O. (2002). Working for free? Motivations for participating in open-source projects. *International Journal of electronic commerce, 6*(3), 25-39.
Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly*, 351-370.
Bonde, M. T., Makransky, G., Wandall, J., Larsen, M. V., Morsing, M., Jarmer, H., & Sommer, M. O. (2014). Improving biotech education through gamified laboratory simulations. *Nature Biotechnology, 32*(7), 694.

Cechanowicz, J., Gutwin, C., Brownell, B., & Goodfellow, L. (2013). *Effects of gamification on participation and data quality in a real-world market research domain.* Paper presented at the Proceedings of the first international conference on gameful design, research, and applications.

Chou, C., & Tsai, M.-J. (2007). Gender differences in Taiwan high school students' computer game playing. *Computers in Human Behavior, 23*(1), 812-824.

Cohen, G. L., Sherman, D. K., Bastardi, A., Hsu, L., McGoey, M., & Ross, L. (2007). Bridging the partisan divide: Self-affirmation reduces ideological closed-mindedness and inflexibility in negotiation. *Journal of personality and social psychology, 93*(3), 415.

De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education, 75*, 82-91.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*(1), 119-142.

Deci, E. L., & Ryan, R. M. (1985). Cognitive evaluation theory *Intrinsic motivation and self-determination in human behavior* (pp. 43-85): Springer.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268.

Deci, E. L., & Vansteenkiste, M. (2004). Self-determination theory and basic need satisfaction: Understanding human development in positive psychology. *Ricerche di psicologia*.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American economic review, 87*(4), 746-755.

Georgellis, Y., Iossa, E., & Tabvuma, V. (2010). Crowding out intrinsic motivation in the public sector. *Journal of Public Administration Research and Theory, 21*(3), 473-493.

Hamari, J., & Eranti, V. (2011). *Framework for Designing and Evaluating Game Achievements.* Paper presented at the Digra conference.

Hofer, J., & Busch, H. (2011). Satisfying one's needs for competence and relatedness: Consequent domain-specific well-being depends on strength of implicit motives. *Personality and Social Psychology Bulletin, 37*(9), 1147-1158.

Hsieh, J. P.-A., Rai, A., & Keil, M. (2008). Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. *MIS Quarterly*, 97-126.

Jeng, S.-P., & Teng, C.-I. (2008). Personality and motivations for playing online games. *Social Behavior and Personality: an international journal, 36*(8), 1053-1060.

Jones, M. A., Mothersbaugh, D. L., & Beatty, S. E. (2002). Why customers stay: measuring the underlying dimensions of services switching costs and managing their differential strategic outcomes. *Journal of business research, 55*(6), 441-450.

Khaleel, F. L., Sahari, N., Wook, T. S. M. T., & Ismail, A. (2016). Gamification elements for learning applications. *International Journal on Advanced Science, Engineering and Information Technology, 6*(6), 868-874.

Kim, H.-W., & Perera, N. (2008). Software continuance from the switching costs perspective: A case of web browser. *PACIS 2008 Proceedings*, 13.

Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management, 45*, 191-210.

Kumar, J. (2013). *Gamification at work: Designing engaging business software.* Paper presented at the International conference of design, user experience, and usability.

Lampel, J., & Bhalla, A. (2007). The role of status seeking in online communities: Giving the gift of experience. *Journal of computer-mediated communication, 12*(2), 434-455.

Landers, R. N., Bauer, K. N., & Callan, R. C. (2017). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior, 71*, 508-515.

Lang, J. W., & Fries, S. (2006). A revised 10-item version of the Achievement Motives Scale. *European Journal of Psychological Assessment, 22*(3), 216-224.

McClelland, D. (1985). Human Motivation (Scott Foresman, Glenview, IL). *McClellandHuman Motivation1985*.

McClelland, D. C. (1965). Toward a theory of motive acquisition. *American psychologist, 20*(5), 321.

Morschheuser, B., & Hamari, J. (2019). The gamification of work: Lessons from crowdsourcing. *Journal of Management Inquiry, 28*(2), 145-148.

Murray, H. A. (1938). Explorations in personality.

Oyserman, D., Elmore, K., & Smith, G. (2012). Self, self-concept, and identity.

Rigby, S., & Ryan, R. M. (2011). *Glued to games: How video games draw us in and hold us spellbound: How video games draw us in and hold us spellbound*: AbC-CLIo.

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons, 58*(4), 411-420.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*: Guilford Publications.

Schoen, J. L. (2015). Effects of implicit achievement motivation, expected evaluations, and domain knowledge on creative performance. *Journal of Organizational Behavior, 36*(3), 319-338.

Schüler, J., Sheldon, K. M., & Fröhlich, S. M. (2010). Implicit need for achievement moderates the relationship between competence need satisfaction and subsequent motivation. *Journal of Research in Personality, 44*(1), 1-12.

Sikhwari, T. (2014). A study of the relationship between motivation, self-concept and academic achievement of students at a University in Limpopo Province, South Africa. *International Journal of Educational Science, 6*(1), 19-25.

Simões, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior, 29*(2), 345-353.

Stewart, D. (2005). Social status in an open-source community. *American sociological review, 70*(5), 823-842.
Stryker, S., & Serpe, R. T. (1994). Identity salience and psychological centrality: Equivalent, overlapping, or complementary concepts? *Social psychology quarterly*, 16-35.
Sussman, N. M. (2000). The dynamic nature of cultural identity throughout cultural transitions: Why home is not so sweet. *Personality and Social Psychology Review, 4*(4), 355-373.
Teo, T. S., Srivastava, S. C., & Jiang, L. (2008). Trust and electronic government success: An empirical study. *Journal of Management Information Systems, 25*(3), 99-132.
Teo, T. S. H., Srivastava, S. C., & Jiang, L. (2008). Trust and Electronic Government Success: An Empirical Study. *Journal of management Information Systems, 25*(3), 99-131.
Turner, J. C., & Reynolds, K. J. (2001). The social identity perspective in intergroup relations: Theories, themes, and controversies. *Blackwell handbook of social psychology: Intergroup processes, 4*, 133-152.
Van Deventer, S. (2016). *The Impact of Relational Switching Costs on the Decision to Retain or Replace IT Outsourcing Vendors.* Auckland University of Technology.
Wang, M.-T., Chow, A., Degol, J. L., & Eccles, J. S. (2017). Does Everyone's Motivational Beliefs about Physical Science Decline in Secondary School?: Heterogeneity of Adolescents' Achievement Motivation Trajectories in Physics and Chemistry. *Journal of youth and adolescence, 46*(8), 1821-1838.
Wang, Y., & Fesenmaier, D. R. (2004). Towards understanding members' general participation in and active contribution to an online travel community. *Tourism management, 25*(6), 709-722.
Zhao, L., Detlor, B., & Connelly, C. E. (2016). Sharing knowledge in social Q&A sites: the unintended consequences of extrinsic motivation. *Journal of Management Information Systems, 33*(1), 70-100.

DECISION SCIENCES INSTITUTE

Streaming Music: How Customers Perceive Quality in a Competitive Market

Vafa Saboori-Deilami
Dominican University of California
Email: vafa.saboori@dominican.edu

ChangSeob Yeo
Dominican University of California
Email: changseob.yeo@dominican.edu

ABSTRACT

This study evaluates factors that influence the perception of quality among customers of music streaming services. Multiple theoretical frameworks have been used to develop a conceptual model. This study analyzed survey data from 247 customers of different music streaming services using multiple regression analysis. Results of this analysis have been discussed to present new insight for different stakeholders in the digital music industry.

KEYWORDS: Music, Streaming, Quality, Technology Management, Decision Making

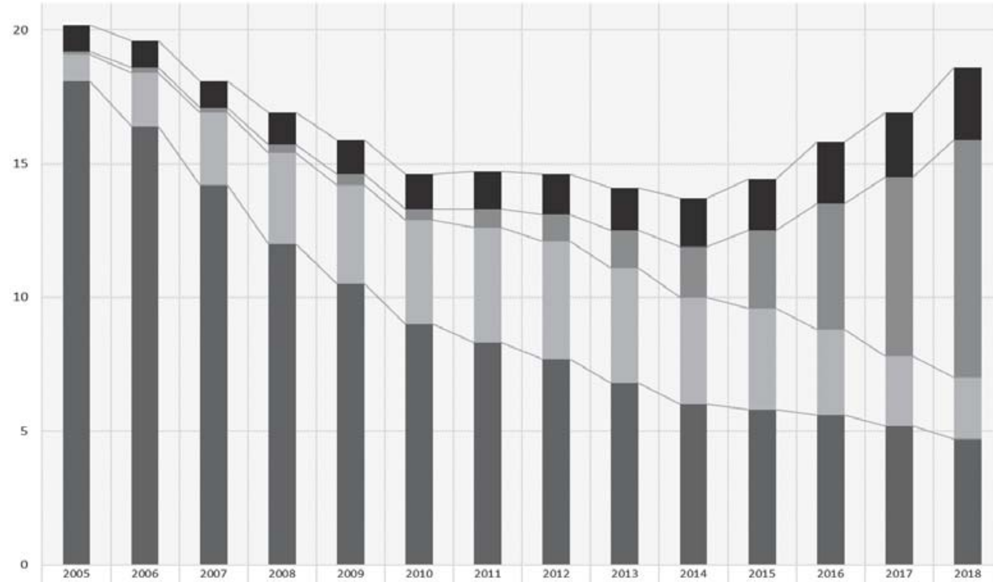
INTRODUCTION

Music streaming is disrupting music industry at an unprecedented rate (Datta et al., 2017). According to a recent report (RIAA, 2018) 75 percent of music industry revenue in the United States is accounted for by streaming services. As shown in figure 1, at the global level the revenue of music streaming services is greater than licencing, downloads, and physical CDs combined.

However, what seems to be missing from the literature, is the congestion of market. In other words, there is an evident gap in the literature when it comes to what defines quality in a highly competitive market. According to a recent report (Wilson, 2018) most streaming services are at par in terms of quality, subscription fee, availability of the free version, offering mobile app, and finally providing non-music content like podcasts. In this saturated and highly competitive market, the ultimate question is how consumers perceive quality.

This study, utilizing multiple theoretical perspectives, investigates the determinants of users' perception of quality with regards to music streaming services. Perception of quality in technology management literature has been proven to be an antecedent of perceived value which in turn affects the likelihood to adapt a technology (Venkatesh and Davis, 2000; Venkatesh and Bala, 2008; Venkatesh et al, 2003; Zeithaml, 1988; Lemon and Verhoef, 2016; Chiu et al., 2014). Drawing on this stream of literature, this study considers perceived quality as a significant proxy for adaptation of a specific streaming service.

Figure 1: Global Recorded Music Industry Revenues 2001-2018 (US$ Billions) Source: IFPI



	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
■ Performance	1	1	1	1.2	1.3	1.3	1.4	1.5	1.6	1.8	1.92	2.3	2.4	2.7
■ Streaming	0.1	0.2	0.2	0.3	0.4	0.4	0.7	1	1.4	1.9	2.9	4.7	6.7	8.9
■ Digital (ex. streaming)	1	2	2.7	3.4	3.7	3.9	4.3	4.4	4.3	4	3.8	3.2	2.6	2.3
■ Physical	18.1	16.4	14.2	12	10.5	9	8.3	7.7	6.8	6	5.8	5.6	5.2	4.7

LITERATURE REVIEW

This study synthesizes multiple theoretical frameworks to propose a conceptual model. These frameworks include theory of planned behavior (Ajzen, 1985, 1991), technology acceptance model (Davis et al, 1989), and objective studies of technology adaptation (Dodds et al, 1991; Cronin et al, 2000).

Theory of planned behavior (Ajzen, 1985, 1991) holds that behavioral intentions are the main driver of human behavior. This theory also states that intentions are driven by Attitude, subjective norms and behavioral control. Ajzen (1991, p.188) defines attitude as "the degree to which a person has a favorable or unfavorable evaluation or appraisal of the behavior in question". Ajzen (1991, p.183) defines perceived behavioral control as "people's perception of the ease or difficulty of performing the behavior of interest." Ajzen (1991, 188), continues to define subjective norm as "the perceived social pressure to perform or not to perform the behavior".

In a nutshell, the theory of planned behavior posits that attitude, subjective norms, and behavioral control, together shape user's intention, which in turn leads to a displayed behavior or act. In other words, if individuals believes that a certain action is advantageous, and if they believe that their immediate social network hold a favorable opinion towards that action, and finally if individuals conclude that they can properly perform that action, then they will form a positive behavioral intention towards that action.

The second theory that this study adapts is technology acceptance model (Davis, 1989). This theory has been widely employed and tested in technology management literature (Koufaris, 2002). Technology acceptance model states that the main reason behind individuals intention to use a computer system is their attitude towards that system. This theory also holds that attitude is formed by perceived usefulness and perceived ease of use. Davis et al. (1989, p.320) defines perceived ease of use as “the degree to which a person believes that using a particular system would be free from effort” and the perceived usefulness as “the degree to which a person believes that using a particular system would enhance his or her job performance”. Essentially, the technology acceptance model argues that what drives people to adapt a technology, is not the technology itself, but people’s perceptions towards technology, which obviously can change.

The third and last theoretical model that will be used in this study to develop a conceptual framework, is the performance-based approach that investigates specific characteristics of service as the main driver of behavioral intentions. According to these studies (Dodds et al, 1991; Cronin et al, 2000; Lemon and Verhoef, 2016) what actually leads to users’ intention to employ a system is not the subjective perception of users, but the objective characteristics of the system, i.e. quality of service. Lewis and Booms (1983) defined service quality as “a measure of how well the service level delivered matches customer expectations.”

According to this stream of literature, the main factor that leads to individuals’ intention to adapt a service is service value. Zeithaml (1988 p. 14) defines service value as “the consumers’ overall assessment of the utility of a product based on perceptions of what is received and what is given.” Service value itself is affected by perceived service quality. The relationship between quality, value, and intention, is the principal reason that this study chose perceived service quality as the dependent variable.

According to Cronin et al (2000), service value is also influenced by perceived sacrifice. Sacrifice in essence is what has to be given up (money, time, effort, etc.) to adapt a specific service. Multiple studies show that perceived sacrifice is directly or indirectly drive customers’ behavioral intention to act in a certain way, e.g. adapt a specific technology.

THEORETICAL DEVELOPMENT

Drawing on the technology acceptance model (Davis et al., 1989; Venkatesh and Davis, 2000; Venkatesh and Bala, 2008; Venkatesh et al, 2003), performance based approach towards intention for service adaptation (Cronin et al, 2000, Dodds et al, 1991), and theory of planned behavior (Ajzen, 1985, 1991), this study proposes a conceptual model that investigates what factors affect customers perceived quality in the context of music streaming services.

This study proposes the perceived quality as dependent variable (Y), and argues that it is affected by five main independent variables: subjective norm (X5), perceived usefulness (X6), perceived ease of use (X7), perceived sacrifice (X8), and perceived control (X9).

In the following section, this study defines each of these variables and provides reasoning for their relationship with perceived quality. Perceived quality in this study is conceptualized as the balance between service performance and customer expectations. In other words, based on the performance-based approach (Cronin et al, 2000, Dodds et al, 1991), this study defines perceived quality as the level and consistency of the characteristics of a specific music streaming service in its fit to subscriber’s expectation.

This study categories subscribers to music services in 5 different groups (X1, X2, X3, and X4) based on their previous usage of a service and decision of switching to a different service brand or staying with the same service brand. First group are “non-users subscribing to a brand” who are characterized as people who are not using any service but decide to subscribe to a brand. Second group are “free version users subscribing to the same brand”. This group are people who are using the free version of a brand and decided to subscribe to the paid version of the same brand. Third group are “free version users subscribing to a different brand”. This group are identified as people who are using the free version of a brand but decided to subscribe to the paid version of a different brand. Fourth group are “subscribed users subscribing to a different brand". These users are people who have subscribed to a paid service but decide to subscribe to a different paid service. Finally, fifth group are "users renewing their subscription". They are users who decide to renew their paid subscription to the same service they have been using.

This study conceptualizes subjective norm (X5) as the social pressures that individuals are under to subscribe to a specific music streaming service. This subjective norm could be exerted from peer or influencers (e.g., Venkatesh et al. 2003, Ajzen 1991, 2015). In other words, if an influencer or peer advocates for a specific music streaming service, individual would conclude that the streaming service bears high quality. Therefore a positive association between subjective norm and perceived quality is proposed.

Perceived usefulness (X6) in this study is construed as the extent to which a person believes that the music streaming service will enhance his or her music listening performance (i.e., accessibility to songs and frequency of listening) and enable him or her to enjoy music better (e.g., Venkatesh and Davis 2000). Similarly, perceived ease of use (X7) is defined as the extent to which a person believes that using the music streaming service will be effortless (e.g., Venkatesh and Davis 2000). Drawing on the technology acceptance model, a positive association between perceived usefulness and ease of use, and positive attitude towards the service is expected. This study defines quality as the balance between service characteristics and customer needs, among which are perceived usefulness and ease of use. Several studies (Landrum and Prybutok, 2004; Parasuraman et al., 1988) propose a positive association between perceived usefulness and ease of use, and perceived service quality. Therefore a positive association between these two variables and service quality is expected.

This study defines perceived sacrifice (X8) as monetary and non-monetary costs that subscribers of a specific music service should bear. This definition is developed based on the works of Lovelock and Patterson (2015), Lemon and Verhoef (2016), Cronin et al (2000), Dodds et al (1991), and Ponte et al. (2015). According to different studies, the relationship between sacrifice and value is an adverse relationship (Monroe, 2009, Dodds et al, 1991; Cronin et al., 2000; Woodruff, 1997; Sirakaya-Turk et al., 2015). Conversely, the relationship between value and quality is stipulated as a positive association (Hauser and Urban, 1986; Kumar and Reinartz, 2016). Therefore an adverse relationship between sacrifice and quality could be conferred. Moreover, this study argues that in a saturated and competitive market with low differentiation, users would be highly price conscious and price (i.e. sacrifice) is expected to negatively affect users' attitude towards quality of a specific music streaming service.

This study conceptualizes perceived control (X9) as the extent to which a person is confident in their ability to subscribe to, use, and unsubscribe from a service (Ajzen 1991). Literature (Parasuraman et al., 1988) holds a positive association between perceived control and perceived quality. This is in accordance with the definition of quality, i.e. balance between customers need and service performance. Customer would consider the quality of a service higher, if the service provider gives them confidence that they have full control over usage, subscription and opting out of the service. Therefore a positive association between perceived control and perceived quality is expected. Based on the above reasoning, this study proposes a multiple linear regression model. This multiple regression model, as shown in Equation 1, proposes that perceived quality is affected by subjective norm, perceived usefulness, perceived ease of use, perceived sacrifice, and perceived control. This model also suggests that different groups of users perceive quality of service in different ways. The relationship between Y and X1 to X9 is formulated as a multiple linear regression model as follow:

$$Y = \beta_0 + \beta 1(X_1) + \beta 2(X_2) + \beta 3(X_3) + \beta 4(X_4) + \beta 5(X_5) + \beta 6(X_6) + \beta 7(X_7) + \beta 8(X_8) + \beta 9(X_9) \quad (1)$$

Where:

Y:	Perceived Quality
X1 to X4:	Dummy variables showing 5 different groups of users
X5:	Subjective Norm
X6:	Perceived Usefulness
X7:	Perceived Ease of Use
X8:	Perceived Sacrifice
X9:	Perceived Control

DATA COLLECTION AND ANALYSIS

Amazon Mechanical Turk (MTurk) was used for data collection. Those who live in the United States and have 89% or higher incentive approval ratio were able to participate and fill out the survey questionnaire. In exchange for their participation, monetary compensation was provided. Respondents were instructed to review the given information of the three brands (i.e., Apple Music, Spotify, and Pandora), and make their service subscription decision. These three brands are chosen because the subscribers of these three brands account for more than half of total subscribers of music streaming services in the United States (Statista, 2018). Moreover, respondents could gain more detailed information by visiting the webpage of each brand by using URLs provided. The total sample size was 247. For the purpose of statistical analysis, along with theoretical constructs, five groups were recorded based on matching between each respondent's current usage of a service (free version user, subscribed user, or non-user; service brand in use) and his or her decision at the end of the survey questionnaire. These five groups are (1) non-user subscribing to a brand, (2) free version user subscribing to the same brand, (3) free version user subscribing to a different brand, (4) subscribed user subscribing to a different brand, and (5) users renewing their subscription. The fifth group is selected as the base case. Multiple Linear Regression was used to test the multiple regression model.

FINDINGS

Multiple linear regression analysis was used to describe the relationship between the dependent variable Y (perceived quality), four dummy variables (X1 to X4) which represent 5 different groups of users, and five explanatory variables (X5 to X9), Subjective Norm, Perceived Usefulness, Perceived Ease of Use, Perceived Sacrifice, and Perceived Control.
As shown in Table 1, in terms of their perception of quality, none of the customer groups showed a significant difference from the group who renewed the subscription (base group). Subjective norm and perceived control did not have a significant impact on perceived quality. Usefulness and ease of use, however, showed a positive and significant association with perceived quality. Finally, perceived sacrifice also showed a negative and significant influence on perceived quality.

Table 1: Summary of Multiple Linear Regression Analysis

Independent Variables	Coefficient	t-stat
β_0 (Intercept)	1.97**	4.39
X1	-0.02	-0.13
X2	-0.17	-1.82
X3	0.15	0.92
X4	0.4	1.83
X5 (Subjective Norm)	0.05	1.4
X6 (Perceived Usefulness)	0.3*	5.56
X7 (Perceived Ease of Use)	0.49*	7.3
X8 (Perceived Sacrifice)	-0.17*	-3.04
X9 (Perceived Control)	0.03	0.81
**: p-value less than .05*		
***: p-value less than .01*		

CONCLUSION

Findings of this study provide preliminary but interesting insights. First, it seems as if brand switching (i.e., free version users subscribing to a different brand and subscribed users subscribing to a different brand) does not occur solely because of the users' perceived quality of service. It may be caused by their variety seeking behavior. The insignificant influence of subjective norm and control shows that, in this saturated and highly competitive market of music streaming services differentiation is relatively low, information is abundantly available for users to evaluate alternatives, and it is easy to unsubscribe and switch to another service. In other words, users may have some amount of experiences with any music streaming service and also possess high levels of confidence in their ability to evaluate different services with respect to their usefulness and easiness to use.

Finally, the results of this study shows that music streaming service providers could no longer use higher prices as a quality signal. Monetary as well as non-monetary costs (perceived sacrifice) negatively influence users' perceived quality toward the service. Currently different music streaming service providers add more features to their service, and increase the price of their premium subscription services. However, the results of this study shows that users may not perceive these added features as a signal of high quality, but rather as unnecessary complexity added to the service. Therefore, this study suggests that the additional complexity, redundant features, or unrecognized benefits not only could not justify higher prices of premium music subscription services, but have a negative impact on users' perception of quality, and as the result, turn them away from the streaming service.

REFERENCES

Aguiar, L., & Waldfogel, J. (2018). As streaming reaches flood stage, does it stimulate or depress music sales?. International Journal of Industrial Organization, 57, 278-307.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In Action control (pp. 11-39). Springer, Berlin, Heidelberg.

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.

Ajzen, I. (2015). Consumer attitudes and behavior: The theory of planned behavior applied to food consumption decisions. Italian Review of Agricultural Economics, 70(2), 121-138.

Anthony, S. D., & Christensen, C. M. (2006). How you can benefit by predicting change. Managing innovation and change, 119.

Ariss, S. S., & Saboori, V. (2012). An integrated framework for the study of organizational innovation. International Journal of Innovation and Technology Management, 9(01), 1250003.

Callaway, S. K. (2010). Internet marketing strategies for music companies: understanding the demographics of an emerging customer segment. International Journal of Internet Marketing and Advertising, 6(2), 199-221.

Chiu, C. M., Wang, E. T., Fang, Y. H., & Huang, H. Y. (2014). Understanding customers' repeat purchase intentions in B2C e- commerce: the roles of utilitarian value, hedonic value and perceived risk. Information Systems Journal, 24(1), 85-114.

Cronin Jr, J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. Journal of retailing, 76(2), 193-218.

Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. Marketing Science, 37(1), 5-21.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management science, 35(8), 982-1003.

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. Journal of marketing research, 28(3), 307-319.

Hauser, J. R., & Urban, G. L. (1986). The value priority hypotheses for consumer budget plans. Journal of consumer research, 12(4), 446-462.

IFPI. (2019). IFPI Global Music Report 2019 Annual State of the Industry. Music Business Association. https://www.ifpi.org/downloads/GMR2019.pdf, (Accessed 26 May 2018).

Koufaris, M. (2002). Applying the technology acceptance model and flow theory to online consumer behavior. Information systems research, 13(2), 205-223.

Kumar, V., & Reinartz, W. (2016). Creating enduring customer value. Journal of Marketing, 80(6), 36-68.

Landrum, H., & Prybutok, V. R. (2004). A service quality and success model for the information service industry. European journal of operational research, 156(3), 628-642.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of marketing, 80(6), 69-96.

Lewis, R.C. & Booms, B.H. (1983). The marketing aspects of service quality, in Berry, L., Shostack, G. and Upah, G. (eds.), Emerging perspectives on services marketing, American Marketing Association Chicago, pp. 99-104.

Lovelock, C., & Patterson, P. (2015). Services marketing. Pearson Australia.

Monroe, Kent B. (2009). Pricing: Making Profitable Decisions. McGraw-Hill Primis Online

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual: A multiple-item scale for measuring consumer perc. Journal of retailing, 64(1), 12.

Ponte, E. B., Carvajal-Trujillo, E., & Escobar-Rodríguez, T. (2015). Influence of trust and perceived value on the intention to purchase travel online: Integrating the effects of assurance on trust antecedents. Tourism Management, 47, 286-302.

RIAA. (2018) Year-End Music Industry Revenue Report. The Recording Industry Association of America. https://www.riaa.com/wp-content/uploads/2019/02/RIAA-2018-Year-End-Music-Industry-
Revenue-Report.pdf, (Accessed 26 May 2018).

Saboori, V., Callaway, SK. (2012). Downloading Music: A Disruptive Innovation examined by Technology Acceptance Model. 43rd Annual Meeting of the Decision Sciences Institute, Decision Sciences Institute, San Francisco, CA, November 17 - 20, 2012, 25-40

Sirakaya-Turk, E., Ekinci, Y., & Martin, D. (2015). The efficacy of shopping value in predicting destination loyalty. Journal of Business Research, 68(9), 1878-1885.

Statista. (2018). Top U.S. Music Streaming Services by Users 2018 | Statistic. Statista, Statista, www.statista.com/statistics/798125/most-popular-us-music-streaming-services-ranked-by-audience, (Accessed 26 May 2019).

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision sciences, 39(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Wilson, J.L. (2018). The Best Online Music Streaming Services for 2019. PC Magazine. https://www.pcmag.com/roundup/260966/the-best-online-music-streaming-services, (Accessed 26, May 2019)

Wlömert, N., & Papies, D. (2016). On-demand streaming services and music industry revenues—Insights from Spotify's market entry. International Journal of Research in Marketing, 33(2), 314-327.

Woodruff, R. B. (1997). Customer value: the next source for competitive advantage. Journal of the Academy of Marketing Science, 25(2), 139.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. Journal of marketing, 52(3), 2-22.

DECISION SCIENCES INSTITUTE

Visa Policies and Innovation: How Powerful Is Your Passport?

Marc S. Mentzer
University of Saskatchewan
Email: mentzer@edwards.usask.ca

ABSTRACT

The growth of isolationist ideologies in the United States and other countries has generated interest in the relationship between international mobility and innovation. This study focuses on visa policies as an indicator of a country's openness to cross-border flows of people, and examines the relationship between visa policies and innovation.

KEYWORDS: Expatriates, Innovation, International markets, Migration, Public policy

INTRODUCTION

The tightening of border controls by the United States has fueled dialogue asking to what extent cross-border mobility affects a country's prosperity, with such discussions often centering on the role of those from other countries in contributing to innovation and creativity. Would the Silicon Valley and Hollywood – or their counterparts in other parts of the world – be the same without the influx of people across borders? These issues have been discussed long before Donald Trump became president; for example see Travern's concerns in 2006 about how the border policies of the George W. Bush administration would hamper innovation. However, the actions of the current administration have renewed concern about this issue.

Regarding the openness of the US to those who might contribute to innovation, the topic of H-1B visas was a contentious issue in the 2016 US presidential election. H-1B visas allow a foreign person to work in the US for a particular, specified employer, for a maximum of three years, although in practice the time period can be much longer. On one hand, Kerr and Lincoln (2010) have demonstrated that H-1B visa holders lead to a net increase in inventions. Others argue that H1-B visa holders are depressing wages and taking jobs that would otherwise go to US residents.

The question is not limited to the United States. The United Kingdom's decision to leave the European Union has sparked similar discussion of how the transborder flow of people helps or hinders the British economy. Several European countries are witnessing the rise of political parties who campaign for tighter border controls. Every country weighs the pros and cons of tight border restrictions to protect jobs and safeguard national security versus the free transborder movement of people, which it might be argued, enhances economic competitiveness through innovation.

Wagner and Jonkers' (2017) essay in *Nature* has sparked an interest in how national "openness" (their term) affects a country's scientific influence. Wagner and Jonkers' analysis used countries as their unit of analysis, and found that national openness correlated strongly with scientific impact, measured though a system of weighting citations in scholarly articles. Their measure of national openness was comprised of variables measuring immigration and emigration of scientists coupled with international co-authorships. They succeeded in documenting that openness correlates with scientific impact. Of particular interest are the outliers to the regression line. In the low-openness high-impact quadrant, the largest outlier was the United States, with extraordinary influence in science despite relatively closed borders. The other quadrant that is an outlier to the regression line, the high-openness low-impact quadrant, was empty except for Hungary. With that exception, Wagner and Jonkers found that high openness, implying a free flow of researchers across borders, always resulted in an above-average impact on scientific research.

The recent findings of Sugimoto, Robinson-Garcia, Murray, Yegros-Yegros, Costas, and Larivière (2017) are parallel to the research of Wagner and Jonkers (2017). Sugimoto *et al.* looked at this issue using individual researchers as their unit of analysis, and asked how Trump's border-tightening might affect US scientific productivity. Their analysis, based on data from 16 million authors from the Web of Science, showed that tighter US borders would cause non-US researchers to move to countries other than the US. In examining patterns of co-authorship, Sugimoto *et al.* found that when a researcher moves to another country, that person constructs a research network connecting both the old country and the new country, thus boosting that person's influence. In other words, the flow of researchers out of a country into another country actually helps develop innovation in both countries, as those in the first country become able to join multinational networks of researchers. Instead of the phrase "brain drain," it would be more useful to think of this as "brain circulation" (Gaillard & Gaillard, 1997; Sugimoto *et al*., 2017). Sugimoto *et al*. conclude that Trump's border policies will damage US influence in the sciences, and that other countries will fill the void by welcoming researchers who seek to advance their careers.

Wagner and Jonkers (2017) point out that their study did not include a potentially important variable, that being ease or difficulty of obtaining a visa. The current study uses measures of visa policies as indicators of a country's openness, and it is anticipated that openness will be associated with measures of innovation.

MEASURES

Independent Variables: Visa Policies

A nation's practices regarding mobility across borders comprise a broad construct that does not easily lend itself to quantitative measurement. One approach to this construct is to examine, for citizens of a given country, how many other countries they can visit without obtaining a visa in advance, a concept sometimes referred to as "Passport Power." For example, a German citizen can visit 161 countries without obtaining a visa in advance, while at the opposite end of the spectrum, an Afghan citizen can visit only 25 counties without obtaining a visa in advance (Arton, 2018b).

The mirror image of Passport Power is whether a country is welcoming to foreigners who do not have visas (Arton, 2018c). In other words, for a given country, how many countries' citizens can visit without obtaining a visa in advance? There are many countries that never require a visa in advance for any visitor. Some of these are small developing countries that cannot maintain a global network of consulates to issue visas; Djibouti, Micronesia, and Samoa are examples. At the other extreme, four countries are “least welcoming” because they require visas in advance of all visitors (Afghanistan, North Korea, Syria, and Turkmenistan).

The country scores for Passport Power and Welcoming Countries have been developed by Arton Capital, a Montreal-based consulting firm that specializes in advising clients about investor visas and second passports. Arton maintains a website, passportindex.org, which displays the scores they have assigned to each country (Arton, 2018a, 2018b, 2018c). Their list includes every jurisdiction that issues passports, including non-sovereign jurisdictions such as Hong Kong and the Palestinian Territories, as well as microstates such as Andorra and Kiribati, although these were dropped from the current analysis due to lack of data for the corresponding dependent variables. Where a country allows a visa to be issued on arrival, Arton treats the situation as one in which a visa is not required; however, as described below, whether a country issues visas on arrival is more ambiguous than one might guess at first glance. For example, the United States has created an on-line procedure known as ESTA (Electronic System for Travel Authorization), which technically is not a visa although it has many of the characteristics of a visa. Where a country allows a traveler to use an on-line procedure such as ESTA, Arton categorizes the situation as one in which a visa is not required.

The scores of Passport Power and Welcoming Countries are statistically unrelated to each other, with an insignificant correlation of -.11. Visa policies are not determined by symmetry or reciprocity. A United States citizen can visit 158 countries without an advance visa, but that does not mean that the US is easy for other citizens to enter. On the other hand, there are small, developing countries who are unselective about who they will admit because their economies depend on tourism, such as the Maldives, Seychelles, and some independent Caribbean islands, yet their own citizens find it difficult to travel abroad without visas. Supply and demand does not seem to determine a country’s visa policies. Some of the most restrictive countries are unlikely to be inundated with foreign visitors, such as Bhutan, Equatorial Guinea, Eritrea, or South Sudan, yet they have remarkably stringent polices on admitting citizens of other countries.

One relatively straightforward approach would be to distinguish between situations where a person could enter a specific country without a visa, versus those situations where a person needs a visa to enter. As experienced international travelers know, the reality of countries’ admissions standards is more nuanced than the dichotomy of visa versus non-visa. For example, one country might be casual about questioning incoming visitors at the border, while another country might subject visitors to a regimen of interrogation, searches, and official harassment. Any variable based on a numerical measure of formal rules will be imperfect in its failure to capture these qualitative shades of gray.

Further complicating the question of measurement is that some countries require a visa in advance, while others issue visas upon arrival, at least to citizens of certain countries. In some instances, this is a pro forma process that is a routine part of the bureaucracy of airport arrivals; in other instances, getting a visa at the border or airport is a roll of the dice. For example, one well-travelled business executive had harshly criticized Mauritania for its bureaucratic hurdles in issuing visas, pointing out that a country is unlikely to get foreign investment if it refuses to issue visas to foreign investors like himself (Biddlecombe, 2004). Mauritania now offers visas on arrival at their capital’s airport for visitors of all nations, although those entering at land borders must still obtain their visa through a consulate. The visa-at-airport procedure yields the highest possible score for Mauritania on Arton’s scale of Welcoming Countries (Arton, 2018c).

Travel guidebooks sometimes bluntly reveal the quirkiness and "vague bureaucracy" surrounding visa procedures, as in this statement about Cameroon:

> "Obtaining a visa is an example of how haggling and vague bureaucracy can affect all areas of African life. Strictly speaking, visitors without a visa may be required to leave Cameroon on the next available flight... Unofficially, visas are usually available at the airports at Doula and Yaoundé [West, 2011: 50]."

In a situation like the above, Arton's scores are based on a country's official visa rules, not practices that are occurring on an informal level.

An additional variation is the introduction of on-line pre-screening procedures in lieu of visas, such as the Electronic System for Travel Authorization (ESTA) procedure, described above, which the US commenced in 2008. ESTA is for those visiting the US from visa-waiver countries, but it would be justifiable to regard ESTA as a sort of "visa-light." ESTA is for short-term visitors and is not available to those who intend to work or study in the US. Other countries are beginning to introduce similar forms of advance, on-line screening that resemble the ESTA process, such as Canada's Electronic Travel Authorization, Australia's Electronic Travel Authority, and the European Travel Information and Authorization System (ETIAS). Inevitably, more pre-screening systems will be developed around the world on the ESTA model, sometimes replacing traditional visas procedures and in other instances creating a visa-like process for travelers who previously did not need a visa.

Despite these shortcomings, the current study is predicated on the assumption that the country scores for Passport Power and Welcoming are indicative of a country's overall policy toward transborder movements of people. Whether a country requires a visa for casual visitors might or might not be related to that country's openness to migrants, short-term residencies of workers, and other border movements. Nevertheless, the current study uses visa policies for casual visitors as a proxy for openness toward transborder movements of people.

A final caveat concerns the measure of Passport Power, in which it is assumed that any citizen of a country can get a passport. In practice, citizens' access to passports varies across countries, even if one sets aside the question of whether people can afford fees for issuing a passport. In some countries, all citizens have an absolute right to a passport; for example, in Canada, it is a right that is embedded in the constitution. Meanwhile, there are places where issuance of a passport to a country's citizens is sometimes denied due to ethic discrimination, sexual discrimination, a criminal record, or the political purity of the applicant.

Dependent Variables: Innovation Outputs

Measuring a construct as broad as innovation is obviously a challenge; however, the World Intellectual Property Organization (WIPO) has released a publication titled *The Global Innovation Index* with various measures of innovation for 127 of the countries of the world. For each country, over 100 highly specific measures are aggregated into seven indices labeled "Institutions," "Human capital and research," "Infrastructure," "Market sophistication," "Business sophistication," "Knowledge and technology outputs," and "Creative outputs" (WIPO, 2017). These seven indices are, in turn, aggregated into a single Global Innovation Index. As WIPO acknowledges, some variables are measures of input and others are measures of outputs. For example, the quality of government institutions, human capital, and infrastructure are the ingredients for innovation rather than measures of innovation (outputs) themselves.

The two measures of output, "Knowledge and Technology Outputs" and "Creative Outputs," are the principal dependent variables in the current study, and both measures are remarkably inclusive in how they capture a country's innovation output. Knowledge and Technology Output includes such items as patents, scientific articles, software spending, and intellectual property receipts, among other times. Creative Output includes trademarks, industrial designs, feature films, publishing, Wikipedia edits, and YouTube uploads, among other items. The composition of these variables is listed in Appendix A, and greater detail regarding how they these measures have been developed is available in Appendix III of The Global Innovation Index (WIPO, 2017). The variables have been scaled in a way to allow an equitable comparison across countries, in many instances expressed in relation to purchasing power parity of gross domestic product (GDP). Variables such as "Wikipedia edits" or "YouTube uploads" have been expressed on a per capita basis, ages 15-69. The approach of WIPO to data means it is unnecessary to adjust their data for differences in population size, purchasing power parity, or size of the economy.

RESULTS AND DISCUSSION

As shown in Table 1, there were strong correlations between one of the independent variables (Passport Power) and the principal dependent variables (Knowledge and Technology Outputs, and Creative Outputs). These results are parallel to the findings of earlier researchers who established a connection between international mobility and innovation.

Table 1: Bivariate Correlation Matrix

		1.	2.	3.	4.	5.	6.
Independent Variables	1. Passport power	1.00					
	2. Welcoming countries	-.11	1.00				
Principal Dependent Variables	3. Knowledge & technology outputs	.66**	-.06	1.00			
	4. Creative outputs	.78**	-.07	.81**	1.00		
Other Dependent Variables	5. Scientific & tech publications	.64**	-.05	.68**	.70**	1.00	
	6. H-Index	.60**	-.12	.76**	.65***	.57**	1.00
	Notes: N = 126 ** *p* < .01						

On the other hand, the ease of entering a country (the variable Welcoming Countries) was not significantly related with any measure of innovation. This is contrary to expectations.

These findings point to the conclusion that open borders are associated with innovation when we are talking about the ease of a country's citizens travelling elsewhere, but openness to other countries' citizens is *not* associated with innovation.

The explanation for such anomalous findings might rest with the phenomenon where extraneous variables lead to two variables being correlated, creating the illusion that a causes b. In other words, a plausible argument is that visa policies are mere proxies for basic, underlying national characteristics that the WIPO scores did not directly measure. Countries like the United States, low on the scale of Welcoming Countries, is likely to maintain high levels of innovation regardless of its visa policies. Countries that are high on the Welcoming scale, like the Maldives and the independent Caribbean islands, as well as troubled countries such as South Sudan, Syria, and Afghanistan, would probably rank low on measures of innovation even if they changed their visa policies. In short, the variable of Passport Power appears to capture the extent to which a country has a wealthy, highly advanced economy. Such countries are innovative for reasons that are unrelated to openness to international travelers.

Admittedly, the measures of innovation outputs used here are extremely broad. In contrast to the broad measures of innovation in the current study, Wagner and Jonkers (2017) and Sugimoto *et al*. (2017) focused specifically on published research articles, using patterns of authorship and citation patterns as their measures of innovation. One of the components of Knowledge and Technology Outputs is the number of scientific and technical articles, expressed in relation to purchasing power parity of GDP to enable easy cross-national comparisons of the data. Only articles in science and engineering journals were included (WIPO, 2017: 412-413). As shown in Table 1, Passport Power has a correlation of .64 with Number of Articles; however, contrary to expectations, Welcoming Countries has a correlation of -.05 with Number of Articles. Another measure of research publication activity is the H-Index, which WIPO operationalized in an unconventional manner as the average number of citations per article (WIPO, 2017: 413). As shown in Table 1, Passport Power has a correlation of .60 with H-Index, while Welcoming Countries has a correlation of -.12 with Number of Articles.

Even among highly educated people, most stay in their country of origin. Sugimoto *et al.* (2017) were able to measure this precisely, finding that of authors of academic articles, 96% maintain the same "country of affiliation" throughout their career. In any case, the openness-innovation relationship is more complicated than it initially appears. The ease with which a country's citizens can travel is associated with innovation, but lowering visa barriers to entry by other countries' citizens has no relationship to innovation. A country's overall state of development probably provides the best explanation of why some countries are more innovative than others.

APPENDIX

Components of Dependent Variables

1. The variable “Knowledge and technology outputs” is an index comprised of:

    a) Patents by origin/bn PPP$ GDP
    b) PCT patent applications/bn PPP$ GDP
    c) Utility models by origin/bn PPP$ GDP
    d) Scientific & technical articles/bn PPP$ GDP
    e) Citable documents H index
    f) Growth rate of PPP$ GDP/worker
    g) New businesses/th pop. 15–64
    h) Computer software spending, % GDP
    i) ISO 9001 quality certificates/bn PPP$ GDP
    j) High- & medium-high-tech manufactures, %
    k) Intellectual property receipts, % total trade
    l) High-tech exports less re-exports, % total trade
    m) ICT services exports, % total trade
    n) FDI net outflows, % GDP

2. The variable “Creative outputs” is an index comprised of:

    a) Trademarks by origin/bn PPP$ GDP
    b) Industrial designs by origin/bn PPP$ GDP
    c) ICTs & business model creation
    d) ICTs & organizational model creation†
    e) Cultural & creative services exports, % of total trade
    f) National feature films/mn pop. 15–69
    g) Global ent. & media market/th pop. 15–69
    h) Printing & publishing manufactures, %
    i) Creative goods exports, % total trade
    j) Generic top-level domains (TLDs)/th pop. 15–69
    k) Country-code TLDs/th pop. 15–69
    l) Wikipedia edits/mn pop. 15–69
    m) Video uploads on YouTube/pop. 15–69

Notes:

The variable labels above are directly quoted from WIPO, 2017, Appendix I.
More detailed explanations of variables are in WIPO, 2017, Appendix III.

Abbreviations used in this appendix:

FDI = Foreign direct investment
GDP = Gross domestic product
H index = Hirsch index (impact of research)
ICT = Information and communication technologies
ISO = International Organization for Standardization
PCT = Patent Cooperation Treaty
PPP = Purchasing power parity
TLD = Top level domain (internet)

REFERENCES

Arton Capital (2018a). About passport index. https://www.passportindex.org/about.php Accessed 10 January 2019.

Arton Capital (2018b). Global passport power rank. Passport index. https://www.passportindex.org/byRank.php Accessed 10 January 2019.

Arton Capital (2018c). Welcoming countries rank. Passport index. https://www.passportindex.org/byWelcomingRank.php Accessed 10 January 2019.

Biddlecombe, P. [pseud.] (2004). *Always feel a friend.* London: Abacus.

Gaillard, J., & Gaillard, A. M. (1997). The international mobility of brains: Exodus or circulation? *Science, technology and society*, 2: 195-228. https://dx.doi.org/10.1177/097172189700200202

Kerr, W. R., & Lincoln, W. F. (2010). The supply side of innovation: H-1B visa reforms and US ethnic invention. *Journal of Labor Economics*, 28: 473-508. http://dx.doi.org/10.1086/651934

Sugimoto, C. R., Robinson-Garcia, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, 550: 29-31. http://dx.doi.org/10.1038/550029a

Traven, M. R. (2006). Restricting innovation: How restrictive US visa policies have the potential to deplete our innovative economy. *Capital University Law Review*, 34: 693-740.

Wagner, C. S., & Jonkers, K. (2017). Open countries have strong science. *Nature*, 550: 32-33. http://dx.doi.org/10.1038/550032a

West. B. (2011). *Cameroon* (3rd ed.). St. Peter, United Kingdom: Bradt.

World Intellectual Property Organization (WIPO) (2017). *The global innovation index 2017: Innovation feeding the world* (10th ed.). http://www.wipo.int/publications/en/details.jsp?id=4193 Accessed 12 January 2019.

DECISION SCIENCES INSTITUTE

Examination Customization: A Simple Procedure to Enhance Academic Honesty

David M. Bowen
California State University, East Bay
Email: david.bowen@csueastbay.edu

ABSTRACT

Analytical courses are susceptible to academic dishonesty as complex analyses are assessed via numerical answers. Such answers are easy to grade and easy to copy from classmates. To combat the copying of answers, we introduce the concept of 'examination customization;' exams that require unique answers for each student and minimal extra time to construct and grade. Anecdotal evidence suggests that students faced with such an exam realize that copying is pointless (and in fact leads to getting caught). The majority of students appreciate test problems that force honesty and minimize peer pressure to participate in answer sharing.

KEYWORDS: Academic honesty, cheating, examination problems, student assessment

INTRODUCTION

Academic honesty is fundamental to the integrity of academic institutions, from research that accurately reports and interprets scientific data, to researchers and students generating their own results and taking credit only for their own work.

According to the International Center for Academic Integrity (ICAI), academic dishonesty has reached epidemic proportions with 39% of undergraduates admitting on surveys that they have cheated on tests and 68% that they have cheated on tests and/or written assignments (ICAI, 2019).

There are a multitude of causes contributing to this epidemic, including deterioration of negative social connotations of cheating, technological advances, and availability of internet based businesses designed to provide papers and homework or test answers for a fee (McCabe, 2012).

In Engineering Education there has been a large emphasis on fostering collaboration skills in graduates (Bowen, 2013), and so mixed messages of 'You must be able to collaborate well with others,' alongside, 'You must do this particular work completely on your own,' are received by students. For international students from cultures less individualistic in nature, the stigma of the transgression of not helping a colleague in need may be considered worse than a transgression against academic honesty.

Analytical course exams in particular can be susceptible to academic dishonesty by virtue of the fact that complex concepts and analyses are often assessed via student generation of correct numerical answers. Such exam questions are typically easy to construct, easy to grade, and easy for students to copy answers from classmates.

FACILITATING ACADEMIC HONESTY

In order to foster academic honesty, I use a three pronged approach that includes an emphasis on behaving ethically, examinations that are always ‘Open Book, Open Note and Closed Neighbor’, and constructing exams that have unique answers for each student. The latter is the focus of this paper, but the former deserve brief explanations.

In emphasizing academic integrity, I take pains to describe the implications of cheating. Of course there is the getting caught and failing or receiving zero credit, though research indicates that most cheating students do not get caught [], so such consequences are not likely to deter many. Rather, I attempt to influence students’ sense of the size, scope and implications that such dishonest conduct leads to. It is not just about fooling the professor and getting a better grade in a class. It is about the integrity of the institution, it is about the value of the degree they receive and it is about the graduates before and after, and how willing employers and others are to respect the knowledge represented by a degree. I let them know that they are relying on the integrity of those that came before them, and they are indebted to prior graduates. It is that existing integrity which gets them the interview or the job offer when an employer that does not know them is willing to take a chance on them due in large part to the stamp of approval bestowed by our institution. If they then do not have the expected professional competencies because they were more concerned with cheating to get a better grade than learning the material, they will fail professionally as well as shut doors for future graduates. At least some students nod their heads in agreement, and it seems, surprisingly, that many of the students have not made those connections before.

One common form of cheating is using prohibited notes or violating ‘closed book’ examination instructions, especially for multiple choice or fill-in-the-blank type examinations. An easy way to avoid this is to not construct exams which are closed book/notes. This does mean that a bit more time and effort is typically spent constructing and grading the examinations. This extra effort is warranted if we consider what skills will be most valued in the future. Based on those future skill considerations, it becomes apparent that closed book/note exams should be a relic of the past.

Memorization, especially of formulas or techniques in the decision sciences, is not the critical skill we need to impart. Rather, valued graduates are those who can encounter an unstructured situation and are able to identify which analytical tools can be used and then to accurately do so, including understanding any underlying assumptions. When they engage in this process professionally in today’s world, they will have virtually infinite resources available to them in real time to assist any application. Therefore it is much more important that they know where/how to find such resources and filter and utilize them correctly than it is to memorize a formula. Furthermore, if we teach them the skills and provide opportunities to practice those skills associated with utilizing available resources, then they are more likely to establish a habit of life-long learning and less likely to experience obsolescence of the facts they memorized in school.

Another easy form of cheating on examinations is looking at the answers of another student and copying that answer. This practice can be stifled if each student by design has a unique correct answer, and all the students know that their answers are *supposed* to be different than their neighbors. This practice can be thought of as ‘Examination Customization.’

Such customization is relatively easy to accomplish with little if any extra burden on faculty or students.

Constructing Custom Examination Questions

Students at all universities should have a unique identifier, usually called a registration number. At CSU East Bay, each student is assigned a 'NetID' which is two letters of the alphabet followed by 4 numbers. Other ways to generate unique input numbers could be to use 'How old you are in months,' 'Last four digits of your Driver's license/Tax Identification Number, etc..

To customize exams, simply specify the input data used in problem solving in terms of the unique identifier. For example in the following decision analysis problem (Figure 1), the ticket price is defined for each student by their unique NetID, resulting in unique correct answers:

Figure 1. Customized Final Examination 'Ticket Purchase' Problem Decision Analysis

SHOW ALL YOUR WORK

Open Book, Open Notes & Closed Neighbor---Good Luck!

Engineering 3841

Final Exam

Fall 2010

NAME______________________________________NetID_____________

Points

1) You have been looking forward to travelling during the holidays (New Orleans?), but it has become so expensive! Right now, you have the choice to buy a cheap (non refundable) ticket, buy a more expensive ticket that is refundable, or wait and buy a ticket next month. You estimate that the chances that the ticket price next month will increase/remain the same/decrease are 35% / 25% / 40%. Ticket prices are:

	Now	Next month
	Cheap = your NetID	Increase of 15% (= NetID*1.15)
	Refundable = your NetID + $250	Same = Your NetID
		Decrease of -20% (= NetID*0.8)

Note: If you buy the refundable ticket, next month you will see if prices increase / remain the same / decrease, then have the option to pay a $75 dollar fee to get a full refund then buy a ticket at the new rate.

20 a) What should you do and why?

10 b) A travel consultant charges a fee and then can tell you now for certain what prices will be next month. What should you be willing to pay for such a service?

Constructing the problem in this way ensures that each student that correctly solves this decision analysis problem will have a unique answer. For grading, it is convenient to construct an Excel spreadsheet to calculate solutions to test problems. Inputting the student's NetID into the appropriately constructed spreadsheet supplies all the correct answers for that student's examination. Figure 2 shows screen shots of the solution spreadsheet for the Ticket Purchase decision analysis problem for students with NetIDs of 1234 and 6789.

Figure 2a. Solution Set Screenshots: Decision Analysis Ticket Purchase Problem, NetID=1234

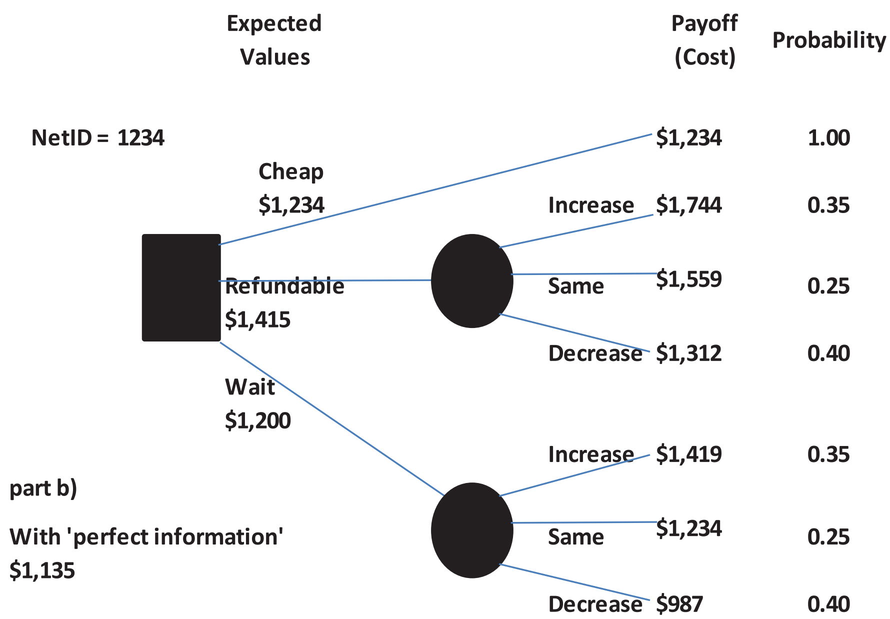
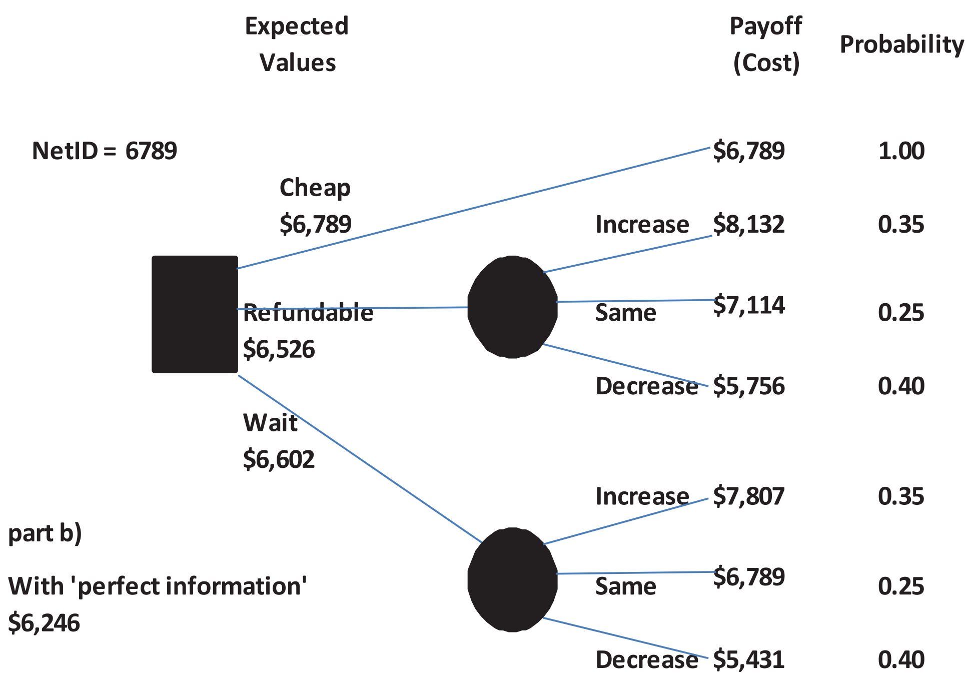


Figure 2b. Solution Set Screenshots: Decision Analysis Ticket Purchase Problem NetID=6789



The solution spreadsheet shows that the best decision, from an expected value perspective, is different depending on the NetID value, with the NetID=1234 best (lowest expected cost) decision is to 'wait' and buy ticket next month, (expected cost of $1,200), while for the student with NetID =6789, the best decision is to purchase a refundable ticket and then either keep it (60% chance) or pay $75 to get a refund and then buy the ticket at a decreased price (40% chance) for an expected value (cost) of $6,526. Similarly, the value of perfect information (part b) of the problem) is different for each case, with EVPI= $65 (=$1,200 - $1,135) for the lower NetID and EVPI = $280 (=$6,526 - $6246) for the higher NetID value.

Resource Consumption Implications for Solution Set Construction

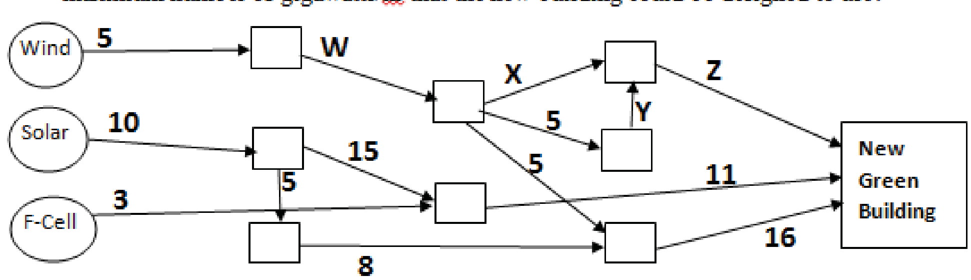
Construction of the Excel Spreadsheet to calculate answers takes approximately 10 minutes, and something similar is typically needed to calculate the answers even if the exam is not customized. Using paper and pencil might require approximately 5 minutes. The added burden on instructors to customize and grade customized exams is minimal. Inputting the NetID information requires 1 to 2 seconds to decipher and enter for each student exam paper. Likewise, the added burden on students ranges from nothing to performing a simple mathematical operation on the NetID to determine input data.

Another benefit to constructing such an Excel Solution Spreadsheet is utilizing it as a tool and visual when reviewing the examination. It can also be used to demonstrate 'sensitivity analysis' to students if there is, for example, a problem that utilized a customized interest rate. Demonstrating to students how solutions/decisions might change based on the specific value of an input such as interest rate (or tax rate or cost) can be enlightening, and is all the more memorable if it is a problem that they have considered in some depth and tried to solve.

Figure 3 shows a network flow problem, which is again based on the student's NetID digits and can result in different answers between students. In this case the students need not perform any calculations on their NetID digits prior to using them to solve the problem. They only need to transfer the W, X, Y & Z values on to the diagram and then determine their answer by identifying the 'min cut set' for their individualized network. In this case the min cut set will be 18 if W>4, or will be 13+W if W<5.

Figure 3. Customized Network Flow Exam Problem.

30 2) The University wants to build a new green building, supplied with power from renewable plants and utilizing the transmission capacities as shown. If the values are in gigawatts/hr, what is the maximum number of gigawatts/hr that the new building could be designed to use?



Linear Programming Example Examination Problem

Figure 4 shows a relatively complex linear programming formulation problem, but the customization itself is simple, straight forward, and results in unique answers. In this case only two digits from the NetID are utilized and will serve as 'right hand sides' in two constraint inequations.

Figure 4. Customized Linear Programming Formulation Exam Problem

SHOW ALL YOUR WORK

Open Book, Open Notes & Closed Neighbor—Good Luck!

Engineering 3841

Midterm #2

Points NAME________________________ NetID____________

W X Y Z

1) A regional University is doing some energy planning to determine the best course of action over the next 2 years. Options are to purchase energy from the local utility, and/or to invest in one or more alternative energy sources, where money is paid initially for the plant, but then energy is produced at no additional cost. For each technology at most one plant can be built, but 'partial' plants can be built with proportional costs and capacities, up to the max capacities shown below. Below, **all $ figures are in millions of dollars**, and capacity/demand figures are in **thousands of gigawatt hours**:

		Max Annual Capacity	
	Cost	**Peak**	**Nonpeak**
Wind	**$3**	**5**	**1**
Solar	**$5**	**10**	**1**
Fuel cell	**$2**	**3**	**3**

Expected Annual Demand		
	Peak	**Nonpeak**
Yr1	**12**	**W**
Yr2	**14**	**X**

Additionally, 'grid' energy costs $0.25 million per thousand Gigawatt hours (non peak) and $1.0 million per thousand Gigawatt hours (peak) both with unlimited capacity.

10 a) Define appropriate decision variables. (HINT: What are you deciding to do or not do?)

10 b) Show the objective function, including units.

15 c) Show the constraints, including units.

10 d) Identify a feasible solution including values of variables and objective function (HINT-No Simplex required for this!)

15 e) Set up the initial tableau and perform the first pivot.

20 f) After the pivot, what are the values of each of the variables and of the objective function? Is the current solution optimal after the first iteration? Explain.

20 g) The University administration has just decided that, independent of the cost, the University will get a fuel cell for demonstration purposes. How does this change your **formulation**? How would it affect your **answer to part d)**?

DISCUSSION

The method for examination customization is simple and straight-forward, but there are a few pitfalls to be avoided. One pitfall can occur when constructing a problem with probabilities. For example you do not want the possibility that related probabilities do not sum to one. Using customization in the face of probabilities is still possible, it just needs some extra thought and attention. Figure 5 displays an information processing problem with state events whose probabilities should sum to one. This is accomplished by specifying some probabilities in terms of the others. Since a digit of the NetID is 9 or less, dividing that number by 10 provides a positive probability less than 1 for Py (probability of 'yellow' result). Subtracting that number from 1 and then dividing the result by 2 will give the probabilities for the probabilities of blue and green (Pb & Pg).

Figure 5. Customized Information Processing Exam Problem Involving Probabilities.

2) The A/H1N1 virus outbreak (Swine Flu) has over-run the testing capacity used to determine if blood samples contain genomic evidence of A/H1N1. The CDC needs to hire more people to interpret the test results. A sample card has been developed that, when a drop of suspect blood is applied, turns either green (no A/H1N1), blue (A/H1N1 present) or yellow (inconclusive, requires re-test). The probability of each of the colors appearing is as follows:
Py=(last digit of NetID)/10 and Pb=Pg. And, Pg+Pb+Py=1
   a) What is the average information in bits of this test?
   b) The CDC wants to hire enough workers to interpret 4 million of these initial tests every day. Accounting for meetings, breaks, lunch, scrubbing in and gowning, workers will be left with 4.5 hours a day of actual processing. If reaction time (in seconds) is equal to (1st digit of NetID) + (2nd digit of NetID)x(average information in bits), how many workers will need to be hired?

Another pitfall to avoid is constructing a problem that has extremely unrealistic values. For example you do not want to construct a problem where the speed of an airplane is 10,000 miles per second or 5 miles per hour. A good rule of thumb is to calculate your NetID-based values using the highest and lowest numbers in the NetID range and see if they give reasonable values. If not, some minor manipulation can probably fix it so that the results are reasonable. For example, if the range of values is too broad, then divide the input data by a number that makes the range reasonable. Let's say you want an input value to be a person's weight in pounds. Using two digits of the NetID gives a range of 11 lbs to 99 lbs, too light unless the people are infants and young children. Using three digits gives a range of 111 lbs to 999 lbs which most likely includes values that are too heavy. Depending on the target population, more realistic ranges could be constructed by specifying "100 plus the last two digits of your NetID," which would give a range of 111 lbs to 199 lbs, or "100 plus 2*(last two digits of NetID)" which would give a range of 122 lbs to 289 lbs. It should also be kept in mind that the customization can be introduced in any numerical value, so it might make more sense to just specify a weight and introduce the customization values elsewhere in the problem. The goal is merely to get unique answers and deter answer copying.

Student Experience and Reactions

I utilize this customization technique in many courses. The first time students are exposed to this sort of customization there are always a few students who are confused by it. To clear up any potential confusion, during the first exam I will introduce the customization by directing students' attention to the problem at the beginning of the exam period. I will ask a student for their NetID and demonstrate how that translates into input data for the problem. I will also note that the input data will be different for each student based on their NetID, and also point out that this means everyone should get a different correct answer for the problem so copying from another student is futile.

Students seem to readily accept these customizations. They understand that customization reduces opportunities for academic dishonesty and realize that it does not really make the problems any harder to solve.

Every once in a while, a student will interpret the question differently from what was intended. For example if the NetID is 1234 and I write that the datum is "YZ, the last two digits of your NetID," some students correctly interpret this as '34' while others multiply 3 times 4 to get '12'. The only real issue here is to be able to understand that it was interpreted in this way so that the student can be given credit for their answer even though the number does not match that generated by the Solution Spreadsheet.

In discussion with alumni, many years after their school days, I have been told that, "Everyone knew it was pointless to try and cheat on your exams." Many students, frustrated by prior cheating by other students, expressed appreciation for the effort that made it much more difficult to copy answers. Of course this is anecdotal, but then again measuring how many students actually cheat (or are willing to self-report cheating to the instructor) is also problematic.

REFERENCES

Bowen, D. M. (2013). "On the Criticality of Communication and Team Skills for Modern Engineering Practice," Chapter 3, pp.85-102 in Agendas for 21st Century Engineers, Cambridge Scholars Publishing, Newcastle, UK.

Educational Testing Service - ETS (2019). Cheating is a Personal Foul. Educational Testing Service and Ad Council, http://www.glass-castle.com/clients/www-nocheating-org/adcouncil/research/cheatingbackgrounder.html Accessed 6/11/19.

McCabe, D. L., Butterfield K. D., & Treviño L. K. (2012). Cheating in College-Why Students Do It and What Educators Can Do about It, Johns Hopkins University Press, Baltimore, MD.

ICAI. 2019. International Center for Academic Integrity, https://academicintegrity.org/statistics/ Accessed 6/10/19.

Introna, L., Hayes, N., Blair L. & Wood E. (2003). Cultural Attitudes Towards Plagiarism: Developing a Better Understanding of the Needs of Students from Diverse Cultural Backgrounds Relating to Issues of Plagiarism. Plagiarism advisory service, Lancaster University.

DECISION SCIENCES INSTITUTE

Home-to-Work Employee Transportation with Escorts for Female Employee Safety

Srikanth Vadde
Kaizen Analytix, 2 Ravinia Dr., Suite 370, Atlanta GA 30046
Email: srikanth.vadde@kaizenanalytix.com

Anand Srinivasan
Kaizen Analytix, 261 Madison Avenue, New York, NY 10016
Email: anand.srinivasan@kaizenanalytix.com

ABSTRACT

Safe transportation for employees is a perk many organizations are offering to mitigate women abuse in public transportation. This research extends the home-to-work employee transportation problem by adding two constraints: (a) an escort must be present if a female passenger is alone in the vehicle; (b) a ride-time threshold for all passengers. Route generation heuristics for inbound (home to work) and outbound (work to home) scenarios are developed. Even distribution of passengers across routes and minimal use of escorts are the hallmarks of the proposed solution. Results from representative datasets provide valuable insights to employee transportation providers.

KEYWORDS: Home-to-work employee transportation, Route optimization, Minimal spanning tree, Route improvement heuristics, Escorts for female employee safety

INTRODUCTION

Empowering women's mobility is a primary concern for many societies around the world. Numerous studies show the rampant nature of this abuse in public and private transportation. A report by the World Health Organization states that 35 percent of women globally undergo some form of abuse or physical violence while commuting (https://www.who.int/news-room/fact-sheets/detail/violence-against-women). Women often face misbehavior, abuse, and harassment from male drivers in public transportation vehicles. For example, in Mumbai, 95% of women reported abuse in public transportation; 23% of them at a bus stop and 46% inside a bus (http://www.aksharacentre.org/ht2/). Unfortunately, the abuse creates a much deeper psychological impact beyond a woman's physical and emotional health – it impacts their ability to enter the workforce. A study conducted by International Labor Organization, states that in Delhi alone, 27% employed and 13% of non-working women see safe mobility as a hindrance to joining the workforce (http://www.aksharacentre.org/wp-content/uploads/2017/01/Akshara-Transport-Report.pdf). Many organizations are realizing the fact that female employee productivity is directly related to their well-being, sense of security and happiness. Several companies are already providing ride sharing services as benefits to employees to ensure their commute to work is safe, especially at night (https://quickride.in/, https://www.onpoint-us.com/services/workforce-transportation/, https://www.gosafr.com/about-safr/), while others are reorganizing company bus routes so that women are never left alone on the bus. Potential safety risks for a female while travelling in a ride-share vehicle are (a) abuse by the male-driver (b) abuse by the fellow male-riders. Escort services are offered by transportation provides to protect female employees from physical abuse and provide a safe environment during the journey. Trained drivers can also act as escorts, but it is usually preferred for drivers to focus on driving and leave the escort duties to someone else. Escorts are usually trained and, in some cases, armed. Many transportation providers are now ensuring at no point women employees are alone in the vehicle either with other male co-passengers or the driver (https://www.routematic.com/safety-measures-for-female-employee-transportation/).

The problem addressed in this work is concerning an Employee Transportation Provider (ETP) who intends to generate efficient vehicle routes to transport employees, from their home to work and vice versa. The ETP provides female employee safety by ensuring that they are always accompanied by an escort if they are the only passenger aboard the vehicle (with the driver). Currently many ETPs are facing challenges and deterrents to operate successfully and profitably. The key challenges are highlighted below:

a) *Dynamic nature of employee shifts*: Employee shift changes, which means different work hours and work start times, happen frequently and mandate rerouting on a short notice.
b) *Limited vehicle capacity*: Capacity constrained vehicle route planning where several constraints must be satisfied. For example, determining the number of vehicles needed to pick-up (or drop-off) employees at the same location such that their capacity is fully utilized.
c) *Forecast driver availability*: Forecasting the number of drivers needed ahead of time such that efficient driver staffing can be performed based on their availability.
d) *Forecast escort availability*: Forecasting the number of escorts needed to accompany female employees and matching that to their availability.
e) *Escort pickup and drop-off locations*: When and where to pick-up or drop-off escorts along a route.
f) *Ensuring on-time arrival to work*: Satisfying the time constraints for pickup and drop-off along with vehicle capacity constraints is challenging.
g) *Last minute requests*: Ability to alter routes to accommodate last minute cancellations and additions on short notice.
h) *Lack of good route planning tools*: Efficient route generation is very hard problem to solve for a route planner. Awareness of the road patterns in the area and its traffic patterns is needed to plan a route. Also, route generation software solutions are costly and may not satisfy all the constraints the ETP has to honor.
i) *Large volume of pickups and drop-offs*: The sheer volume of employees increases route planning time.
j) *Traffic conditions*: Unpredictable traffic conditions can disrupt existing route plans and may require real-time rerouting to a different vehicle or adding an extra vehicle.
k) *Inbound and outbound trips*: Difficulty in routing early morning trips and late-night trips back to home because of limited escort availability and unsafe conditions for female riders.

Plagued by the above problems many new ETPs are hesitant to enter the market. ETP profit margins are affected due to: (i) over-staffing escorts and drivers because of the inherent uncertainty around how many employees are transported per shift; (ii) high fuel costs because of poorly designed routes; (iii) revenue loss from paltry employee service rate. Another important aspect in routing ETPs that is usually considered is whether to originate routes from a central depot or the last known location of the vehicle. In the inbound scenario where employees are picked up from their home, many ETPs usually allow drivers to start from their last know location such as their home. Whereas in the outbound scenario, where employees are transported back to their home, many ETPs have vehicles starting from a centralized depot location. Also, many ETPs have poor data collection such as taking requests over telephone and text messages, and poor data storage methods like use of spreadsheets and handwritten notes. Such methods only add to the woes of ETPs. Moreover, route optimization is a NP-hard problem which means designing truly optimal routes is not possible in finite time. Hence, heuristic and meta-heuristic approaches are employed to solve the route optimization problem.

The main objective of this research is to design efficient inbound and outbound vehicle routes for an ETP. The goal is to find routes with the least total route distance under these constraints (a) employee pickup and drop-off is with a certain time window; (b) number of employees in a vehicle must not exceed capacity; (c) an escort must be present if a female employee is the only passenger in the vehicle; (d) ride-time of employees cannot exceed a certain threshold. Furthermore, the number of escorts needed is reduced by ensuring female passengers are not alone in the vehicle.

RELATED WORK

Perugia et al. (2011) studied the HWTP at ENEA Casaccia Research Center in Rome with fair route generation for all employees, acceptable to the company management and worker unions. Their objective was to determine routes such that total costs are minimized and quality of service, as perceived by the employee, is maximized. Route fairness was imposed by adhering to time windows on the arrival time of a bus at a stop. The problem was modeled as a multi-objective generalized vehicle routing problem with time windows (GVTPTW) which was solved using a novel incremental Tabu search algorithm. The next closest work related to HWTP is by Leksakul et al. (2017) who addressed the employee transportation for a large industrial facility in Thailand framing it as a SBRP. Their problem was to design efficient bus routes to pick-up workers who live far away from a large-scale factory. They solved the problem in two stages: (1) bus stops are allocated based on employee home location and proximity to the bus stop using cluster-based methods; (2) an ant-colony-optimization approach was used to find the optimal route between all proposed bus stop locations. Apart from the work of the above two authors there is hardly any research work related to HTWP. Hence, problems that have some commonality with HTWP are discussed below.

HWTP is related to the school bus routing problem (SBRP) which aims to design optimal routes to transport students, who board the bus near the vicinity of their home to get transported to their school. Constraints of the problem are (a) number of students in the bus cannot exceed its capacity, (b) upper threshold on the maximum ride time of a student in a bus, (c) students must reach school before school start time, (d) upper limit on the number of student at stops. Desrosiers et al (1981) decomposed SBRP into five steps: data preparation, bus stop selection (student assignment to stops), bus route generation, school bell time adjustment, and route scheduling. Data preparation loads all the necessary input data such as the locations of student homes, school, bus depot, and the origin-destination matrix. In the bus stop selection step, students are assigned to bus stops using techniques such as location-allocation-routing and allocation-routing-location which use the maximum distance from home to bus stop as the constraint (Laporte, 1988; Bowerman et al., 1995; Bodin and Berman, 1979; Dulac et al., 1980; Chapleau et al., 1985). In bus route generation, mainly two techniques are used to generate routes – “route-first, cluster-second” and “cluster-first, route-second” (Laporte and Semet, 2002). In school bell time adjustment, the objective is to determine the school starting and ending time such that the number of students served is maximized by the same bus and number of buses is minimized (Desrosiers et al., 1981, 1986; Bodin et al., 1983; Fügenschuh, 2009). In route scheduling, a sequence of successive routes is determined with the exact starting and ending time of each route (Graham and Nuttle, 1986; Braca et al., 1997; Li and Fu, 2002; Spada et al. 2005).

The current work introduces two constraints to the HTWP – escorts to accompany female employees; and a ride-time constraint. The current has some commonalities with the dial-a-ride problem. Dial-a-ride-problem (DARP) is the transportation of people between an origin and destination with the following constraints (a) pickup at the origin stop and drop-off at the destination must occur within a certain time window, (b) there is an upper limit on the maximum ride time of a person in the vehicle, (c) satisfy level-of-service requirements such as door-to-door service, wheelchair assistance, etc. (Cordeau and Laporte, 2003, 2007; Parragh et al., 2008; Ho, 2018; Molenbruch et.al., 2017a, 2017b, 2017c, 2017d). Unlike in DARP where the driver assists passengers, in the present work escorts can be considered as service providers of safety for female passengers. Presence of a ride-time constraint is another common feature with DARP.

This research is unique in three aspects:

1. It solves both inbound and outbound scenarios unlike many SBRP studies that are only focused on solving the inbound scenario (Braca et al., 1997; Li and Fu, 2002). The outbound or the afternoon problem, where students are dropped to their bus stops, received very little attention in the SBRP literature.
2. It extends the HWTP to include two constraints – escort for women employees and a ride-time limit.
3. It extends HWTP, by considering different starting locations for vehicles in the inbound scenario.

PROBLEM DESCRIPTION

A corporate company is providing safe transportation services as a benefit to its employees to ensure their wellbeing. It outsources this job to a third-party ETP, who must transport employees from their home to the office and vice versa in a metropolitan area. The ETP aims to design optimal routes to maximize the number of employees transported, reduce operational costs, improve employee satisfaction and increase profit margins. ETP executes two types of trips – inbound and outbound. In the *inbound trip*, employees are transported from their home to work; in the *outbound trip* they are transported from work to home. Inbound and outbound scenarios are depicted in Figure 1 and 2 respectively.

The following assumptions are made to limit the scope of this work:

- ETP has fixed number of vehicles.
- Drivers report their location to the ETP prior to inbound trip planning. Vehicles are assumed to be with the driver in this case.
- Drivers and vehicles are stationed at the office before the outbound trip starts.
- Seat capacity of each vehicle can be different.
- Vehicles do not return to their starting location at the end of a route.
- Work location is same for all employees.
- Employees live in a city where there are road and traffic restrictions.
- Employees living in the same residential premises are transported together if possible.
- Single load i.e., male or female employees; and mixed load i.e., a mix of female and male employees is possible.
- Special need employees are not considered, only general employees are assumed.
- Finite number of shifts are serviced in a day. For example, an organization can operate, 3 shifts of 8-hour durations.
- Starting and ending time of an employee shift is provided.

- The shifts do not overlap with each other (i.e., the starting and ending time of a shift does not overlap with another shift).
- Employees are contacted by email or text message about their pickup time for the inbound trip and the drop-off time for the outbound trip.
- Escort pickup location is known before route planning.
- Escort pickup location can either be their home or anywhere between the office and employees' home.
- Escort scheduling and optimizing the number of escorts across all routes is not considered.
- Escorts are either male or female security personnel who are trained and skilled to protect the female employee during the ride.
- Escort drop-offs is not considered in route planning, i.e., their drop-off is not part of the route sequence.

Inbound Model

Given a fixed number of vehicles, what is the optimal routing sequence to pick up employees and drop them all at the same work location such that the total route distance and number of escorts are minimized?

Figure 1: Inbound scenario

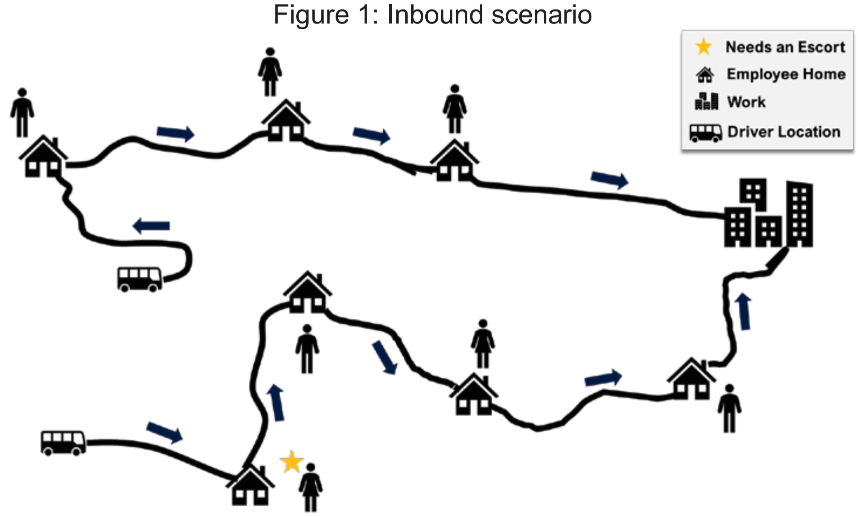


Figure 2: Outbound scenario

The constraints of this problem are:

a. employees must be picked up in a certain time window and dropped off at the office by the start of the employee shift time
b. vehicle capacity is not exceeded
c. an escort must be in the vehicle if a female passenger is the first employee to be picked up in the route
d. each pickup location is only visited once, i.e., only one vehicle services a pickup stop
e. upper limit on the number of employees transported in a vehicle (usually set by the ETP)
f. threshold limit on the ride-time of each employee in a vehicle

Escorts are provided when the female passenger is the first person in the route sequence. This service is used by many ETPs in practice. In a situation where the first person in the route is a male and the next employee in the route is a female, it is assumed to be a safe practice by the ETP because the employer has strict disciplinary rules to prevent employee harassment.

Outbound Model

Given a fixed number of vehicles, what is the optimal routing sequence to drop off employees after being picked up from the same work location such that the total route distance and number of escorts are minimized? The constraints of this problem are:

a. employees must be picked up at the end of their shift and dropped off at their home in a certain time window
b. vehicle capacity is not exceeded
c. an escort must be in the vehicle if a female employee is the last passenger to be dropped-off in the route
d. each drop-off location is only visited once
e. upper limit on the number of employees transported in a vehicle (usually set by the ETP)
f. threshold limit on the ride-time of each employee in a vehicle

Escorts are provided if the employee is the last person in the route sequence. Again, similar to the inbound case, ETP considers it safe if the penultimate passenger is a female followed by a male employee.

SOLUTION METHODOLOGY

A heuristic solution is proposed to solve the above two problem scenarios since they both are NP hard problems.

Inbound Model

The inbound problem is modeled as a VRP (Vehicle Routing Problem). Since employee actual pickup time is provided only after route planning is finished, it is safe to assume that there is a reasonably large time window for their pickup. Unlike the actual VRP problem, here vehicles are distributed across an area. The route planning algorithm is described below.

1. Get input data – employee home locations, driver locations, escort locations, vehicle-seating limit threshold, shift start time, employee ride time threshold.
2. Group employees at the same location so that they are picked up at the same time.
3. Calculate the midpoint of all employee pickup locations.
4. Calculate the distance from each vehicle's current location to the midpoint of all pickup locations. Sort the vehicles in the ascending order of this distance.
5. Calculate the distance of each employee's home location to office. Sort the employees in the ascending order of this distance.
6. If the employee pickup consists of both male and female passengers or only male passengers, then:
    a. Determine the least number of vehicles required to transport all employees while maintaining an even distribution of employees in a route. This is formulated as integer linear optimization problem where the objective function is the sum of distances between each vehicle's location and all employee pickup locations subject to vehicle available capacity constraint. The decision variable is binary, if a vehicle is assigned to an employee then the decision variable is 1 otherwise it is 0. The optimal solution is a set of vehicles that are assigned a set of employee pickups. The main purpose of calculating the number of vehicles upfront is to reduce the number of drivers and escorts required.

    If the employee pickups are only females, then:
    b. Assign an escort to a vehicle and reduce the available capacity by 1.
    c. Selection of vehicles is identical to 6 (a)
7. Create an initial route sequence for each vehicle from its corresponding assigned stops in 6 (a).
8. For each pickup stop in the sorted list of employee pickups from step 5, assign it to a vehicle such that the incremental change in route distance is at minimum. This is achieved by inserting the stop in the best possible position within a vehicle's route and recording the incremental change in route distance of the vehicle.
9. Apply the Relocate Operator on each pair of vehicle routes (Braysy and Gendreau., 2005)
10. Apply the 2-opt operator on every vehicle's route (Braysy and Gendreau., 2005)
11. Identify routes where the first pickup is a female employee and an escort is not provided.
    a. If there is available capacity in vehicle, insert the escort pickup stop that is closest to the vehicle's starting location or the first female pickup stop
    b. If the vehicle capacity is reached, rearrange the route sequence such that incremental change in route distance is least and the first pickup stop is a male employee.

Outbound Model

The outbound problem is also modeled as a VRP. Since employee actual drop-off time is provided only after route planning is finished, it is safe to assume that there is a reasonably large time window for their drop-off. Route planning algorithm when employee pickups are both male and female is described below.

1. Get input data – in all employee home locations, driver locations, escort locations, vehicle-seating limit threshold, shift end time, employee ride time threshold.
2. Group employees at the same drop-off location so that they are dropped at the same time.
3. Calculate the distance of each employee's home location to office. Sort the employees in the ascending order of this distance.
4. If the employee drop-off consists of both male and female passengers or only male passengers, then:
    a. Estimate the number of vehicles needed, ensuring an even distribution of employees in a vehicle's route using a cluster-based approach:
        i. Construct a minimal spanning tree with the drop-off stops as the nodes of the tree and the work location as the root of the tree.
        ii. Traverse the tree from the deepest node to the root node.
        iii. Assign each node traversed to a vehicle and terminate when the vehicle's capacity is reached.
        iv. The output is a set of employee pickups assigned to a vehicle.

    If only female passengers have to be dropped, then:
    b. Assign an escort to a vehicle and reduce the available capacity by 1.
    c. Estimation of vehicles needed is identical to 4(a)
5. Create an initial route sequence for each cluster in 4(a).
6. Apply the Relocate Operator on each pair of vehicle routes.
7. Apply the 2-opt operator on every vehicle route.
8. Identify routes where the last drop-off is a female employee and an escort is not provided.
    a. If there is available capacity in the vehicle, insert the escort pickup stop at the right spot in the route sequence that is before the last female drop-off stop
    b. If the vehicle capacity is reached, rearrange the route sequence such that incremental change in route distance is least and the last drop-off stop is a male employee.

COMPUTATIONAL EXPERIMENTS

An example dataset is constructed to illustrate the solution methodology because no benchmark datasets are available for the HTWP with female employee safety constraint. The dataset is adopted from a real-world situation faced by an ETP, however due to data sovereignty some data elements are masked. The results and insights from this dataset are significant for an ETP.

Inbound Scenario

4 test instances are considered here. In the first 2 test instances, the employee home locations are uniformly distributed and in the remaining 2 test instances, employee home locations form distinct clusters, see Figures 3-6. The office location is assumed to be located between employee home locations or far away from it. Driver locations and escort locations are uniformly distributed; vehicles reside at driver home locations. All vehicles have the same capacity of 10 seats. Employee gender at each stop is randomly generated for all test instances. All employees must in office by their shift start time of 7 am. They are picked up after 5 am and before 7 am. The ride-time threshold is 30 minutes. A constant vehicle velocity of 40 miles per hour is assumed and all distances are assumed to be Euclidian.

Outbound Scenario

4 test instances are considered here. The employee home and office locations are identical to the inbound scenario. Vehicles, drivers, and escorts are all present at the office location before the route start. Vehicle capacity and employee gender for all test instances is identical to the inbound scenario. The employee shift end time is 2 pm. They should be dropped before 4 pm at their home and their ride-time threshold is 30 minutes. A constant vehicle velocity of 40 miles per hour is assumed and all distances are assumed to be Euclidian and in miles.

Table 1: Input parameters inbound and outbound for test instances

Scenario	Test Instance	Number of Vehicles/ Drivers	Number of Escorts	Number of Male Employees	Number of Female Employees	Vehicle capacity
Inbound	5501	20	20	49	51	10
Inbound	5502	20	20	45	55	10
Inbound	5503	20	20	51	49	10
Inbound	5504	20	20	49	51	10
Outbound	4401	20	20	49	51	10
Outbound	4402	20	20	45	55	10
Outbound	4403	20	20	51	49	10
Outbound	4404	20	20	49	51	10

Results of the test instances are summarized in Table 2. It can be observed that

a) vehicle capacity is utilized completely in test instances 5502, 4401, 4402, and 4404.
b) maximum ride-time in a route across all test instances is below the ride-time threshold.
c) escorts were not needed in all inbound test instances because either the capacity was reached, or the route distance was at minimum by having a male employee as the first pickup passenger (see Figure 7). Two routes in outbound test instance needed escorts (see Figure 8).
d) even load of employees across all routes, which means there is a fairness among all drivers on their workload.

Table 2: Summary of optimization results for test instances

Scenario	Test Instance	Number of Routes	Avg. Stops Per Route	Total Distance (miles)	Max Ride Time (mins)	Escorts Used
Inbound	5501	10	9	922.9	20	0
Inbound	5502	10	10	687.0	20	0
Inbound	5503	10	9	1132.0	20	0
Inbound	5504	10	9	870.9	20	0
Outbound	4401	10	10	1050.0	21	0
Outbound	4402	10	10	972.9	21	0
Outbound	4403	11	9	1078.3	21	2
Outbound	4404	10	10	829.5	21	0

The rearrangement of a route to that the female is not alone in the vehicle certainly reduced the need for more escorts however it did increase the route time. In such instances the ETP should evaluate the fuel cost and escort wages and decide if having more escorts is beneficial.

CONCLUSIONS AND FUTURE WORK

This work presented a HWTP by adding a ride-time constraint and providing escorts for female passengers if they are alone in the vehicle. In the inbound and outbound scenario, a cluster-first route second approach is applied to attain an even distribution of stops across vehicle routes. A linear optimization is used to perform clustering in the inbound scenario. A minimal spanning tree is used to perform clustering in the outbound scenario. A feasible initial solution is obtained and 2-opt and relocate route improvement heuristics are applied.

As part of future work, several extensions to the current work are planned:

a) Exploring the use of variable neighborhood search, adaptive variable neighborhood search, Tabu search, and other meta-heuristics to attain better solutions
b) Including a constraint where escorts must be on a vehicle if male passengers are also traveling with female passengers.
c) Relaxing the criteria for escort drop-off – they can be dropped off when no female passengers are present on the vehicle.
d) Including a constraint where the ETP operates exclusive vehicles for female passengers on certain routes.

Figure 3: Inbound test instance 1.

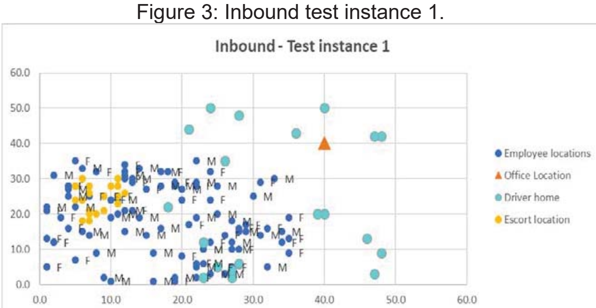


Figure 5: Inbound test instance 3.

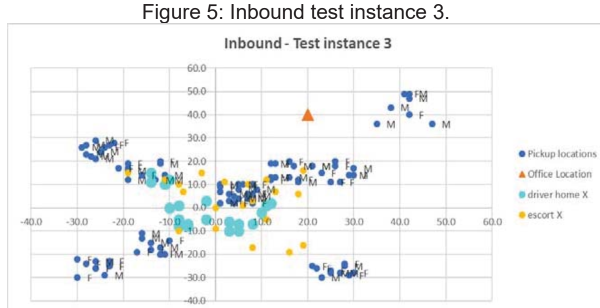


Figure 4: Inbound test instance 2.

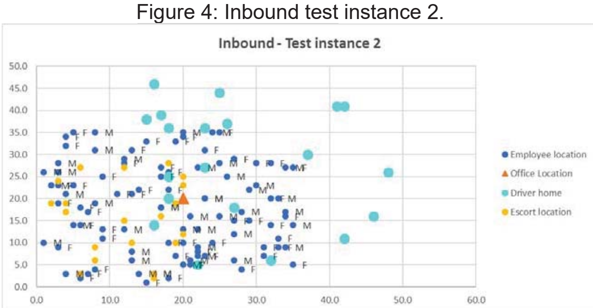


Figure 6: Inbound test instance 4.

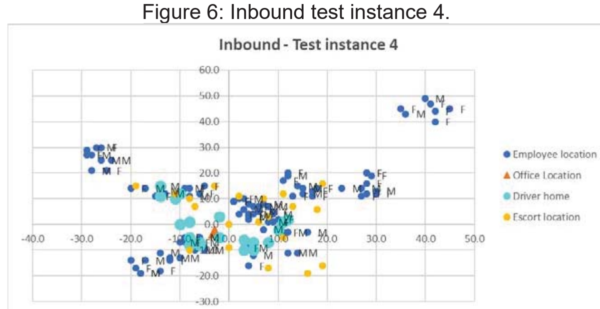


Figure 7: Inbound test instances depicting employee gender in route sequence

Test Instance	Route ID	seq-1	seq-2	seq-3	seq-4	seq-5	seq-6	seq-7	seq-8	seq-9	seq-10
5501	1	M	F	F	F	M	M	M	M	M	F
5501	2	M	F	F	M	M	M	F	F		
5501	3	M	F	F	F	M	F	F	F	F	
5501	4	M	M	M	M	M	F	F	M	M	F
5501	5	M	F	M	F	F	M	M	F	M	M
5501	6	M	M	M	F	F	F	F	F	M	
5501	7	M	M	F	F	F	M	F	M	M	
5501	8	M	M	F	M	F	F	M	F		
5501	9	M	F	M	F	M	F	M	F	M	F
5501	10	M	M	F	M	F	F	M	F	F	F
5502	1	M	F	F	F	F	M	F	F	M	
5502	2	M	F	F	F	F	F	M	M	F	F
5502	3	M	F	M	F	F	M	M	M	M	
5502	4	M	F	M	F	F	M	F	F	F	F
5502	5	M	F	M	M	F	F	M	M	F	M
5502	6	M	F	M	F	M	M	M	F	F	F
5502	7	M	F	F	F	F	F	M	M	M	
5502	8	M	F	F	F	M	M	M	M	F	
5502	9	M	F	F	F	F	M	F	M	M	F
5502	10	M	F	F	F	M	M	M	F	F	F
5503	1	M	F	F	M	F	M	F	M	F	F
5503	2	M	F	F	M	F	F	F	M	M	
5503	3	M	F	F	F	F	M	M	F	M	F
5503	4	M	F	F	F	M	M	F	M	F	
5503	5	M	M	F	F	F	M	M	M		
5503	6	M	F	F	F	F	F	M	F	M	
5503	7	M	F	F	M	M	M	M	M	M	F
5503	8	M	F	F	M	F	M	F	F	M	F
5503	9	M	M	F	F	F	F	M	M	F	
5503	10	M	M	M	M	M	M	M	M	F	
5504	1	M	F	F	M	M	F	F	M	F	M
5504	2	M	M	F	F	F	F	M	M	F	
5504	3	M	F	M	F	M	F	F	F	M	
5504	4	M	F	F	F	F	M	F	M	F	
5504	5	M	F	M	F	M	M	M	F	F	M
5504	6	M	M	F	M	M	M	M	M	F	F
5504	7	M	F	M	M	F	M	M	F	F	F
5504	8	M	F	F	M	M	F	F	M	M	
5504	9	M	F	F	F	F	F	M	F	F	
5504	10	M	F	M	M	M	F	F	M	M	

Figure 8: Outbound test instances depicting employee gender in route sequence

Test Instance	Route ID	seq-1	seq-2	seq-3	seq-4	seq-5	seq-6	seq-7	seq-8	seq-9	seq-10
4401	1	M	M	M	F	F	F	F	F	F	M
4401	2	F	M	M	M	F	F	M	M	F	M
4401	3	M	M	M	M	F	M	M	M	F	M
4401	4	M	M	M	M	F	F	M	F	M	M
4401	5	F	F	M	M	F	F	F	F	F	M
4401	6	F	M	F	F	F	F	M	M	M	M
4401	7	M	F	F	M	M	F	F	F	F	M
4401	8	M	F	F	F	F	M	F	F	F	M
4401	9	M	F	M	F	F	M	F	F	F	M
4401	10	F	F	M	F	M	M	M	F	F	M
4402	1	M	F	M	F	F	F	F	F	F	M
4402	2	F	M	F	F	M	F	M	F	M	M
4402	3	M	M	M	F	F	M	M	F	M	M
4402	4	F	F	M	M	F	F	M	M	F	M
4402	5	M	F	F	M	M	F	F	M	F	M
4402	6	F	F	M	F	M	F	F	F	F	M
4402	7	M	M	M	F	M	F	M	F	F	M
4402	8	M	F	M	F	F	F	F	F	F	M
4402	9	M	F	M	F	M	F	F	F	F	M
4402	10	M	M	F	F	F	F	M	F	F	M
4403	1	ESCORT	M	F	F	M	F	F	F	F	F
4403	2	M	F	M	F	M	M	F	F	M	M
4403	3	M	M	M	M	M	F	F	F	F	M
4403	4	F	F	F	F	M	M	M	F	M	M
4403	5	M	M	M	M	M	M	F	M	M	
4403	6	M	F	F	F	F	F	F	F	F	M
4403	7	F	F	M	M	M	F	F	F	F	M
4403	8	F	M	F	M	M	M	F	F	F	M
4403	9	F	M	M	F	F	F	F	M	F	M
4403	10	M	F	F	M	M	M	M	M	M	M
4403	11	ESCORT	M	F							
4404	1	M	M	M	M	F	F	F	F	F	M
4404	2	M	F	M	F	M	M	F	F	F	M
4404	3	F	F	M	F	F	F	F	F	F	M
4404	4	M	F	F	F	F	F	M	M	F	M
4404	5	M	M	F	M	M	F	F	F	F	M
4404	6	M	M	F	F	F	M	M	M	F	M
4404	7	F	F	M	M	M	F	F	M	F	M
4404	8	M	F	M	M	F	M	F	M	M	M
4404	9	F	M	M	M	M	M	F	F	F	M
4404	10	F	F	M	F	M	M	F	F	F	M

REFERENCES

Bodin, L.D., Berman, L., 1979. Routing and Scheduling of School Buses by Computer. *Transportation Science* 13 (2), 113–129.

Bodin, L.D., Golden, B., Assad, A., Ball, M.O., 1983. Routing and Scheduling of Vehicles and Crews: the State of the Art. *Computers and Operations Research* 10, 63–211.

Bowerman, R., Hall, B., Calamai, P., 1995. A Multi-Objective Optimization Approach to Urban School Bus Routing: Formulation and Solution Method. *Transportation Research Part A,* 29 (2), 107–123.

Braca, J., Bramel, J., Posner, B., Simchi-Levi, D., 1997. A Computerized Approach to the New York City School Bus Routing Problem. *IIE Transactions* 29, 693–702.

Braysy O., and Gendreau., M, (2005). Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms. *Journal of Transportation Science,* 29 (1), 104–118.

Chapleau, L., Ferland, J.A., Rousseau, J.-M., 1985. Clustering for Routing in Densely Populated Areas. *European Journal of Operational Research* 20 (1), 48–57.

Cordeau, J.-F., and Laporte, G. (2003). A Tabu Search Heuristic for the Static Multi-Vehicle Dial-a-Ride Problem. *Transportation Research Part B: Methodological*, 37(6), 579–594.

Cordeau, J.-F., and Laporte, G. (2007). The Dial-a-Ride Problem: Models and Algorithms. *Annals of Operations Research*, 153(1), 29–46.

Desrosiers, J., Ferland, J.A., Rousseau, J.-M., Lapalme, G., Chapleau, L., (1981). An Overview of a School Busing System. In: Jaiswal, N.K. (Ed.), *Scientific Management of Transport Systems*. North-Holland, Amsterdam, pp. 235–243.

Desrosiers, J., Ferland, J.A., Rousseau, J.-M., Lapalme, G., Chapleau, L., (1986). TRANSCOL: A Multi-Period School Bus Routing and Scheduling System. *TIMS Studies in the Management Sciences*, 22, 47–71.

Dulac, G., Ferland, J.A., Forgues, P.A., 1980. School Bus Routes Generator in Urban Surroundings. *Computers and Operations Research*, 7 (3), 199–213.

Fügenschuh, A., 2009. Solving a School Bus Scheduling Problem with Integer Programming. *European Journal of Operational Research*, 193 (3), 867–884.

Graham, D., Nuttle, H.L.W., (1986). A Comparison of Heuristics for a School Bus Scheduling Problem. *Transportation Research B*, 20 (2), 175–182.

Ho, C. S, Szeto, Y. W, Leung, Y. M. J, Petering, M, and Tou, H. W. T, (2018). Typology and Literature Review for Dial-a-Ride Problems. *Transportation Research Part B: Methodological*, 111, 395–421.

Laporte, G., (1988). Location-Routing Problems. In: Golden, B.L., Assad, A.A. (Eds.), *Vehicle Routing: Methods and Studies*. North-Holland, Amsterdam, pp. 163–198.

Laporte, G., Semet, F., 2002. Classical Heuristics for the Capacitated VRP. In: Toth, P., Vigo, D. (Eds.), The Vehicle Routing Problem. *SIAM*, Philadelphia, PA, pp. 109–128.

Leksakul, K., Smutkupt., U, Jintawiwat., R., and Phongmoo., S, (2017). Heuristic Approach for Solving Employee Bus Routes in a Large-Scale Industrial Factory. *Advanced Engineering Informatics,* 32, 176-187.Li, L., and Fu, Z., 2002. The School Bus Routing Problem: A Case Study. *Journal of the Operational Research Society* 53, 552–558.

Molenbruch, Y, Braekers, K, Caris, A, (2017a). Benefits of Horizontal Cooperation in Dial-a-Ride Services. *Transportation Research Part E: Logistics*. *Transp. Rev*. 107, 97–119.

Molenbruch, Y, Braekers, K, Caris, A, (2017b). Operational Effects of Service Level Variations for the Dial-a-Ride Problem. *Central European Journal of Operations Research*, 25 (1), 71–90.

Molenbruch, Y, Braekers, K, Caris, A, Berghe, G.V., (2017c). Multi-Directional Local Search for a Bi-Objective Dial-a-Ride Problem in Patient Transportation. *Computers and Operations Research* 77, 58–71.

Molenbruch, Y, Braekers, K, Caris, A, (2017d). Typology and Literature Review for Dial-a-Ride Problems. *Annals of Operations Research,* 259 (1-2), 259–325.

Parragh, S.N., (2011). Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem. Transp. Res. Part C: Emerg. Technol. 19 (5), 912–930.

Perugia, A., Moccia, L., Cordeau, J.-F., and Laporte, G (2011). Designing a Home-to-Work Bus Service in a Metropolitan Area. *Transportation Research Part B: Methodological*, 45 (10), 1710–1726.

Spada, M., Bierlaire, M., Liebling, Th.M., 2005. Decision-Aiding Methodology for the School Bus Routing and Scheduling Problem. *Transportation Science* 39, 477–490.

DECISION SCIENCES INSTITUTE

Evaluation of a Continuous Pipe Production System for Existing and Optimized Configuration

Md. Shahriar J. Hossain
Northwestern State University
Email: hossains@nsula.edu

Bhaba R. Sarker
Louisiana State University
Email: bsarker@lsu.edu

ABSTRACT

This paper is an extension of a detailed research which is primarily conducted to make an overall improvement of a steel pipe manufacturing system. In order to achieve that goal, four distinct problems namely, buffer capacity, tool magazine size and reload timing, pipe sizing, and inspection and rework problems, in the system are solved optimally. To validate that the overall effectiveness of the system significantly improves while the optimum solutions are implemented in the production floor, the whole system is simulated with existing and optimum configurations. The comparison is presented in terms of overall equipment effectiveness.

KEYWORDS: Overall Effectiveness, Simulation, Continuous Production.

INTRODUCTION

This paper presents the numerical application of some locally optimized operational solutions in a continuous steel pipe manufacturing system. The optimum solutions for four different operational problems are compared with existing non-optimal situations. The optimum solutions to those four problems are generated based on some data imitating a practical industrial environment. The first problem was focused in obtaining an optimum buffer capacity for the steel coil accumulator (Hossain and Sarker 2018a). The second problem focused in finding the optimum tool order quantity in addition to the tool magazine size and magazine reload timing (Hossain and Sarker 2018b, 2018c). The cutting length (size—standard/substandard) of the pipes and the rework policy are optimized in the third problem (Hossain and Sarker 2017). Wherein, the fourth problem of this research dealt with off-line inspection-rework facility to find the optimum number of servers and space for the inspection and rework stations (Hossain and Sarker 2018d). Subsequently an additional evaluation system called weighted *overall equipment effectiveness* (OEE) is developed, based on the monetary loss based regression method (Hossain and Sarker 2016, 2019), to measure the economic performance of the individual or composite performance of the system. It is expected that, OEE of the system is improved when the optimum solutions for these four problems are fully implemented in the production floor. In order to check this expectation, the whole system is simulated for two cases: (1) with the existing configuration as well as (2) with the improved or optimum configuration.

Non-disclosure of data proprietorship: This paper presents a detailed information of the simulation results that mimics a steel pipe manufacturing system. The system parameters used in this paper do not represent the original data of any industrial setup, but the data chosen here capture a close picture of the aforementioned industrial sector. The name of the industry is not disclosed to preserve the confidentiality of its proprietorship.

Literature Survey

The recent available literatures on buffer capacity optimization are either evolved from approximation of discrete systems [Alvarezvargas et al. (1994), Rajaram and Tian (2009), Bierbooms et al. (2013), Ozdogru and Altiok (2015)], or focused in modeling the system as a fluid flow. None of the available researches is specifically oriented with the situation of a coil accumulator in continuous steel pipe manufacturing systems. Hence, the first optimization problem was oriented in determining the optimum buffer capacity.

In the second operational problem a multiple-tool magazine system is proposed for the steel pipe manufacturing system. Some research works are reported in the literature, that deals with tool life prediction [Karandikar et al. (2012), Yao and Chien (2014), Wang and Gao (2015)], or tool replacement scheduling [Lamond and Sodhi (2006), Sodhi and Sarker (2003), Hippalgaonkar and Shin (2011)]. However, a generalized model for multiple-tool magazine system in terms of magazine reload timing, magazine size, tool order quantity, and generalized tool-life distribution, still absent in the available literatures. In this paper, an optimization model is evaluated that deals with the impact of tool magazine system on the cost savings.

Some research works are found, which deal with the one-dimensional slitting problems [Sarker (1988), Hsieh (1997)], knapsack problems [Yan (2003), Lee et al. (2013), Della Croce et al. (2017), Feng et al. (2017), Furini et al. (2017)], or discrete rework optimization [Selim and Al-Zu'bi (2011), Hossain and Sarker (2016a, 2016b)]. However, the on-line defect detection and rework policy in a continuous process, as well as the sizing policy for the pipes cut are not precisely addressed in the available research works. The third operational problem here, is concerned about determining the optimum pipe cutting length and pipe rework policy.

The inspection and rework queues, in the fourth problem of pipe manufacturing, depend several system parameters and constraints. Though several literatures on the production line queues are reported [Laghaie et al. (2012), Ravid et al. (2013)], the appropriate model for the intertwined inspection-rework queuing network in the pipe manufacturing system, is not specifically exemplified in the available researches. A specific model for determining the number of inspection/rework server(s), with the appropriate system parameters for a steel pipe manufacturing line, is presented in this research paper with a practical field data.

Finally, a suitable method for evaluating the effectiveness of the pipe manufacturing system is implemented in this paper. The old measures [Raouf (1994), Wudhikarn (2010), Yuniawan et al. (2013)] of weighted overall equipment effectiveness (OEE) do not provide a concrete method for determining the weights in the expression. Moreover, the old performance measures do not consider the weights in terms of cost parameters. Hence, this research is followed with a new approach of measuring the weighted OEE that suits with the continuous pipe manufacturing system. The overall effectiveness's (OEE) of the pipe manufacturing system, for both existing and improved configurations, are evaluated by comparing the WOEE (weighted overall equipment effectiveness) and PEE (production equipment effectiveness) methods. The newly developed MLBR (monetary loss based regression) method presented in by Hossain (2019), is used for finding the WOEE and PEE. The numerical example taken here, is solved for obtaining optimum solutions, following the approaches described by Hossain and Sarker (2017-2018).

Notations

Some notations listed below are used in this paper to explain the system.

T_{br}	Buffer runout duration in a cycle (hours)
T_c'	Coil feed cycle time (hours)
$T_{cr,i}$	Remaining cycle time in Day-i , for next day use (hours)
$T_{r,i}$	Remaining tool life in Day-i , for next day use (hours)
n_t	Number of magazine reloads in a day
l_1^s	Scrapped pipe length from the online rework-scrap decision (feet)
n^{st}	Number of standard size pipes
n^i	Number of inspected pipes in the off-line inspection-rework facility
n^{ex}	Number of pipe external failures that customer identifies
n^r	Number of reworked pipes in the inspection-rework facility
n^s	Number of scrapped pipes in the inspection-rework facility
l^s	Total length of the scrapped pipe in a day (feet)

The parametric values of the system for the four optimization problems, the existing configuration of the line and the corresponding optimum solutions are presented in the next section.

SYSTEM PARAMETERS

Two different scenarios are evaluated for comparison purposes. In one scenario the line is configured with the traditional non-optimum parameters; this situation is termed as *existing* configuration. In another scenario, the line is configured with the optimum solutions from the four optimization problems. The system described with the optimum solutions is termed as *improved* configuration. The numerical values of the parameters, corresponding to the (1) buffer capacity, (2) tool life, (3) pipe sizing and rework polity, and (4) inspection-rework facility configuration problems are listed in Table 1, 2, 3 and 4, respectively.

System Parameters Relevant to Buffer Capacity

The coil feeding system in the pipe manufacturing line consists of a coil joining station and a coil accumulator. The joining time (T_J) follows normal distribution with mean $\overline{T}_J$ = 0.85 hours and standard deviation σ_T = 0.35 hours. All other parameters which are involved in the feeding stage, are listed in Table 1. Here, c_B indicates the operational cost for the coil accumulator, α denotes the acceleration of strip input to the coil accumulator, c_{td} indicates the line stoppage or downtime cost, h_B defines buffer strip holding cost at the coil accumulator, D_c expresses the demand of coils in a year, and L_c is the average length of the steel strip coil. The strip output rate from the coil accumulator is the same as the actual rate of material flow in the line or the rate of pipe production λ .

Table 1. System parameters involved in the coil feeding system.

c_B \$/feet/year	$\bar{T}_J$ hours	σ_T hours	α feet/hour^2	λ feet/hour	c_{td} \$/hour	h_B \$/feet/year	D_c coils/year	L_c Feet
100	0.85	0.35	3600	300	600	250	700	800

System Parameters Relevant to Tool Life

In the pipe manufacturing industries several cutting tools are used. Flash trimming tool is one of them which is used for removing extra weld metals from the welding zone. The manager plans to purchase an automated tool magazine for handling the flash trimming tools. In the market, several tool magazines are available with variable sizes. The tool magazine capacity ranging from n = 2 to 15 are considered here, among which the optimum size for the tool magazine has to be selected. The distribution of the life (T) of a flash trimming tool is given as $T \sim N$(6, 1.5^2) hours. Other relevant information about this tool magazine system are listed in Table 2 which includes the operating cost for an *n*-tool magazine (c_o), service rate of the tool (μ), tool storage cost at tool crib (h), scheduled tool replacement cost (c_s), unscheduled tool replacement cost (c_u), price of a tool (c_p) and the tool ordering cost (A_o).

Table 2. System parameters involved in the flash trimming tool magazine system.

c_o \$/tool/year	μ feet/hour	h \$/tool/year	c_s \$/magazine	c_u \$/magazine	c_p \$/tool	A_o \$/order
$10n$	500	30	10	900	20	60

System Parameters Relevant to Pipe Sizing and Rework Polity

At the end of the manufacturing line the pipes are cut into standard and/or substandard lengths. During production some welding defects occur randomly. The linear distance between two consecutive defects (L) follows exponential distribution with a mean rate of defects δ. Here, L can also be defined as the defect free length. All other relevant information about the on-line inspection facility and the pipe sizing system are listed in Table 3. In this table, L_s and L_b represents the desired length of a standard and substandard pipe, respectively. On the other hand, P_o and P' are denoted for the standard size and substandard size pipes, respectively.

The cost parameters in this problem are pipe cutting cost (c_t), fixed cost (C) and variable cost (c) of the pipe manufacturing line, online (c^{on}) and offline (c^{off}) repair costs for a defect, and the scrapping cost (c_s). Here, $L_b^{\max}$ is the maximum demand of substandard pipes, which limits the amount of total substandard pipes.

Table 3. Relevant numerical information about the pipe sizing and online inspection.

L_s feet	L_b feet	P_o \$/feet	P' \$/feet	c_t \$/cut	C \$/hour	c \$/feet	c^{on} \$/defect	c^{off} \$/defect	c_s \$/feet	$L_b^{\max}$ feet/hour	δ defects /hour
40	9	180	350	15	1000	100	30	20	10	50	3

System Parameters Relevant to Inspection-Rework Facility Configuration

Now, after partitioning the pipes into standard and/or substandard lengths, some pipes are randomly selected for ultrasonic inspection. Any defective pipe that is repairable by grinding or polishing operation is sent to the rework station, and the unrepairable pipes are scrapped. The important parameters which influence the optimal decision on the number of inspection/rework servers are noted in Table 4. The parameters listed in this table are defined for the pipes holding costs h^r and h^{os} at the regular and overstock holding area; the required floor spaces a_p , a_1 and a_R for a pipe, an inspection server and a rework server; the variable costs c_1 and c_R for inspection and rework of a pipe; the fixed costs C_1 and C_R for maintaining an inspection and a rework servers; service rates μ_1 and μ_R at the inspection and rework servers, all respectively. The other parameters are available floor space A_T , pipe arrival rate at the inspection-rework facility λ' , and repairable and unrepairable defect proportion p_1' and p_1 , respectively.

Table 4. Important parameters involved in the inspection-rework facility.

Parametric description	h^r \$/pipe/hour	h^{os} \$/pipe/hour	a_p (feet²)	a_1 (feet²)	c_1 \$/pipe	C_1 \$/hour	μ_1 pipes/hour
	3.50	5.50	200	2000	100.00	120.00	2.55
A_T (feet²)	λ' pipes/hour	p_1'	p_1	a_R (feet²)	c_R \$/pipe	C_R \$/hour	μ_R pipes/hour
17,000	5	0.0820	0.050	4000	120	200	0.5

EXISTING CONFIGURATION OF THE SYSTEM

The existing state of the steel pipe manufacturing system is discussed here from the perspective of the four optimization problems. The system is currently equipped with a coil accumulator. The coil feeding system is designed considering the average coil joining time ($\bar{T}_J$). A minimum buffer capacity level, $B_{\min} = \lambda^2/\alpha + \bar{T}_J\lambda$ = 280 feet, is used in the system. On the other hand, the current system is equipped with single-tool magazine for all kind of cutting tools. Thus, for the flash trimming operation, a single-tool holder is used and the tool is replaced once it fails. For a single-tool magazine system the optimum tool reload timing is found as τ_1^* = 10.97 hours. Reloading a new tool takes T^R = 0.5 hours. In other words, if a tool fails or about to exceed 10.97 hours of active operation, it should be replaced, and the replacement process takes on an average 0.5 hours. In place of a single-tool holder for the flash trimming tool, a multiple-tool magazine is proposed to be used.

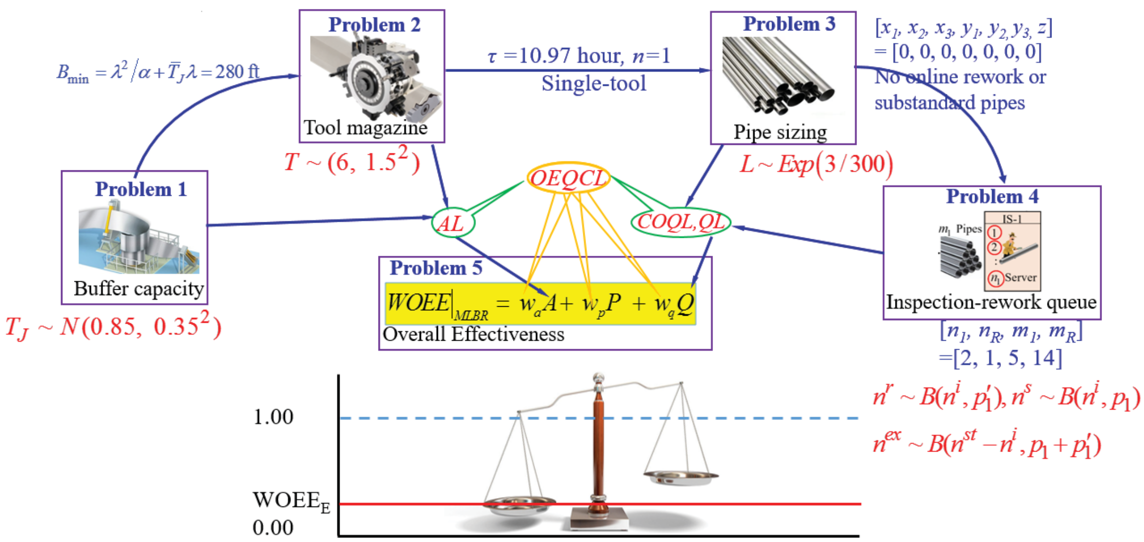


Figure 1. Simulation model for the existing system.

In the existing condition the pipe manufacturing line is not facilitated with an online inspection device. In addition to that, no online rework facility is available in the system. As a result, any defect can only be detected after the pipes are cut into standard pieces. In other words, no rework decision can be taken online, and no substandard size pipe can be cut directly from the continuous line. After cutting the pipes into standard lengths it goes to the inspection-rework facility and that facility is equipped with 2 ultrasonic testing machines (n_1 = 2) that detects hidden welding defects, and a rework station (n_R =1). Rework operations includes cutting off the defective portion of the pipe if the defect falls within 5 feet from any end of that pipe. Some surface defects are removed by grinding or polishing the pipe. Both of these rework operations are currently done in a single off-line rework station. The existing system configuration is graphically shown in Figure 1.

PROPOSED IMPROVED CONFIGURATION OF THE SYSTEM

The pipe manufacturing line is proposed to be improved with some local optimum solutions. The corresponding optimum solutions for (1) buffer storage, (2) flash trimming tool magazine size and reload timing, (3) pipe sizing and rework policy, and (4) the inspection-rework servers are calculated from the information provided in Tables 1 to 4.

The optimum buffer size is obtained from Eq. (1) as [Hossain, 2018]

$$B^* = \lambda\left[\frac{\lambda}{\alpha} + \bar{T}_J + \sigma_T F_{T_J}^{-1}\left(1 - \lambda\frac{c_B + h_B}{c_{td} D_c}\right)\right] = 350.82 \text{ feet.} \quad (1)$$

On the other hand, the optimum magazine size for the flash trimming tool, is found as n^* = 12, and the magazine reload timing is found as τ_{12}^* = 121.29 hours. These solutions are obtained with help of *Tool Heuristic* presented in by Hossain (2018). In the proposed configuration an online inspection facility is planned, which creates an opportunity to decide whether a defective portion should be scrapped or reworked, or a substandard pipe should be made from the defective portion. Making of substandard pipes depend on the sales price of the pipes. With the online inspection facility, only one type of defect (surface scars) is detectable. From the information provided in Table 3, the optimum policies for dealing the defects are found as (x_1, x_2, x_3, y_1, y_2, y_3, z) = (1, 0, 1, 1, 0, 1, 1). This solution indicates that, the defective spots should be reworked at the off-line rework station if the defect free length L falls in the range $L < L_b$ or $L \geq L_s$. However, the defect-free portion should be cut at substandard lengths if $L_b \leq L < L_s$.

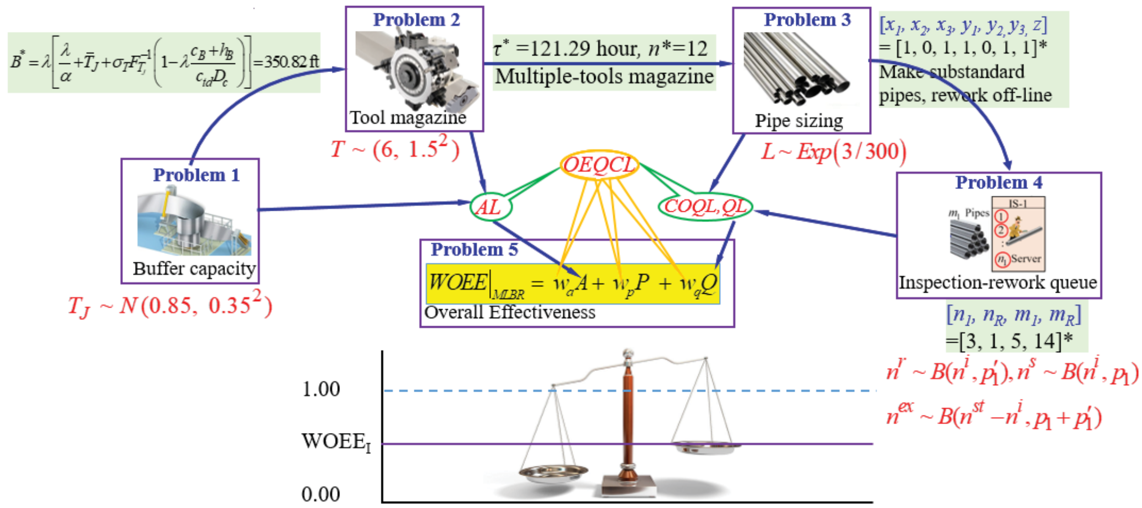


Figure 2. Simulation model for the improved system.

From the information listed in Table 4 the optimum number of servers can be found as n_1 = 3 and n_R =1 for the inspection and rework stations, respectively. These solutions are obtained using the IR-Algorithm presented by Hossain and Sarker (2018). From these algorithm, the optimum space allocations for the inspection and rework stations are found as m_1 = 5 and m_R = 14 pipes, respectively. The improved system configuration based on the local optimum solutions, is presented in Figure 2.

SIMULATION RESULTS

In the pipe manufacturing system several parameters are probabilistic, including coil joining time $T_J \sim N\,(0.85, 0.35^2)$ hours, tool life $T \sim N\,(6, 1.5^2)$ hours and defect free length $L \sim$ Exp (3/300). The arrival of discrete pipes at the inspection-rework facility is also probabilistic, follows approximately Poisson distribution with mean λ' = 5 pipes/hour. Again, the detection of defective pipes in the inspection-rework facility follows binomial distribution with probabilities p_1' and p_1 for the repairable and unrepairable defects, respectively. These uncertainties are captured by simulating the system for several days. The whole pipe manufacturing system is simulated over a month (effectively 22 working days). It is to be noted that the factory is kept running for $T_p = 8$ hours in a day (single shift).

Simulation of the Line with Current Configuration

The system is first simulated with the existing configuration. The coil joining time T_J , and tool life T are randomly generated for each day. All other columns in Appendix 1 are obtained from these two random variables. The coil joining times for the first day are obtained as 0.9776, 0.8679 and 0.7952 hours, respectively. With the buffer capacity $B_{\min}$ = 280 feet, a maximum joining time of $B_{\min}/\lambda - \lambda/\alpha$ = 0.85 hours can be supported. The buffer runout duration in a cycle (T_{br}) is defined as $T_{br} = \max\{0, (T_J - B_{\min}/\lambda + \lambda/\alpha)\}$. As, no joining time should exceed 0.85 hours, buffer runouts will occur for the first two coil joining in the first day. Hence T_{br} = 0.1276, 0.0179 and 0 hours for three coil joining, respectively. Now, the coil feed cycle time will be $T_c' = t_5 + T_J - T_{br} = L_c/\lambda - \bar{T}_J + T_J - T_{br}$ = 800/300 - 0.85 + $T_J - T_{br}$ = 1.8167 $+ T_J - T_{br}$ hours. Hence, the cycle time will be 2.6667, 2.6667 and 2.6119 hours, respectively [Day 1, Column 4 in Appendix 1]. Thus, the available time of production using these three coils is found as 2.6667+2.6667+ 2.6119 = 7.9452 hours, for Day 1.

On the other hand, the first flash trimming tool can run for 5.5423 hours; hence, the second tool needs to be loaded. The life of the second tool is found as 6.0015 hours. These two tool loadings require T^R + T^R = 0.5 + 0.5 = 1 hour, where T^R is the time required for a tool reloading. Thus, the actual production time for Day 1, is found as $T_a = T_p - \sum T_{br} - 2T^R$ = 6.8545 hours, where the potential production time is T_p = 8 hours. From this result, the remaining tool life in Day 1 ($T_{r,1}$) can be found as $T_{r,i}$ = 5.5423 + 6.0015 - 6.8545 = 4.6893 hours [see Day 1, Column 7, in Appendix 1]. Similarly, the remaining coils in the line can be used for $T_{cr,i} = \sum T_c' - T_a$ = 2.6667+ 2.6667+ 2.6119 - 6.8545 = 1.0907 hours in the next day (Day 2). Now, the availability rate (A) for Day 1 can be found as $\mathrm{A} = T_a/T_p$ = 6.86/8 = 0.8568.

Again, if the production line does not produce anything, the industry still has to bear an overhead cost of $600 per hour. Hence, the monetary value for the availability loss becomes AL^c = \$($T_p - T_a$)600 = $687.29 for Day 1. Similarly, the actual production is found as $P_a = T_a\lambda$ = 2056.36 feet. The system is designed with a theoretical capacity of production, λ_{th} = 325 feet/hour. In a real environment this designed capacity is not achieved. So, in general, the theoretical production in Day 1 should be $P_t = T_a\lambda_{th}$ = 2227.72 feet. Hence, the performance rate of the line can be found as $\mathrm{P} = P_a/P_t$ =2056.36/2227.72 = 0.9231.

Now, it is estimated that if the theoretical production rate λ_{th} could be achieved, the manufacturer could make additional profit of $30 from the additional production. So, the monetary value of the availability loss should be $PL^c = \$30(P_t - P_a)$ =$5,140.89. All values in Appendix 1 are evaluated in a similar way.

Appendix 2 is presented to calculate the scrapped pipe lengths, due to the defects that can be detected with the online inspection device. The random defects are simulated and the distance between two consecutive defects (L) are listed in Appendix 2. In the existing situation the defects detected online, are marked on the pipe. Then the pipes are cut into standard lengths ($L_s = 40$ feet). Some defective pipes can be shortened not less than 35 feet by cutting off the defective portion from one end. However, when the defect on the pipe does not fall within 5 feet from any end, the whole pipe is scrapped. This scenario is replicated for all of the defect free lengths in Appendix 2 and the summation of the lengths of all scrapped pipes are noted in that table. The cost of quality loss (QL_1^c) at this stage includes the sales price of the pipe $(P_p = \$180$ /feet) which is scrapped and the scrapping cost of the pipe ($c_s = \$10$ /feet). The cost for the online inspection is considered with a flat rate of $1/feet.

The quality loss and the cost of quality are listed in Appendix 3. The number of pipes produced (n^{st}) in each day are obtained from the simulation of defect free length (L). The number of inspected pipes (n^i) depends on the random arrival process of pipes at the inspection-rework facility. This arrival process follows approximate Poisson distribution with mean λ' = 40 pipes/day, from where n^i is simulated. The number of reworked pipes (n^r), the number of scrapped pipes (n^s) and the number of external failures (n^{ex}) are also simulated, from their corresponding binomial distribution parameters, n^r ~B(n^i , p_1'), n^s ~B(n^i , p_1), n^{ex} ~B(n^{st} - n^i , $p_1 + p_1'$). The external failure cost includes customer dissatisfaction, loss of business goodwill and warranty, which is approximately estimated as $10,000/pipe. From these simulated and estimated information, the monetary equivalent of *QL* (QL_2^c) at the inspection-rework server, is evaluated.

Finally the total quality loss ($QL^c = QL_1^c + QL_2^c$) in monetary unit, is found as the sum of quality losses from the defects detected at the on-line inspection (QL_1^c) and the quality losses in the inspection-rework facility. The cost of quality (COQ) is different from the quality losses. This cost includes all inspection costs (on-line and off-line) as well as external failure costs.

Simulation of the Line with Improved Configuration

The steel pipe manufacturing system is simulated with the optimum solutions and proposed improved conditions. The optimum solution suggests that the buffer capacity should B^* = 350.82 feet, and there should be multiple-tool magazine facility with magazine size n^* = 12 tools and magazine reload timing τ^* = 121.29 hours. For the improved configuration the substandard pipes should be made when $9 \le L < 40$ feet. On the other hand, if the $L < 9$ feet and $L > 40$ feet, the repairable defects, detected with the on-line inspection device, should be marked on the pipes and then repaired at the off-line rework station. The improvement for the system also includes n_1 = 3.

Now, with these locally optimized parameters listed above the simulation model is modified and the system is simulated again. The similar tables are generated for this improved configuration, and corresponding A , P , Q , AL^c , PL^c , QL^c , COQL are evaluated. The values of availability rate A , performance P , the monetary equivalents of availability loss AL^c , and performance loss PL^c are listed in Appendix 4. On the other hand, Appendix 6 summarizes the quality rate Q , the monetary equivalent of quality loss QL^c and the cost of quality loss (COQL). Appendix 5 is presented here to list the random numbers for the defect free length *L* (feet). These results are finally used to evaluate the WOEE, and PEE values.

Comparing Overall Effectiveness of the Existing and Improved Systems

The overall effectiveness of the system, with both existing and improved conditions, is evaluated for 22 days in terms of WOEE and PEE. For this WOEE and PEE calculations, the monetary loss based regression (MLBR) method is used to obtain the weights. Detailed calculations are reported in Table 5. In this table, the overall equipment and quality cost losses (OEQCL) are calculated by summing up all cost losses, as OEQCL= AL^c + PL^c + QL^c +COQL.

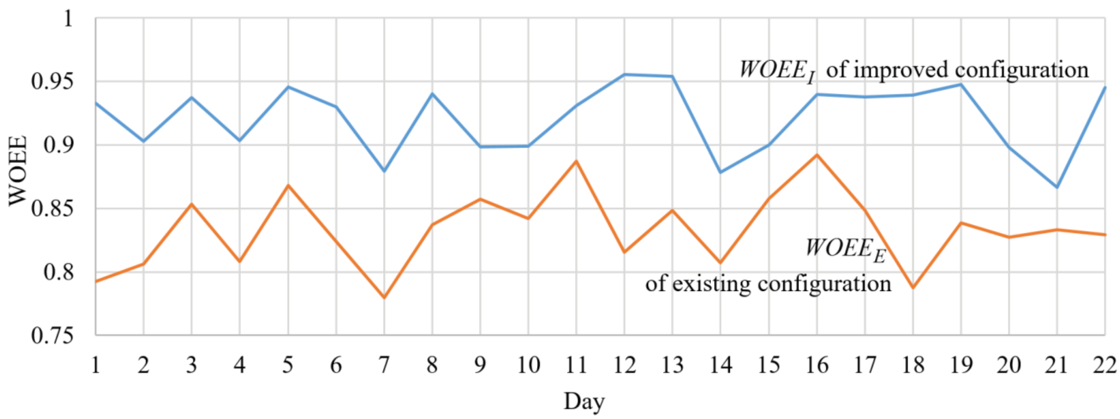


Figure 3. WOEE values of the system with existing and improved configuration.

From the WOEE as well as PEE values it is observed that, the overall effectiveness of the system improves almost 10% when the suggested optimum solutions are incorporated in the system. Figure 3 presents a graphical presentation of WOEE of the system for 22 days. It shows that the WOEE is better for each of the 22 days, when the system is configured with the optimum solution. A paired comparison is done to check the difference with statistical evidence [see Table 5]. Calculating the mean and standard deviation of the differences gives, $E|WOEE_I\text{-}WOEE_E|$= 0.0873 ≈ 8.73% and $SD_{WOEE_I - WOEE_E}$ = 0.0341 ≈ 3.41%.

The *t*-statistic for the difference in WOEE become *t* = 11.9952. On 21 degree of freedom the *p*-value is found as *p*-value = 5.84×10^{-08}. Therefore, there is a strong evidence that, on average, the proposed solution does lead to significant improvements in the overall system. Similarly, in the last column of Table 5, the differences between PEE values for existing and improved system, are listed. The mean and standard deviation of the differences in PEE are found as, $E|PEE_I\text{-} PEE_E|$= 0.0890 and $SD_{PEE_I - PEE_E}$ = 0.0352. The *t*-statistic for the difference in PEE become *t* = 11. 8475 with *p*-value = 6.63×10^{-08}. Hence, the same conclusion can be drawn when the comparison is made in terms of PEE.

Table 5. OEE and PEE of the line in existing and improved configuration.

Day	Existing configuration									Improved configuration									Paired *t*-test	
	A	P	Q	OEQCL	$w_{(1)}$	$w_{(2)}$	$w_{(3)}$	$WOEE_E$	PEE_E	A	P	Q	OEQCL	$w_{(1)}$	$w_{(2)}$	$w_{(3)}$	$WOEE_I$	PEE_I	$\|WOEE_I - WOEE_E\|$	$\|PEE_I - PEE_E\|$
1	0.8568	0.9231	0.7258	107448.70	0.5987	0.2111	0.1902	0.7924	0.7881	0.9375	0.9231	0.9523	26749.85	0.5669	0.2265	0.2066	0.9327	0.9326	0.1403	0.1445
2	0.9045	0.9231	0.7336	102621.10	0.5977	0.2116	0.1907	0.8063	0.8016	0.9966	0.9231	0.8610	56526.23	0.5841	0.2182	0.1977	0.9029	0.9012	0.0965	0.0997
3	0.8892	0.9231	0.8166	82613.86	0.5928	0.2140	0.1932	0.8534	0.8522	0.9812	0.9231	0.9275	30834.26	0.5702	0.2249	0.2049	0.9371	0.9368	0.0836	0.0845
4	0.8517	0.9231	0.7538	101534.10	0.5975	0.2117	0.1908	0.8083	0.8054	0.9396	0.9231	0.8830	41212.8	0.5769	0.2217	0.2015	0.9036	0.9033	0.0953	0.0979
5	0.8808	0.9231	0.8447	97365.69	0.5965	0.2122	0.1913	0.8683	0.8677	0.9898	0.9231	0.9621	42408.4	0.5775	0.2214	0.2011	0.9457	0.9453	0.0774	0.0776
6	0.7954	0.9231	0.8023	86076.13	0.5937	0.2135	0.1928	0.8240	0.8224	0.9271	0.9231	0.9504	56260.27	0.5840	0.2182	0.1978	0.9298	0.9298	0.1059	0.1073
7	0.8848	0.9231	0.6960	113691.80	0.6000	0.2105	0.1895	0.7796	0.7730	0.9890	0.9231	0.8245	65830.66	0.5876	0.2165	0.1959	0.8794	0.8768	0.0998	0.1038
8	0.9245	0.9231	0.7779	92318.83	0.5953	0.2127	0.1919	0.8370	0.8340	1.0000	0.9231	0.9252	73127.36	0.5900	0.2153	0.1947	0.9401	0.9396	0.1031	0.1056
9	0.8696	0.9231	0.8295	83573.86	0.5930	0.2138	0.1931	0.8572	0.8564	0.9834	0.9231	0.8599	76352.43	0.5910	0.2148	0.1942	0.8987	0.8973	0.0415	0.0409
10	0.8914	0.9231	0.7960	73204.43	0.5900	0.2153	0.1947	0.8419	0.8401	0.9763	0.9231	0.8621	65198.7	0.5874	0.2166	0.1960	0.8988	0.8976	0.0568	0.0575
11	0.9717	0.9231	0.8455	83753.87	0.5931	0.2138	0.1931	0.8875	0.8859	0.9229	0.9231	0.9581	33130.63	0.5718	0.2241	0.2041	0.9308	0.9307	0.0434	0.0448
12	0.8697	0.9231	0.7597	93971.55	0.5957	0.2126	0.1917	0.8155	0.8126	1.0000	0.9231	0.9930	18160.84	0.5579	0.2308	0.2113	0.9556	0.9549	0.1401	0.1423
13	0.9200	0.9231	0.7992	102267.80	0.5976	0.2116	0.1908	0.8485	0.8464	0.9944	0.9231	0.9945	26041.8	0.5663	0.2268	0.2070	0.9540	0.9534	0.1056	0.1070
14	0.9024	0.9231	0.7354	100159.90	0.5972	0.2118	0.1910	0.8071	0.8025	0.9736	0.9231	0.8272	54119.66	0.5831	0.2187	0.1982	0.8782	0.8760	0.0711	0.0735
15	0.8816	0.9231	0.8264	85132.23	0.5935	0.2136	0.1929	0.8577	0.8568	1.0000	0.9231	0.8564	88070.74	0.5942	0.2133	0.1925	0.8998	0.8980	0.0422	0.0413
16	0.9249	0.9231	0.8709	106928.10	0.5986	0.2111	0.1902	0.8922	0.8919	1.0000	0.9231	0.9241	92207.23	0.5953	0.2128	0.1920	0.9396	0.9391	0.0474	0.0473
17	0.8750	0.9231	0.8130	90205.36	0.5948	0.2130	0.1922	0.8484	0.8472	0.9628	0.9231	0.9533	56909.91	0.5843	0.2181	0.1976	0.9377	0.9376	0.0893	0.0904
18	0.8708	0.9231	0.7134	106546.10	0.5986	0.2112	0.1903	0.7876	0.7824	0.9721	0.9231	0.9493	38777.21	0.5755	0.2224	0.2022	0.9393	0.9391	0.1517	0.1567
19	0.8540	0.9231	0.8028	75753.53	0.5908	0.2149	0.1943	0.8386	0.8373	1.0000	0.9231	0.9630	53931.84	0.5830	0.2187	0.1983	0.9478	0.9473	0.1092	0.1100
20	0.9223	0.9231	0.7651	134919.40	0.6039	0.2086	0.1875	0.8275	0.8240	0.9375	0.9231	0.8756	74331.93	0.5904	0.2151	0.1945	0.8982	0.8977	0.0706	0.0737
21	0.9266	0.9231	0.7718	102903.30	0.5978	0.2116	0.1907	0.8334	0.8301	0.9627	0.9231	0.8117	60038.16	0.5855	0.2175	0.1970	0.8665	0.8640	0.0331	0.0339
22	0.9014	0.9231	0.7728	98318.89	0.5967	0.2121	0.1912	0.8293	0.8265	1.0000	0.9231	0.9510	57928.84	0.5847	0.2179	0.1974	0.9453	0.9449	0.1161	0.1184
																		Mean	**0.0873**	**0.0890**
																		SD	**0.0341**	**0.0352**
																		t	**11.9952**	**11.8475**
																		***p*-value**	**5.84×10^{-08}**	**6.63×10^{-08}**

CONCLUSION

This paper numerically shows the overall improvement of the system when proposed local-optimum solutions are implemented. An impression of a pipe manufacturing system is simulated for 22 working days, for the existing and the proposed optimum conditions. It is statistically proven that the overall effectiveness of the system improves with the optimum solution. On an average, the WOEE and PEE improve more than 10%. The proposed local optimum solutions can lead to the improvement in availability rate (A) and quality rate (Q) only. The first and second optimization problems were concerned about buffer capacity and cutting tool magazine, jointly improve the availability rate by reducing the productive time losses. While, the third and fourth problems were dealing with pipe sizing and inspection-rework facility, collectively improve the quality rate by optimally suggesting some rework processes. Another element of OEE is performance (P), which is an internal technical issue of the machine, and hence, it can be improved by technical modification of the machine. Besides these important observations, this simulation result presents a clear insight of the system and explains the implementation phase in a virtual environment.

REFERENCES

Alvarezvargas, R., Dallery, Y. and David, R. (1994) “A study of the continuous-flow model of production lines with unreliable machines and finite buffers,” *Journal of Manufacturing Systems* **13**(3): 221-234.

Bierbooms, R., Adan, I. J. B. F. and Vuuren, M. V. (2013) “Approximate Performance Analysis of Production Lines with Continuous Material Flows and Finite Buffers,” *Stochastic Models* **29**(1): 1-30.

Della Croce, F., Salassa, F. and Scatamacchia, R. (2017) “A new exact approach for the 0-1 Collapsing Knapsack Problem,” *European Journal of Operational Research* **260**(1): 56-69.

Feng, Y., Wang, G. G., Deb, S., Lu, M. and Zhao, X. J. (2017) “Solving 0-1 knapsack problem by a novel binary monarch butterfly optimization,” *Neural Computing & Applications* **28**(7): 1619-1634.

Furini, F., Ljubic, I. and Sinnl, M. (2017) “An effective dynamic programming algorithm for the minimum-cost maximal knapsack packing problem,” *European Journal of Operational Research* **262**(2): 438-448.

Hippalgaonkar, R. R. and Shin, Y. C. (2011) “Robust optimization of machining conditions with tool life and surface roughness uncertainties,” *International Journal of Production Research* **49**(13): 3963-3978.

Hossain, M. S. J. and Sarker, B. R. (2016a) “Optimal locations of on-line and off-line rework stations in a serial production system,” *International Journal of Production Research* **54**(12): 3603-3621.

Hossain, M.S.J. and Sarker, B.R. (2016b) “Overall equipment effectiveness measures of engineering production systems,” *Proceedings of the 2016 Annual Meeting of the Decision Sciences Institute*, November 19-22, Austin, Texas, USA.

Hossain, M.S.J. and Sarker, B.R. (2017b) “Optimization of rework policy and length of defective pipes in a continuous metal pipe manufacturing system,” *Proceedings of IISE*

Annual Conference and Expo 2017, May 20-23, Pittsburgh, Pennsylvania, USA, pp. 127-132.

Hossain, M.S.J. and Sarker, B.R. (2018a) “Buffer capacity of coil accumulator in the seam welded pipe manufacturing mill,” 2018 MIE Graduate Student Conference, April 14, Louisiana State University, Baton Rouge, LA 70803.

Hossain, M.S.J. and Sarker, B.R. (2018b) “Optimal reload timing and size of tool magazine in steel pipe manufacturing,” *POMS 29th Annual Conference*, May 4-7, Houston, Texas, USA.

Hossain, M.S.J. and Sarker, B.R. (2018c) “Optimal usable time of tool magazines and required number of tools in steel pipe manufacturing,” for *AAER International Conference on Communication, Engineering, Data Mining, Information Technology & Applied Sciences* (*CEDIA*)*.* May 17-18, Manila, Philippines.

Hossain, M.S.J. and Sarker, B.R. (2018d) “Surrogate model based optimization of inspection/rework server allocation for generalized distribution of arrival time,” *POMS 29th Annual Conference*, May 4-7, Houston, Texas, USA.

Hossain, M.S.J. and Sarker, B.R. (2019) “Monetary-Loss Based Regression Method for Overall Equipment Effectiveness Measure,” *Northwestern State University's 32nd Annual Research Day*, April 25, Natchitoches, LA 71497.

Hsieh, Y. C. (1997) “Alternative approaches for one-dimensional slitting problems,” *Journal of the Operational Research Society* **48**(3): 314-318.

Karandikar, J. M., Schmitz, T. L. and Abbas, A. E. (2012) “Spindle speed selection for tool life testing using Bayesian inference,” *Journal of Manufacturing Systems* **31**(4): 403-411.

Laghaie, K. S., Mansouri, M., Larijani, A. M. and Naini, G. J. (2012) “Combining a maintenance center M/M/c/m queue into the economic production quantity model with stochastic machine breakdowns and repair,” *Computers & Industrial Engineering* **63**: 864–874.

Lamond, B. F. and Sodhi, M. S. (2006) “Minimizing the expected processing time on a flexible machine with random tool lives,” *IIE Transactions* **38**(1): 1-11.

Lee, J., Kim, B. I. and Johnson, A. L. (2013) “A two-dimensional bin packing problem with size changeable items for the production of wind turbine flanges in the open die forging industry,” *IIE Transactions* **45**(12): 1332-1344.

Ozdogru, U. and Altiok, T. (2015) “Continuous material flow systems: analysis of marine ports handling bulk materials,” *Annals of Operations Research* **231**:79-104.

Rajaram, K. and Tian, Z. L. (2009) “Buffer location and sizing to optimize cost and quality in semi-continuous manufacturing processes: Methodology and application,” *IIE Transactions* **41**(12): 1035-1048.

Raouf, A. (1994), “Improving capital productivity through maintenance,” *International Journal of Operations and Production Management,* **14**(7), 44–52.

Ravid, R., Boxma, O. J. and Perry, D. (2013) “Repair systems with exchangeable items and the longest queue mechanism,” *Queueing System* **73**:295–316.

Sarker, B. R. (1988) “An optimum solution for one-dimensional slitting problems: a dynamic programming approach,” *Journal of the Operational Research Society* **39**(8): 749–755.

Selim, S. Z. and Al-Zu'bi, W. K. (2011) “Optimal means for continuous processes in series,” *European Journal of Operational Research* **210**(3): 618-623.

Sodhi, M. S. and Sarker, B. R. (2003), “Configuring Flexible Flowlines,” *International Journal of Production Research,* **41**(8): 1689-1706.

Wang, P. and Gao, R. X. (2015) “Adaptive resampling-based particle filtering for tool life prediction,” *Journal of Manufacturing Systems* **37**(SI) 528-534.

Wudhikarn, R. (2010) “Overall Weighting Equipment Effectiveness,” *Proceedings of the 2010 IEEE IEEM*, 23-27.

Yan, L. X. (2003) “Solving combinatorial optimization problems with line-up competition algorithm,” *Computers & Chemical Engineering* **27**(2): 251-258.

Yao, C. W. and Chien, Y. X. (2014) “A diagnosis method of wear and tool life for an end mill by ultrasonic detection,” *Journal of Manufacturing Systems* **33**(1): 129-138.

Yuniawan, D., Ito, T. and Bin, M. E. (2013) “Calculation of overall equipment effectiveness weight by Taguchi method with simulation,” *Concurrent Engineering,* **21**(4), 296-306.

APPENDIX 1. Simulation results for evaluating availability A, and performance P of the system with existing configuration.

Day i	Coil joining time T_J	Buffer runout time T_{br}	Coil feed Cycle time T_c'	Remaining cycle time for next day use $T_{cr,i} = T_{cr,i-1} + \sum T_c' - T_a$	Tool life T	Remaining Tool life for next day use $T_{r,i} = \sum T + T_{r,i-1} - T_a$	Number of tool loads n_t	Actual production time $T_a = T_p - \sum T_{br} - n_t T^R$	A	AL^c $/day	Actual Production P_a feet	Theoretical production P_t feet	PL^c $/day @$30 /feet	P $= P_a / P_t$
1	0.9776 0.8679 0.7952	0.1276 0.0179 0.0000	2.6667 2.6667 2.6119	1.0907	5.5423 6.0015	4.6893	2	6.8545	0.8568	687.29	2056.36	2227.72	5140.89	0.9231
2	1.1136 0.5325 0.6818	0.2636 0.0000 0.0000	2.6667 2.3492 2.4985	1.3686	4.5035	1.9564	1	7.2364	0.9045	458.18	2170.91	2351.82	5427.28	0.9231
3	1.2364 0.8459 0.5876	0.3864 0.0000 0.0000	2.6667 2.6626 2.4043	1.9885	8.6871	3.5299	1	7.1136	0.8892	531.83	2134.08	2311.92	5335.21	0.9231
4	0.8080 1.5363	0.0000 0.6863	2.6247 2.6667	0.4661	7.1204	3.8365	1	6.8137	0.8517	711.76	2044.12	2214.46	5110.30	0.9231
5	1.1192 0.7961 1.0345	0.2692 0.0000 0.1845	2.6667 2.6127 2.6667	1.3658	4.6557	1.4459	1	7.0463	0.8808	572.20	2113.90	2290.06	5284.75	0.9231
6	1.1679 0.5115 1.6690	0.3179 0.0000 0.8190	2.6667 2.3282 2.6667	2.6643	8.5354	3.6183	1	6.3630	0.7954	982.18	1908.91	2067.99	4772.28	0.9231
7	0.9474 1.1740	0.0974 0.3240	2.6667 2.6667	0.9191	5.6336	2.1733	1	7.0786	0.8848	552.86	2123.57	2300.53	5308.92	0.9231
8	0.9538 0.8486 0.7327	0.1038 0.0000 0.0000	2.6667 2.6653 2.5494	1.4042	5.9845	0.7615	1	7.3962	0.9245	362.28	2218.86	2403.76	5547.15	0.9231
9	0.8464 0.8836 0.9430	0.0000 0.0336 0.0930	2.6630 2.6667 2.6667	2.4439	6.1951 7.1211	7.1211	2	6.9566	0.8696	626.01	2086.99	2260.91	5217.48	0.9231
10	1.2186 0.2800	0.3686 0.0000	2.6667 2.0966	0.0758	7.7837	7.7734	1	7.1314	0.8914	521.18	2139.41	2317.69	5348.52	0.9231
11	0.8061 0.3687 0.4977 0.4994	0.0000 0.0000 0.0000 0.0000	2.6228 2.1854 2.3143 2.3161	1.7411	5.0126	5.0126	1	7.7734	0.9717	135.96	2332.02	2526.36	5830.05	0.9231
12	1.3925 0.4553 0.7782	0.5425 0.0000 0.0000	2.6667 2.2720 2.5949	2.3170	6.5970	4.6521	1	6.9575	0.8697	625.48	2087.26	2261.20	5218.15	0.9231

(To be continued…)

APPENDIX 1. continued.

Day i	Coil joining time T_J	Buffer runout time T_{br}	Coil feed Cycle time T_c'	Remaining cycle time for next day use $T_{cr,i} = T_{cr,i-1} + \sum T_c' - T_a$	Tool life T	Remaining Tool life for next day use $T_{r,i} = \sum T + T_{r,i-1} - T_a$	Number of tool loads n_t	Actual production time $T_a = T_p - \sum T_{br} - n_t T^R$	A	AL^c \$/day	Actual Production P_a feet	Theoretical production P_t feet	PL^c \$/day @\$30 /feet	P $= P_a/P_t$
13	0.1219 0.5704 0.9902	0.0000 0.0000 0.1402	1.9385 2.3870 2.6667	1.9494	6.3811	3.6735	1	7.3598	0.9200	384.13	2207.94	2391.93	5519.84	0.9231
14	0.8184 1.1305	0.0000 0.2805	2.6350 2.6667	0.0316	5.4314	1.8854	1	7.2195	0.9024	468.30	2165.85	2346.34	5414.63	0.9231
15	0.8439 1.2975 0.7033	0.0000 0.4475 0.0000	2.6605 2.6667 2.5199	0.8263	7.4436	2.2765	1	7.0525	0.8816	568.50	2115.75	2292.06	5289.38	0.9231
16	0.4558 0.4333 0.9507	0.0000 0.0000 0.1007	2.2725 2.2500 2.6667	0.6162	6.6735	1.5507	1	7.3993	0.9249	360.43	2219.78	2404.77	5549.46	0.9231
17	0.8127 0.8307 0.6037	0.0000 0.0000 0.0000	2.6294 2.6474 2.4204	1.3133	4.2921 7.3236	6.1664	2	7.0000	0.8750	600.00	2100.00	2275.00	5250.00	0.9231
18	0.5134 1.3839 0.7356	0.0000 0.5339 0.0000	2.3301 2.6667 2.5523	1.8963	7.3298	6.5301	1	6.9661	0.8708	620.33	2089.84	2263.99	5224.59	0.9231
19	0.7219 1.3092 1.0589	0.0000 0.4592 0.2089	2.5386 2.6667 2.6667	2.9363	5.6892	5.3874	1	6.8319	0.8540	700.89	2049.56	2220.35	5123.89	0.9231
20	0.3845 0.9714 0.5987	0.0000 0.1214 0.0000	2.2012 2.6667 2.4154	2.8409	7.0683	5.0770	1	7.3786	0.9223	372.83	2213.59	2398.05	5533.96	0.9231
21	0.8517 0.4354 0.9355	0.0017 0.0000 0.0855	2.6667 2.2520 2.6667	3.0134	4.8928	2.5570	1	7.4129	0.9266	352.29	2223.86	2409.18	5559.64	0.9231
22	0.8177 1.1387	0.0000 0.2887	2.6344 2.6667	1.1031	5.8004	1.1460	1	7.2113	0.9014	473.19	2163.40	2343.69	5408.51	0.9231

APPENDIX 2. Quality gain in pipe sizing and rework policy problem (no online rework).

Day *i*	Defect free length, *L* (feet) $L \sim Exp(1/100)$	Scrapped pipe length l_1^s feet	QL_1^c $/day	Inspection cost @$1
1	156.88, 109.11, 269.28, 65.34, 47.51, 0.09, 146.67, 13.73, 244.66, 16.33, 76.84, 63.87, 47.08, 21.22, 86.52, 43.84, 33.82, 28.15, 30.99, 200.22, 95.18, 0.22	323.81	61524.18	2056.36
2	287.46, 60.12, 178.90, 27.77, 115.06, 15.48, 71.11, 267.82, 59.39, 23.16, 235.71, 102.88, 283.85, 222.73, 81.10, 44.33, 132.11	258.24	49064.74	2170.91
3	287.92, 67.24, 75.34, 76.13, 211.80, 103.87, 195.56, 315.21, 22.37, 113.98, 201.89, 225.92, 67.72, 15.37, 70.03, 283.21	231.33	43952.73	2134.08
4	30.08, 186.10, 70.41, 171.20, 54.81, 20.80, 20.54, 72.45, 73.03, 24.33, 53.60, 179.78, 132.12, 243.96, 28.56, 146.93, 175.73, 60.22, 83.33, 12.61, 12.82, 77.35, 89.80	423.20	80407.89	2044.12
5	66.87, 93.31, 37.11, 7.67, 83.30, 12.18, 92.28, 8.09, 51.66, 47.11, 207.92, 6.03, 20.45, 79.00, 173.50, 151.32, 61.84, 152.53, 122.00, 90.14, 199.76, 75.45, 39.29	248.18	47154.84	2113.90
6	445.59, 51.25, 25.92, 64.70, 72.88, 5.74, 29.97, 28.11, 39.03, 96.97, 86.94, 208.28, 16.78, 12.08, 109.53, 86.17, 4.44, 58.97, 197.06, 197.53, 38.33, 131.59, 38.21, 41.53	297.44	56512.76	1908.91
7	152.41, 96.79, 67.18, 11.21, 91.03, 172.29, 101.87, 163.30, 20.99, 116.93, 7.32, 10.01, 10.63, 178.06, 46.29, 154.94, 193.43, 119.58, 5.46, 7.99, 88.21, 180.61, 34.44, 71.64, 19.41	405.51	77046.46	2123.57
8	105.93, 48.95, 192.01, 157.72, 149.28, 85.58, 10.85, 346.57, 105.61, 11.79, 62.48, 1.39, 20.10, 24.14, 350.07, 128.74, 42.27, 24.53, 219.54, 2.97, 116.22	332.79	63230.54	2218.86
9	158.06, 4.60, 219.83, 51.46, 62.41, 44.41, 54.78, 300.83, 48.91, 34.17, 155.09, 41.17, 30.42, 117.28, 71.89, 111.64, 235.66, 88.01, 191.37	275.91	52423.37	2086.99
10	301.17, 81.99, 119.41, 6.36, 324.91, 99.42, 378.78, 64.40, 265.36, 100.72, 30.07, 543.21	196.40	37315.32	2139.41
11	174.83, 18.45, 5.71, 185.01, 96.15, 79.31, 164.62, 242.09, 222.67, 41.96, 40.11, 12.35, 149.51, 110.27, 153.68, 197.22, 115.69, 3.31, 14.03, 125.78, 3.51, 5.30, 30.95	280.29	53255.84	2332.02
12	247.69, 134.71, 280.23, 154.29, 1.13, 167.87, 92.67, 168.30, 10.81, 71.51, 24.26, 18.40, 28.43, 55.58, 59.87, 90.85, 216.95, 145.29, 24.94, 187.69	341.48	64880.66	2087.26
13	57.78, 233.42, 150.49, 307.64, 73.77, 87.61, 247.16, 221.36, 28.02, 2.33, 291.88, 12.27, 121.58, 26.87, 92.70, 4.30, 84.15	283.35	53835.86	2207.94
14	280.08, 49.01, 213.50, 8.00, 72.22, 36.29, 60.98, 98.98, 28.35, 154.76, 57.88, 29.23, 97.40, 185.04, 56.85, 84.26, 75.21, 137.79, 55.16, 95.98, 237.58, 167.53	333.01	63271.11	2165.85
15	199.86, 9.06, 126.64, 72.77, 21.57, 31.46, 223.66, 176.21, 714.63, 8.52, 64.97, 36.27, 171.50, 26.43, 16.90, 142.95, 0.09	287.36	54598.60	2115.75
16	723.44, 29.36, 34.83, 116.69, 180.14, 86.67, 45.68, 62.61, 10.15, 3.84, 35.77, 9.32, 70.97, 319.03, 68.02, 231.68, 28.81, 174.76, 6.21	286.52	54438.38	2219.78
17	212.62, 64.28, 15.87, 237.23, 1.95, 89.59, 107.37, 86.74, 13.23, 173.35, 109.33, 12.90, 7.37, 115.84, 64.12, 101.46, 14.62, 36.03, 10.36, 195.54, 7.50, 244.48, 48.31, 23.80, 2.93, 100.43	312.61	59395.36	2100.00
18	357.42, 77.32, 24.96, 181.80, 17.88, 65.84, 17.93, 12.78, 78.89, 20.66, 292.99, 44.16, 34.09, 117.87, 57.54, 87.83, 327.52, 103.33, 139.78, 49.91	279.01	53011.32	2089.84
19	329.72, 354.68, 101.09, 15.75, 19.94, 66.84, 106.40, 97.27, 80.29, 89.34, 101.58, 82.42, 533.63, 65.15	244.10	46379.19	2049.56
20	264.88, 296.10, 20.74, 53.96, 18.89, 242.43, 57.96, 32.71, 60.35, 160.82, 33.24, 76.65, 114.22, 34.37, 0.26, 68.36, 10.95, 54.27, 52.23, 128.66, 2.78, 198.01, 68.03, 17.41, 48.32, 87.58, 24.58, 40.58, 5.35	439.99	83599.00	2213.59
21	139.50, 12.04, 115.48, 63.27, 11.42, 88.40, 261.62, 24.98, 26.69, 34.63, 273.89, 152.57, 8.24, 61.52, 132.79, 32.87, 353.65, 63.83, 298.22, 27.39	427.51	81227.52	2223.86
22	347.67, 140.44, 98.38, 13.15, 79.95, 244.90, 30.14, 66.56, 17.25, 143.07, 19.46, 385.80, 32.34, 20.41, 25.92, 140.70, 121.06, 69.76, 10.52, 44.25, 0.43, 106.78, 12.43	411.44	78173.79	2163.40

APPENDIX 3. Quality loss, and cost in the final inspection-rework facility (existing configuration).

Day i	Standard pipes count n^{st}	Inspected pipes count n^{i}	Number* of External Failures n^{ex}	Number* of Reworked pipes, n^{r}	Number* of Scrapped pipes, n^{s}	External Failure cost @$10,000 /pipe	$QL_2^c = n^s(P_p + c_s)L_s + c_R n^r$	Off-line Inspection cost $c_1 n^i$ + 8(C_1)	Total Scrapped length $l^s = n^s L_s + l_1^s$	Length of pipes $P_g = P_a - l^s - n^s L_s$	Actual Production P_a feet	Q = P_g / P_a	QL^c $QL_1^c + QL_2^c$ $/day	COQL $/day
1	43	38	1	4	3	10000	23280	4760	443.81	1492.54	2056.36	0.7258	84,804.18	16816.36
2	48	39	1	2	4	10000	30640	4860	418.24	1592.68	2170.91	0.7336	79,704.74	17030.91
3	48	39	1	5	2	10000	15800	4860	311.33	1742.75	2134.08	0.8166	59,752.73	16994.08
4	41	41	0	5	1	0	8200	5060	463.20	1540.92	2044.12	0.7538	88,607.89	7104.12
5	46	32	3	4	1	30000	8080	4160	288.18	1785.72	2113.90	0.8447	55,234.84	36273.90
6	41	31	1	2	1	10000	7840	4060	337.44	1531.48	1908.91	0.8023	64,352.76	15968.91
7	43	43	0	5	3	0	23400	5260	525.51	1478.06	2123.57	0.6960	100,446.46	7383.57
8	47	42	0	5	2	0	15800	5160	412.79	1726.07	2218.86	0.7779	79,030.54	7378.86
9	46	37	1	8	1	10000	8560	4660	315.91	1731.08	2086.99	0.8295	60,983.37	16746.99
10	49	40	0	1	3	0	22920	4960	316.40	1703.01	2139.41	0.7960	60,235.32	7099.41
11	51	34	1	2	1	10000	7840	4360	320.29	1971.73	2332.02	0.8455	61,095.84	16692.02
12	44	44	0	5	2	0	15800	5360	421.48	1585.78	2087.26	0.7597	80,680.66	7447.26
13	48	38	2	3	2	20000	15560	4760	363.35	1764.59	2207.94	0.7992	69,395.86	26967.94
14	46	46	0	4	3	0	23280	5560	453.01	1592.85	2165.85	0.7354	86,551.11	7725.85
15	46	34	1	5	1	10000	8200	4360	327.36	1748.39	2115.75	0.8264	62,798.60	16475.75
16	48	28	4	5	0	40000	600	3760	286.52	1933.27	2219.78	0.8709	55,038.38	45979.78
17	45	37	1	5	1	10000	8200	4660	352.61	1707.39	2100.00	0.8130	67,595.36	16760.00
18	46	40	1	2	4	10000	30640	4960	439.01	1490.83	2089.84	0.7134	83,651.32	17049.84
19	45	45	0	7	2	0	16040	5460	324.10	1645.46	2049.56	0.8028	62,419.19	7509.56
20	44	38	3	7	1	30000	8440	4760	479.99	1693.59	2213.59	0.7651	92,039.00	36973.59
21	45	45	0	4	1	0	8080	5460	467.51	1716.34	2223.86	0.7718	89,307.52	7683.86
22	44	33	0	2	1	0	7840	4260	451.44	1671.96	2163.40	0.7728	86,013.79	6423.40

*Simulated assuming binomial distributions, $n^r \sim$ Bin(n^i , p_1'), $n^s \sim$ Bin(n^i , p_1), $n^{ex} \sim$ Bin(n^{st} - n^i , $p_1 + p_1'$).

APPENDIX 4. Simulation results for evaluating availability A, and performance P of the system with improved configuration.

Day i	Coil joining time T_J	Buffer runout time T_{br}	Coil feed Cycle time T_c'	Remaining cycle time for next day use $T_{cr,i} = T_{cr,i-1} + \sum T_c' - T_a$	Tool life T	Remaining Tool life for next day use $T_{r,i} = \sum T + T_{r,i-1} - T_a$	Remaining tools in the magazine at the end of the day	Number of magazine reload n_t	Actual production time $T_a = T_p - \sum T_{br} - n_t T^R$	A	AL^c $/day	Actual Production P_a feet	Theoretical production P_t feet	PL^c $/day @$30 /feet	P $= P_a / P_t$
1	0.9776 0.8679 0.7952	0.0000 0.0000 0.0000	2.7942 2.6846 2.6119	0.5907	5.5423 6.0015	4.0438	10	1	7.5000	0.9375	300.00	2250.00	2437.50	0.9231	5625.00
2	1.1136 0.5325 0.6818	0.0276 0.0000 0.0000	2.9027 2.3492 2.4985	0.3686	4.5035	0.5749	9	0	7.9724	0.9966	16.54	2391.73	2591.04	0.9231	5979.33
3	1.2364 0.8459 0.5876	0.1503 0.0000 0.0000	2.9027 2.6626 2.4043	0.4885	8.6871	1.4123	8	0	7.8497	0.9812	90.19	2354.90	2551.15	0.9231	5887.26
4	0.8080 1.5363 1.1192	0.0000 0.4502 0.0331	2.6247 2.9027 2.9027	1.4019	7.1204	1.0159	7	0	7.5167	0.9396	289.99	2255.00	2442.92	0.9231	5637.51
5	0.7961 1.0345 1.1679	0.0000 0.0000 0.0819	2.6127 2.8512 2.9027	1.8504	4.6557 8.5354	6.2889	5	0	7.9181	0.9898	49.11	2375.44	2573.40	0.9231	5938.61
6	0.5115 1.6690 0.9474	0.0000 0.5830 0.0000	2.3282 2.9027 2.7641	2.4284	5.6336	4.5054	4	0	7.4170	0.9271	349.78	2225.11	2410.53	0.9231	5562.77
7	1.1740 0.9538	0.0879 0.0000	2.9027 2.7705	0.1896	5.9845	2.5779	3	0	7.9121	0.9890	52.76	2373.62	2571.42	0.9231	5934.05
8	0.8486 0.7327 0.8464	0.0000 0.0000 0.0000	2.6653 2.5494 2.6630	0.0672	6.1951	0.7730	2	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00
9	0.8836 0.9430 1.2186	0.0000 0.0000 0.1326	2.7003 2.7596 2.9027	0.5624	7.1211	0.0267	1	0	7.8674	0.9834	79.54	2360.23	2556.91	0.9231	5900.57
10	0.2800 0.8061 0.3687 0.4977	0.0000 0.0000 0.0000 0.0000	2.0966 2.6228 2.1854 2.3143	1.9712	7.7837	0.0000	0	0	7.8104	0.9763	113.79	2343.11	2538.36	0.9231	5857.76
11	0.4994 1.3925 0.4553	0.0000 0.3064 0.0000	2.3161 2.9027 2.2720	2.0787	5.0126 6.5970	4.2263	10	1	7.3832	0.9229	370.05	2214.97	2399.56	0.9231	5537.44
12	0.7782 0.1219	0.0000 0.0000	2.5949 1.9385	0.9992	6.3811	2.6075	9	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00

(To be continued…)

Hossain & Sarker Evaluation of a Continuous Production System

APPENDIX 4. continued.

Day i	Coil joining time T_J	Buffer runout time T_{br}	Coil feed Cycle time T_c'	Remaining cycle time for next day use $T_{cr,i} = T_{cr,i-1} + \sum T_c' - T_a$	Tool life T	Remaining Tool life for next day use $T_{r,i} = \sum T + T_{r,i-1} - T_a$	Remaining tools in the magazine at the end of the day	Number of magazine reload n_t	Actual production time $T_a = T_p - \sum T_{br} - n_t T^R$	A	AL^c \$/day	Actual Production P_a feet	Theoretical production P_t feet	PL^c \$/day @\$30 /feet	P $= P_a/P_t$
	0.5704	0.0000	2.3870												
13	0.9902 0.8184 1.1305	0.0000 0.0000 0.0444	2.8069 2.6350 2.9027	1.3882	5.4314	0.0833	8	0	7.9556	0.9944	26.66	2386.67	2585.56	0.9231	5966.68
14	0.8439 1.2975 0.7033	0.0000 0.2114 0.0000	2.6605 2.9027 2.5199	1.6829	7.4436 6.6735	6.4119	6	0	7.7886	0.9736	126.86	2336.57	2531.29	0.9231	5841.43
15	0.4558 0.4333 0.9507	0.0000 0.0000 0.0000	2.2725 2.2500 2.7674	0.9728	4.2921	2.7040	5	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00
16	0.8127 0.8307 0.6037	0.0000 0.0000 0.0000	2.6294 2.6474 2.4204	0.6699	7.3236	2.0276	4	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00
17	0.5134 1.3839 0.7356	0.0000 0.2978 0.0000	2.3301 2.9027 2.5523	0.7529	7.3298	1.6552	3	0	7.7022	0.9628	178.69	2310.66	2503.21	0.9231	5776.64
18	0.7219 1.3092 1.0589	0.0000 0.2231 0.0000	2.5386 2.9027 2.8756	1.2929	5.6892 7.0683	6.6358	1	0	7.7769	0.9721	133.88	2333.06	2527.48	0.9231	5832.64
19	0.3845 0.9714 0.5987	0.0000 0.0000 0.0000	2.2012 2.7880 2.4154	0.6975	4.8928	3.5286	0	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00
20	0.8517 0.4354 0.9355	0.0000 0.0000 0.0000	2.6683 2.2520 2.7521	0.8700	5.8004	1.8289	11	1	7.5000	0.9375	300.00	2250.00	2437.50	0.9231	5625.00
21	0.8177 1.1387 1.3318	0.0000 0.0526 0.2457	2.6344 2.9027 2.9027	1.6082	7.1532	1.2804	10	0	7.7017	0.9627	178.99	2310.50	2503.05	0.9231	5776.26
22	1.0354 0.6453 0.4184	0.0000 0.0000 0.0000	2.8521 2.4620 2.2351	1.1573	2.7970 7.3247	3.4020	8	0	8.0000	1.0000	0.00	2400.00	2600.00	0.9231	6000.00

APPENDIX 5. Quality gain in pipe sizing and rework policy problem (improved configuration).

Day i	Defect free length, L (feet) $L \sim Exp(1/100)$	Substandard pipe count	Scrapped pipe length	Rework cost	QL_1^c $/day	Inspection cost @$1
1	156.88, 109.11, 269.28, 65.34, 47.51, 0.09, 146.67, 13.73, 244.66, 16.33, 76.84, 63.87, 47.08, 21.22, 86.52, 43.84, 33.82, 28.15, 30.99, 200.22, 95.18, 0.22	13	27.24	320	5494.85	2250.00
2	287.46, 60.12, 178.90, 27.77, 115.06, 15.48, 71.11, 267.82, 59.39, 23.16, 235.71, 102.88, 283.85, 222.73, 81.10, 44.33, 132.11	6	12.41	280	2638.63	2391.73
3	287.92, 67.24, 75.34, 76.13, 211.80, 103.87, 195.56, 315.21, 22.37, 113.98, 201.89, 225.92, 67.72, 15.37, 70.03, 283.21	3	10.75	280	2321.90	2354.90
4	30.08, 186.10, 70.41, 171.20, 54.81, 20.80, 20.54, 72.45, 73.03, 24.33, 53.60, 179.78, 132.12, 243.96, 28.56, 146.93, 175.73, 60.22, 83.33, 12.61, 12.82, 77.35, 89.80	14	23.74	320	4830.30	2255.00
5	66.87, 93.31, 37.11, 7.67, 83.30, 12.18, 92.28, 8.09, 51.66, 47.11, 207.92, 6.03, 20.45, 79.00, 173.50, 151.32, 61.84, 152.53, 122.00, 90.14, 199.76, 75.45, 39.29	11	10.03	380	2285.24	2375.44
6	445.59, 51.25, 25.92, 64.70, 72.88, 5.74, 29.97, 28.11, 39.03, 96.97, 86.94, 208.28, 16.78, 12.08, 109.53, 86.17, 4.44, 58.97, 197.06, 197.53, 38.33, 131.59, 38.21, 41.53	22	30.43	320	6102.61	2225.11
7	152.41, 96.79, 67.18, 11.21, 91.03, 172.29, 101.87, 163.30, 20.99, 116.93, 7.32, 10.01, 10.63, 178.06, 46.29, 154.94, 193.43, 119.58, 5.46, 7.99, 88.21, 180.61, 34.44, 71.64, 19.41	10	16.69	380	3550.23	2373.62
8	105.93, 48.95, 192.01, 157.72, 149.28, 85.58, 10.85, 346.57, 105.61, 11.79, 62.48, 1.39, 20.10, 24.14, 350.07, 128.74, 42.27, 24.53, 219.54, 2.97, 116.22	8	19.41	320	4007.36	2400.00
9	158.06, 4.60, 219.83, 51.46, 62.41, 44.41, 54.78, 300.83, 48.91, 34.17, 155.09, 41.17, 30.42, 117.28, 71.89, 111.64, 235.66, 88.01, 191.37	6	10.59	340	2352.09	2360.23
10	301.17, 81.99, 119.41, 6.36, 324.91, 99.42, 378.78, 64.40, 265.36, 100.72, 30.07, 543.21	3	3.07	220	804.05	2343.11
11	174.83, 18.45, 5.71, 185.01, 96.15, 79.31, 164.62, 242.09, 222.67, 41.96, 40.11, 12.35, 149.51, 110.27, 153.68, 197.22, 115.69, 3.31, 14.03, 125.78, 3.51, 5.30, 30.95	7	12.78	380	2808.17	2214.97
12	247.69, 134.71, 280.23, 154.29, 1.13, 167.87, 92.67, 168.30, 10.81, 71.51, 24.26, 18.40, 28.43, 55.58, 59.87, 90.85, 216.95, 145.29, 24.94, 187.69	10	16.85	300	3500.84	2400.00
13	57.78, 233.42, 150.49, 307.64, 73.77, 87.61, 247.16, 221.36, 28.02, 2.33, 291.88, 12.27, 121.58, 26.87, 92.70, 4.30, 84.15	6	13.17	280	2781.79	2386.67
14	280.08, 49.01, 213.50, 8.00, 72.22, 36.29, 60.98, 98.98, 28.35, 154.76, 57.88, 29.23, 97.40, 185.04, 56.85, 84.26, 75.21, 137.79, 55.16, 95.98, 237.58, 167.53	10	3.87	380	1114.80	2336.57
15	199.86, 9.06, 126.64, 72.77, 21.57, 31.46, 223.66, 176.21, 714.63, 8.52, 64.97, 36.27, 171.50, 26.43, 16.90, 142.95, 0.09	13	24.69	220	4910.74	2400.00
16	723.44, 29.36, 34.83, 116.69, 180.14, 86.67, 45.68, 62.61, 10.15, 3.84, 35.77, 9.32, 70.97, 319.03, 68.02, 231.68, 28.81, 174.76, 6.21	14	22.25	260	4487.23	2400.00
17	212.62, 64.28, 15.87, 237.23, 1.95, 89.59, 107.37, 86.74, 13.23, 173.35, 109.33, 12.90, 7.37, 115.84, 64.12, 101.46, 14.62, 36.03, 10.36, 195.54, 7.50, 244.48, 48.31, 23.80, 2.93, 100.43	11	27.81	380	5663.93	2310.66
18	357.42, 77.32, 24.96, 181.80, 17.88, 65.84, 17.93, 12.78, 78.89, 20.66, 292.99, 44.16, 34.09, 117.87, 57.54, 87.83, 327.52, 103.33, 139.78, 49.91	10	38.30	280	7557.62	2333.06
19	329.72, 354.68, 101.09, 15.75, 19.94, 66.84, 106.40, 97.27, 80.29, 89.34, 101.58, 82.42, 533.63, 65.15	3	8.69	240	1891.84	2400.00
20	264.88, 296.10, 20.74, 53.96, 18.89, 242.43, 57.96, 32.71, 60.35, 160.82, 33.24, 76.65, 114.22, 34.37, 0.26, 68.36, 10.95, 54.27, 52.23, 128.66, 2.78, 198.01, 68.03, 17.41, 48.32, 87.58, 24.58, 40.58, 5.35	17	39.88	420	7996.93	2250.00
21	139.50, 12.04, 115.48, 63.27, 11.42, 88.40, 261.62, 24.98, 26.69, 34.63, 273.89, 152.57, 8.24, 61.52, 132.79, 32.87, 353.65, 63.83, 298.22, 27.39	15	35.01	260	6912.41	2310.50
22	347.67, 140.44, 98.38, 13.15, 79.95, 244.90, 30.14, 66.56, 17.25, 143.07, 19.46, 385.80, 32.34, 20.41, 25.92, 140.70, 121.06, 69.76, 10.52, 44.25, 0.43, 106.78, 12.43	16	37.63	280	7428.84	2400.00

APPENDIX 6. Quality gain, and cost in the final inspection-rework facility (improved configuration).

Day i	Standard pipes count n^{st}	Inspected pipes count n^{i}	Number* of External Failures n^{ex}	Number* of Reworked pipes, n^{r}	Number* of Scrapped pipes, n^{s}	External Failure cost @$10,000/pipe	$QL_2^c = n^s(P_p + c_s)L_s + c_R n^r$	Off-line Inspection cost $c_1 n^i$ + 8(C_1)	Total Scrapped length l^s	Length of pipes $P_g = P_a - l^s - n^s L_s$	Actual Production P_a feet	Q = P_g / P_a	QL^c $QL_1^c + QL_2^c$ $/day	COQL $/day
1	53	38	0	6	1	0	8320	4760	67.24	2142.76	2250.00	0.9523	13814.85	7010.00
2	58	39	1	2	4	10000	30640	4860	172.41	2059.32	2391.73	0.8610	33278.63	17251.73
3	58	39	0	1	2	0	15320	4860	90.75	2184.16	2354.90	0.9275	17641.90	7214.90
4	53	42	0	2	3	0	23040	5160	143.74	1991.27	2255.00	0.8830	27870.30	7415.00
5	57	32	2	0	1	20000	7600	4160	50.03	2285.42	2375.44	0.9621	9885.24	26535.44
6	50	31	3	3	1	30000	7960	4060	70.43	2114.67	2225.11	0.9504	14062.61	36285.11
7	57	46	1	3	5	10000	38360	5560	216.69	1956.93	2373.62	0.8245	41910.23	17933.62
8	58	42	4	3	2	40000	15560	5160	99.41	2220.59	2400.00	0.9252	19567.36	47560.00
9	57	37	3	5	4	30000	31000	4660	170.59	2029.64	2360.23	0.8599	33352.09	37020.23
10	58	40	2	6	4	20000	31120	4960	163.07	2020.03	2343.11	0.8621	31924.05	27303.11
11	53	34	1	2	1	10000	7840	4360	52.78	2122.20	2214.97	0.9581	10648.17	16574.97
12	57	47	0	5	0	0	600	5660	16.85	2383.15	2400.00	0.9930	4100.84	8060.00
13	58	38	1	1	0	10000	120	4760	13.17	2373.50	2386.67	0.9945	2901.79	17146.67
14	56	49	0	7	5	0	38840	5860	203.87	1932.70	2336.57	0.8272	39954.80	8196.57
15	56	34	4	0	4	40000	30400	4360	184.69	2055.31	2400.00	0.8564	35310.74	46760.00
16	56	28	6	3	2	60000	15560	3760	102.25	2217.75	2400.00	0.9241	20047.23	66160.00
17	55	37	3	6	1	30000	8320	4660	67.81	2202.85	2310.66	0.9533	13983.93	36970.66
18	55	40	1	3	1	10000	7960	4960	78.30	2214.75	2333.06	0.9493	15517.62	17293.06
19	59	46	3	4	1	30000	8080	5560	48.69	2311.31	2400.00	0.9630	9971.84	37960.00
20	51	38	3	5	3	30000	23400	4760	159.88	1970.12	2250.00	0.8756	31396.93	37010.00
21	54	53	0	5	5	0	38600	6260	235.01	1875.49	2310.50	0.8117	45512.41	8570.50
22	55	33	3	2	1	30000	7840	4260	77.63	2282.37	2400.00	0.9510	15268.84	36660.00

*Simulated assuming binomial distributions, n^r ~ Bin(n^i , p_1'), n^s ~ Bin(n^i , p_1), n^{ex} ~ Bin(n^{st} - n^i , $p_1 + p_1'$).

DECISION SCIENCES INSTITUTE

An Empirical Investigation of Marketing Analytics Orientation

Jeen-Su Lim
The University of Toledo
Email: jlim@utnet.utoledo.edu

John H. Heinrichs
Wayne State University
Email: ai2824@wayne.edu

ABSTRACT

In order to build relationships, personalize products, and enhance customer loyalty, it is important for organizations to utilize data to generate meaningful, fact-based insights. This study focuses on identifying the organization's marketing analytics orientation and the adopted analytical tools used to generate insights. Using the statistical techniques of text-mining and cluster analysis, the marketing analytics orientations of the organization were uncovered. Also, the relationships among organizational characteristics, marketing analytics orientation, and adopted analytical tools were assessed. The results provide insights into organizations utilizing marketing analytics capability and analytical tools.

KEYWORDS: Marketing Analytics Orientation, Analytical Tools, Organizational Characteristics, Text-Mining, Cluster Analysis

INTRODUCTION

The top reason cited for organizations implementing marketing analytics processes was to improve customer relationships. This was followed by two additional reasons: gaining the ability to translate real-time insights into marketing and competitive advantage (Ertemel, 2015) and gaining the ability to process and evaluate the ever increasing, massive amounts of unfiltered and unstructured data (Ghasemaghaei et al, 2018). Further, since organizations gain competitive advantage through improved problem-solving and decision-making performance, advantage goes to organizations that can effectively using the wide-ranging talents of analysts performing varied insight generation tasks (Ghasemaghaei et al, 2018). So, exploiting the vast quantity of available structured and unstructured data simultaneously is both a monumental challenge and a tremendous opportunity as organizations strive to achieve action-oriented, data-driven, decision-making (Doumpos & Zopounidis, 2016).

When organizations operate in a rapidly changing and highly competitive market, studies have reported the positive impact based upon the various insights generated from marketing analytics (Popovič et al, 2012). Yet, which particular marketing analytics orientations are utilized by organizations to produce the insights are not well understood and continues to be an ongoing area of investigation. Further, limited information exists in the literature regarding what and how organizations utilize analytics (Doumpos & Zopounidis, 2016; Kakatkar & Spann, 2019). Therefore, the focus of this research is to provide an awareness of how organizations use analytics and which specific marketing analytics orientations are demanded by organizations. This research contributes to the literature as it seeks to identify which decision areas organizations require marketing analytics for by exploring various marketing analytics orientations. This study also evaluates the various analytical tools managers are provided to perform their insight generation activities. Further, this study intends to investigate whether the marketing analytic orientations utilized by organizations vary by industry classification. The results of this study can provide insights into the determinants and nature of marketing analytics orientations adopted as well as providing insights on the types of selected marketing analytical tools.

LITERATURE REVIEW

Marketing Analytics Orientation

The use of marketing analytics by the organization is of vital importance as marketing analytics are increasingly being used to create knowledge and generate meaningful insights thus leading to more informed, fact-based decisions (Doumpos & Zopounidis, 2016). The marketing analytical tools used by the organization are viewed as well-designed data stores containing volumes of high-quality information connected with user-friendly software tools that can provide managers timely access and intuitive presentation of sought-after information enabling the managers to develop marketing strategic actions or provide the managers with insights to better grasp marketing problems or make better marketing decisions (Popovič et al, 2012).

When defining marketing analytics, it is important to recognize the integration of the concepts of business intelligence and big data along with a strong focus on data collection, model creation, and data analysis all fostered within an evidenced-based organizational culture (Davenport & Kim, 2013). As such, an organization’s marketing analytics orientation can be viewed as a decision area focus utilizing marketing analytics with the practice of first collecting, then measuring, managing, and finally analyzing market performance with the goal to obtain useful insights and maximizing the effectiveness of a particular marketing objective (Cao et al, 2019). The concept of big data adds massive volumes of structured and unstructured data from numerous sources which can then be used to support multiple types of strategic marketing decisions over specified time-frames. Organizations also need to ensure that marketing analytics using these data for fact-based decision-making and problem-solving should be actionable (Davenport & Kim, 2013; Vidgen et al, 2017). To become actionable, the generated insights need to be connected to all aspects of the processes used to enhance problem-solving and marketing decision-making in the organization (Baesens et al, 2016).

Conceptual Framework

This study investigates marketing analytics orientations that organizations may adopt. While organizations can utilize any permutation of these orientations separately or in combination, it is expected that the organization will primarily focus on one orientation when completing marketing analytics tasks. It is also thought that the chosen marketing analytics orientation is influenced by organizational characteristics such as industry classification type. In addition, it has been suggested that the primary marketing analytics orientation utilized by the organization will have an impact on the specific marketing analytics tools utilized.

<u>Marketing Analytics Orientation</u>

An organization can select from a variety of marketing analytic orientations. Several common action-oriented marketing orientations include campaign analytics, digital presence, and product research as well as general marketing and model creation analytics. The general premise for selecting a marketing orientation is that the use of marketing analytics by the organization should lead to improved marketing decision-making and organizational competitiveness, hence, the use of marketing analytics should provide the organization with valuable insights. Thus, an organization's use of marketing analytics can be expected to permit it to achieve sustained competitive advantage from its sensing, seizing, and reconfiguring capabilities as it creates its knowledge base to better sense threats and opportunities in the marketplace (Cho et al, 2019).

Monitoring the competitors' products and gauging the marketplace's reaction to them is an important marketing analytic task. Managers can monitor the competitors' sites, relevant blogs, potential suppliers, and distribution channels to gather product and user information (Vreeman, 2014). This information provides the organization with insight into product features, pricing strategies, and customer feedback thus focusing the organization on discovering competitive advantage (Cho et al, 2019). These insights can also be useful in determining new product strategies (Comm & Mathaisel, 2018) and information for developing product launches. Other important marketing analytic tasks include developing marketing campaigns and sales promotions with key tasks involving determining which campaigns are the most convincing and which analytic measures should be used for the marketing campaigns (Liu & Burns, 2018).

General modeling and insight generation is viewed as a marketing task when data mining is a central part of analyst's knowledge discovery and seeks patterns and knowledge from large amounts of data. The goal of this marketing task is to enhance customer relationships with more accurate response models. While using the explanatory response model is something most marketing managers are accustomed to, additional care should be taken when using predictive analytic models as they are different from explanatory modeling in many aspects.

The organization also seeks to optimize their digital presence to improve their rankings with search engines as it has been shown that positive optimization increases both search rankings and visitor satisfaction. Therefore, understanding the various elements contributing to search engine rankings and optimizing those elements can help the organization promote its products and improve its brand visibility, hopefully leading to an increase of sales and profits (Liu & Burns, 2018; Rosenkrans & Myers, 2018). This study suggests that organizations have specific marketing analytics orientations as the managers generate marketing insights, solves problems, and focuses on decision making. The particular marketing analytics orientation employed, focuses the organization on various marketing analytics tasks in terms of what is to be accomplished by the specific marketing program.

Organizational Characteristics

It has been argued that an organization's size, number of skilled personnel, and organizational culture contribute to the successful adoption and usage of marketing analytics (Trieu, 2017). Further, Gupta & George (2016) highlight that an organization needs to possess a unique blend of cultural, financial, and human resources to create a marketing analytic capability so that their competitors will find it difficult to replicate. To accomplish creating this capability, however, requires that an organization either needs to improve the managers' analytic skills via training or search for individuals who can fulfill the organization's marketing analytics demands. Ideally, marketing managers with improved competencies will provide improved insight generation capabilities for the organization (Ayyagari et al, 2011; Davenport & Patel, 2012). To accomplish this goal of creating a marketing analytics capability for the organization requires that the organization has a clear understanding of both the marketing analytics objectives and the connection between the expected outcome and the skills required to achieve it (Ghasemaghaei et al, 2018). In various studies, an increase in investments related to the budget for education and training have been reported by successful organizations (Trieu, 2017).

Yet, not everyone embraces this data-driven culture. Certain industries are characterized as more data-driven and have adopted advanced marketing analytics tools and approaches. For example, although many healthcare organizations have adopted marketing analytics to support their processes, few of these organizations had significant experience using marketing analytics (Wang et al, 2018). Thus, various industries have adopted different levels of fact-based decision-making and problem-solving which ultimately leads to the use of diverse marketing analytics orientations. Yet, for many organizations, the marketing analytics use is still in its initial stages (Lismont et al, 2017).

Marketing Analytics Tools

Organizations acquire or develop appropriate marketing analytics tool capabilities considering their specific marketing analytics orientation. Ghasemaghaei et al. (2018) argue that there is a strong connection between the marketing analytic tool sophistication and the quality of insights capable of being achieved by the organization. This suggests that these analytical tools can help the organization improve its performance. Wang and Byrd (2017) report that the effective use of marketing analytics tools indirectly influences problem-solving and decision-making efficiency and effectiveness. Thus, it is important that the tools have the right capabilities that can facilitate the marketing insight generation.

Marketing analytics capabilities encompass a portfolio of various analysis methods, tools, and data transformations. Three specific types of analytic capabilities employed by organizations include descriptive, predictive, and prescriptive methods (Bayrak, 2015). Additional analytic capability examples include content analytics, speech analytics, text analytics, dashboards, and visualizations. The decision capabilities of analytic tools include applications to help analysts learn, share, and understand data. Decision capabilities examples include scenario modeling and simulation thus allowing the analysts to communicate the results that could optimally balance required trade-offs. The information capabilities focus on the technology infrastructure that ensures proper interfaces with other parts of the organization permitting the exchange of data, documents and other content. The level of capabilities achieved by organizations must support the goals established for the marketing analytics orientation. Organizations may also combine these tool capabilities by establishing proactive use in which insight generation is based on rationality as it is believed that managers will adopt and use the organization's analytics tools to a greater extent if the right analytics culture is in place (Popovič et al, 2012).

METHODS

Data Collection

Marketing analytics orientation descriptions were created through the extraction of analytics job postings. During 2017, 286 marketing analytics job postings were collected by using various online job search sites. The job postings showing a high relevance or containing marketing analytics job titles were then selected. The organization's name from the job postings were used to identify industry type. The identified organizations were classified into one of the four industry classification types (consumer products, industrial products, financial and health services, and retail) by finding their SIC code as well as exploring their primary markets served.

Text Mining Process

A four-step text mining process (Inzalkar & Sharma, 2015) was used in this study. The first step in this process was to collect marketing analytics descriptions from the job postings. Over a two-month period, marketing analytics job postings that were highly relevant for marketing analytics work were collected. The second step in the process was to prepare the collected text using text refining and pre-processing. Text refining and pre-processing was accomplished by clearing the unnecessary formatting as well as implementing tokenization and stemming for text reduction. The third step in the process was to perform text-mining analysis using the refined and pre-processed texts. The text mining analysis generated document-based intermediate output that contains the relevant term concepts related to the marketing analytics components. The final step was to perform clustering analysis. The clustering analysis provided the marketing analytics orientation clusters to be used for insight generation.

RESULTS

Cluster Analysis Results

The k-means clustering procedure was chosen to perform this analysis. Five unique clusters from the clustering procedure were generated. 19 relevant word concept terms were used in generating the five clusters. These word concept terms are related to five marketing analytics orientations. To characterize the resulting clusters, the cluster centroids were examined thus allowing interpretation of what each cluster grouping represented.

The product research analytics orientation cluster with 44 organizations became cluster 1. The word concept terms showing the highest values for this cluster were product and research. This cluster shows large centroid values greater than .500 for word concepts which are relevant for product development research analytics. The second cluster contained 142 organizations and represented a general marketing analytics orientation. The word concept term list showing the highest values greater than .240 for this cluster were the following three term concepts: market, data, and insight respectively. This cluster does not have the highest centroid values indicating no specific analytics focus and this cluster was characterized by relatively small centroid values for most of the word concept terms. The third cluster contained 37 organizations and is referred to as the campaign and sales analytics orientation cluster. The word concept term list showing the highest values greater than .500 include test, campaign, report, insight, and lead respectively. These concepts are primarily related to promotional campaigns as well as sales lead analytics. The fourth cluster represents data modeling analytics orientation and contains 39 organizations. The word concept term list showing the highest values greater than .500 include the following terms: model, data, and insight. These concepts are related to developing insightful models and generating forecasts using data. The fifth cluster contains 24 organizations and was referred to as the search engine and web analytics orientation cluster. The word concept term list showing the highest values greater than .500 include SEO, SEM, campaign, web, and traffic respectively. These concepts were primarily related to online and web search related analytics. These results highlight that approximately half of organizations are classified into the general marketing analytics orientation cluster whereas the other half are classified into clusters related to the four specific analytics orientation clusters.

Cross-tabulation Results

This study evaluates the relationship between the marketing analytics orientations and organizational characteristics as well as between marketing analytics orientations and marketing analytic tools. Chi-square statistics for cross-tabulations were used to test these relationships. The chi-square test also indicates a significant relationship between marketing analytics orientations and industry type (chi-square = 21.54, $p < .05$).

Insights can be ascertained by first examining each cluster to determine the key industry, then examining each industry to understand the desired marketing analytic orientation pattern. Thus, examining each cluster, the results show that for the first cluster, industrial product and retail organizations belong to this group and utilize product-related analytic orientations. Examining the second cluster highlights that approximately 55% of organizations in all four industry types are in this group suggesting the universal utilization of general marketing analytics orientation. More service organization belong to the third cluster which highlights the utilization of promotional campaign and sales lead analytics orientation. Consumer product and retail organizations tend to dominate the fourth cluster as those organizations utilize data and forecasting model analytic orientation. The fifth cluster primarily includes industrial product organizations as they utilize search and web related analytic orientation.

All organization utilize general marketing analytic orientation. In addition, consumer product organizations primarily focus on data modeling analytic orientation indicating required insight generation through data modeling. Industrial product organizations focus on product research analytics as well as search engine and web analytics as they focus on product research and competitive intelligence and then on ensuring they have a digital presence to get found on the internet. Service industry organizations are focused on campaign and sales analytic orientations to gain insights utilizing A/B testing and lead generation in a relationship-oriented industry, whereas retail organizations primarily focus on product analytic and modeling/forecasting orientations to capture consumer trends and manage inventories. The chi-square test for the relationship between marketing analytics orientations and marketing analytics tools used by organizations indicates a significant relationship between marketing analytics orientations and the marketing analytics tools used (chi-square = 247.01, $p < .01$). The selected organizations reported a total of 21 different analytics tools in their job postings. The results show that organizations are using a diverse set of marketing analytics tools. These tools include generic business analytics tools such as Python and SQL as well as more commonly used marketing analysis software such a SAS or SPSS.

DISCUSSION AND CONCLUSIONS

This study developed and tested a conceptual model capturing the relationships among marketing analytics orientations, organizational characteristics, and marketing analytics tool adoption. Text mining analysis of analytics job postings together with cluster analysis results provide interesting insights into the nature of marketing analytics orientations and tools being adopted by organizations. Evaluating job postings provides a leading indicator view of where organizations desire to be on their marketing analytics adoption and usage journey. These are the marketing analytics orientations that organizations view as important to possess to meet their strategic goals. Each word concept term the organizations uses in their job postings belongs to a specific cluster. The cluster analysis results identified five marketing analytics orientations that organizations are utilizing for their insight generation. These results provide important theoretical implication.

This study empirically identifies five marketing analytics orientations. Text mining analysis was performed to identify the relevant concepts used by organizations to describe marketing analytics orientations required for marketing analytics professionals. Text mining output was further analyzed using clustering analysis to identify marketing analytics orientations of the organizations. The cluster analysis results show that the largest cluster containing about half of the organizations focused on the general marketing analytics orientation. The next largest clusters, in order, were the product research analytics, data modeling analytics, and campaign and sales analytics orientation clusters with 15.4%, 13.6%, and 12.9%. The smallest cluster with 8.4% of organizations was the search engine and web analytics orientation cluster. These results suggest that organizations are utilizing various marketing analytics orientations for marketing decision-making. This is important because as Wedel and Kannan (2016) discuss, “one-size-fits-all analytic solutions are neither desirable or likely to be effective”. For organizations, the movement toward marketing analytics is impacting organizational processes, rebalancing the power of relationships in decision-making, and altering the scope and scale of optimization challenges (Baesens et al, 2016). Drawing out decision-ready inferences from marketing analytics influences and enhances the organization’s marketing decision-making (Baesens et al, 2016).

This study also evaluates antecedents and consequence of an organization’s marketing analytics orientations. The antecedent industry classification type is evaluated to assess whether they are related to marketing analytics orientations. The results show that marketing analytics orientations are related to the industry classification type. There exist significant differences in utilizing various marketing analytics orientations by the type of industry the organization resides. As certain industries utilize different marketing analytics orientations, organizations need to develop unique marketing analytics competencies to remain competitive in the market. The results show a significant relationship between marketing analytics orientations and marketing analytics tools used by organizations. Interestingly, the results show that there are no dominant marketing analytics tools used by organizations. These findings are consistent to the reported statistics suggesting that the marketing analytics tool market is highly fragmented (Columbus, 2016). Turck (2016) highlights that the analytic tool landscape focused on various areas such as artificial intelligence, log analytics, machine learning, mobile analytics, social analytics, statistical computing, visualization, and web analytics. Yet, as Wheatley (2016) reports, the software market that drives these marketing analytics tools is expected to surge by a 23% compounded annual growth rate by 2026 with analytics tools brought about by the emergence of cloud-native business analytic startup firms.

The findings from this study provide important managerial implication regarding marketing analytics training and education. The results can be used in developing the appropriate approach in marketing analytics training and education as multidisciplinary skill set are required for effective marketing analysts (Wedel & Kannan, 2016). As no specific marketing analytics tools are dominantly used by organizations, tool independent training may be needed for marketing analytics competence. This means that marketing analytics training should focus on the basic mechanisms and processes of marketing analytics rather than tool specific marketing analytics knowledge (Turel & Kapoor, 2016). The cluster analysis results suggest that various marketing analytics orientations are utilized by organizations in different industries. These findings shed light on the nature and content of marketing analytics knowledge and training required. While this study provides valuable insights, the results are based on the marketing analytics job posts of 286 organizations. The findings may not be generalizable given the limited nature of this sample. Future research should further validate the findings of this study and extend the study by evaluating marketing analytics orientations utilized for specific marketing problem-solving and marketing decision-making tasks.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Predicting Consumer Behaviors: Active Innovation Resistance Modeling using Agent-Based Simulation

Shehrose Javaid
Illinois State University
Email: sjavaid@ilstu.edu

Dr. Borinara Park
Illinois State University
Email: bpark@ilstu.edu

ABSTRACT

Innovation rejection is one of the biggest issues of marketing research and active innovation resistance has been identified as one of the main driver behind conscious innovation rejection. The current research that is focused on understanding active innovation resistance, has produced significant findings. A typology of functional and psychological barriers has been developed to comprehensively explain active innovation resistance. However, no effort has been made to apply these findings to the real world. Simulation is a great tool that can be used to apply theory to real world scenarios. A simulation model that uses the aforementioned typology to explain innovation rejection remains unexplored. We propose a model that uses agent-based simulation to predict the rejection of innovation by a group of consumers. This model provides insights into the reasoning behind these rejections and how they relate to the proposed typology. The model uses statistical data about the typology to predict the active innovation resistance values of its population of agents. The model also uses agent networks to inform the values of the barriers that depend on the social groups of the consumer. The proposed model can be beneficial to companies aiming to reduce the uncertainty associated with launching innovation. This study opens the gate to further study into active innovation resistance working in tandem with simulation.

KEYWORDS: Agent-based Simulation, Active innovation resistance, Simulation modeling, Structural equation modeling, marketing research

INTRODUCTION

For companies with an established consumer base, introducing an innovation can be a risky and uncertain process. This innovation can be a new product or a new service but the reaction of consumers remains largely unpredictable (Castellion & Markham, 2013). For companies that want to remain to be on the cutting edge and retain their competitiveness, this uncertainty is troublesome (Hansen, McDonald, & Mitchell, 2013). A failed innovation can cause a company to hurt their brand's image and lose the trust of investors (Liao, Chou, & Lin, 2015). Product failure rates are on the rise, and big companies are becoming more and more reluctant to innovate (Castellion & Markham, 2013). Therefore, reducing innovation failure rate is one of the biggest challenges of marketing research (Henard & Szymanski, 2001; Joachim, Spieth, & Heidenreich, 2018). This study will attempt to tackle the aforementioned problem of innovation failure. Our objective is to reduce the unpredictability of consumer behavior and provide insights into the potential causes behind consumer rejection.

Using the literature, we can conclude that the majority of innovation failures are caused by consumer rejection (Kleijnen, Lee, & Wetzels, 2009; Laukkanen, Sinkkonen, & Laukkanen, 2009). Consumer rejection occurs after the consumer has evaluated the proposed innovation. Talke and Heidenreich (2014) talk about how rejection is a direct result of conscious decision-making by the consumer. According to their research, this conscious decision is a vital step in the evaluation phase of the adoption process of the consumer, they call it active innovation resistance (AIR). AIR is the dependent variable that we will be studying in this research. The independent variables will be the typology of barriers, introduced by Talke and Heidenreich (2014), that directly affect AIR.

The purpose of this study is to propose an agent-based simulation (ABS) model that can determine the probability of rejection of an innovation by predicting the active innovation resistance of a group of consumers. It will also highlight the reasons behind the rejection and provide valuable insights into innovation rejection. ABS is a paradigm of simulation that focuses on the dynamic behavior of agents in an environment. The individual decision-making model of agents determines the behavior of the overarching system (Gilbert, 2008). The focus on micro-behavior makes ABS a great tool for marketing research (Meyer, Lorscheid, & Troitzsch, 2009). This is because these micro-behaviors can be used to simulate observed market phenomena. Marketing research has made significant strides by using simulation to better understand the adoption process of the consumer (Meyer et al., 2009; Schubring, Lorscheid, Meyer, & Ringle, 2016). However, the same cannot be said for consumer rejection and there is still a lot of room for improvement (Talke & Heidenreich, 2014).

The hyper-competitive world of business makes innovation a necessity for companies to in order to be competitive. Even big companies face uncertainty in terms of innovation success, and with failure rates standing at 40%, the risk factor seems to be uncomfortably high (Castellion & Markham, 2013). A simulation model that can map consumer behavior could be a very useful tool in marketing research (Meyer et al., 2009). It can help companies reduce the level of uncertainty, associated with innovation, and quantify their performance expectations for said innovation. To accomplish the aforementioned objectives of this study, a set of research questions have been proposed. These research questions are listed in Table 1.

Simulation modeling is a great tool for marketing research. A simulation is only as good as the data that is used to build it. The data used in this study will be taken from a previously published research study done by Joachim, Spieth, and Heidenreich (2018). The data sample is a cross-sectional data sample and represents a population of 1000 individuals. Additionally, the model assumes that consumer behavior is limited to the prescribed variables. This is a big assumption because in the real world there are many other factors that affect consumer behavior. However, this assumption is necessary because one of the goals of this study is to determine which of the independent variables has the most significant effect on AIR. This study chooses to limit agent behavior in this way because of the findings of Joachim et al. (2018). They provide evidence for the comprehensiveness of the AIR barrier typology that we will use in this model. Based on this evidence, we believe that the limitation of agent behavior is valid. To conclude, the limitations of the proposed model are as follows. The model is a meager representation of the real world due to the limited behavior of its agents. The model is not too generalizable because of the limited nature of the data available. The model does not take other factors that affect innovation resistance into account. However, the proposed model will still offer insight into innovation rejection and is a worthwhile effort.

The structure of the paper is as follows. First, we elaborate on the concept of ABS and its application in the field of marketing research. Second, through the use of literature, we discuss the context of consumer behavior that is relevant to this study. Additionally, we operationalize our definition of active innovation rejection and define the relationship between our dependent and independent variables. Third, we describe the methodology and the tools that will be used to make the ABS model. Finally, we discuss the outcomes and benefits of the proposed model.

Table 1
The proposed research questions.

- **Research question 1.** What is the likelihood that the proposed innovation will fail?
- **Research question 2.** What portion of the consumer base will reject the innovation?
- **Research question 3.** What are the reasons behind the rejection of the proposed innovation?

LITERATURE REVIEW

Agent-based simulation

Agent-based simulation modeling uses simulation experiments to investigate the dynamic behavior of real world situations. This dynamism has led to its firm establishment in social sciences, as a useful simulation paradigm in the field (Meyer et al., 2009). The rudimentary idea behind ABS modeling is to specify micro behavior of agents and use them to model a developing phenomenon, such as the behavior of a consumer market base or social group (Gilbert, 2008). ABS is a great tool for studying Innovation diffusion and other marketing phenomena (Garcia & Jager, 2011). This paper intends to use ABS to simulate consumer rejection to the innovation. In this scenario consumer networks hold particular relevance to the study; they are ideal for simulating social groups of consumers. ABS is the perfect tool to explore the interactions and behavior of consumers at a micro-level (Gilbert, 2008). Despite the availability of data and theory, the failure of innovations is still largely unpredictable (Hansen et al., 2013; Hui, Goldberg, Magdon-Ismail, & Wallace, 2010; Repenning, 2002). ABS can provide verifiable insights that can help assess the causes and effects of rejection. Therefore, this study will use data and theory to create an ABS model that explores the reasoning behind consumer behavior, with respect to innovation.

Consumer Behavior in the context of innovation

The study of consumer behavior, with regards to innovation, has evaluated many causes and effects. These range from acceptance to rejection, delay, and even opposition (Davidson & Walley, 1985; Eric A. Greenleaf & Donald R. Lehmann, 1995; Nabih, Bloem, & Poiesz, 1997; Rogers, 2003). These are behavioral outcomes that classify a consumer’s response to innovation. Therefore, it can be said that the success or failure of a new product or service depends on the consumer’s perception of it (Heidenrich & Spieth, 2013). During the evaluation phase of the adoption process (Talke & Heidenreich, 2014) the consumer forms attitudes towards the innovation. A positive attitude can lead to innovation acceptance, whereas a negative attitude can lead to innovation resistance (Nabih et al., 1997). Thus, the consumer’s mindset in the decision stage of the adoption process (Kuisma, Laukkanen, & Hiltunen, 2007) is a direct consequence of these attitudes (Rogers, 2003). These negative attitudes lead to active innovation resistance by the consumer (Talke & Heidenreich, 2014).

The barriers of active innovation resistance

The term “active innovation resistance” refers to the conscious decision of rejecting the innovation, that the consumer makes, after evaluating a new product or innovation (Kleijnen et al., 2009). Ram and Sheth (1989) first proposed a framework of five barriers and discussed its relationship with innovation resistance. Later, Talke and Heidenreich (2014) provided a much more comprehensive framework of seventeen barriers that contribute to active innovation resistance. The framework of barriers introduced by Talke and Heidenreich (2014) is a thorough and complete typology of barriers that directly affect active innovation resistance (Joachim et al., 2018). There are eight psychological barriers, that “arise as soon as the innovation conflicts with a consumer's social norms, values, or individual usage patterns, or if its usage is perceived as being too risky” (Talke & Heidenreich, 2014, p. 899). Research by Joachim, Spieth, & Heidenreich (2018) provides an excellent collection of definitions for the aforementioned typology. The following is an elaboration of the definitions compiled by Joachim et al. (2018). The psychological barriers are a conceptual take on the consumer’s reaction to the innovation and how it relates to their social environment. The *personal risk barrier* is a psychological response to the perception, that the innovation could cause personal harm to the consumer of their property. Similarly, the *usage barriers* arise if the consumer perceives that the innovation will cause a disruption in their regularly established usage patterns and routines. Both of these barriers arise from the natural fear of change or the unknown that humans feel. The *social risk barrier* arises from a fear of rejection, from the consumer’s social group, as a result of adoption to the innovation. Similarly, the *norm barrier* results from the perception that the innovation conflicts with social norms or established traditions within the consumer’s social environment. If the innovation is associated with a brand or country that is perceived negatively by the consumer and their social group, this culminates in the rise of the *image barrier*. These three barriers are strongly dependent on the reaction of the consumer’s social group to the adoption of the innovation (Kleijnen et al., 2009). If adoption of the innovation means hurting their social standing, the consumer will be more likely to reject the innovation. The *economic risk barrier* stems from the fear that the innovation is not worth the financial investment that it demands. Finally, rejection can also be attributed to the fear of the innovation being dysfunctional. This is known as the *functional risk barrier*. All of these psychological barriers affect the customer’s decision to adopt an innovation or to reject it (Ram & Sheth, 1989). Additionally, there are nine functional barriers, that come into play “as soon as a consumer perceives any product attribute as dysfunctional or inadequate for their individual requirements and usage expectations (Talke & Heidenreich, 2014, p. 899). Functional barriers are all about how the consumer perceives the traits of the innovation. The *value barrier* depends on how the consumer compares the value of the innovation to its competitors and its precursor. If an innovation is complex and difficult to understand it will face the *complexity barrier*. Furthermore, if the consumer feels that the innovation doesn’t work well with the other products that they use this will result in the *compatibility barrier*. Additionally, if the consumer finds that the innovation is incomplete or needs to be supplemented, to realize its full benefits, then this will lead to the *Co-dependence barrier*. If it takes too long for the innovation to provide the consumer with noticeable benefits, then it makes it more likely that the consumer will reject it, this is the essence of the *realization barrier*. If the innovation is difficult to observe it will face the *visibility barrier*, and if the consumer finds it hard to share the benefits of the innovation with others, it will face the *communicability barrier.* If the consumer perceives that they have missed the chance, or have not gotten the chance, to try the innovation before adopting it, this will result in the *trialability barrier.* Finally, if the innovation cannot be modified to fit the requirements of the consumer than the *amenability barrier* applies, consumers want to use products and services that fit their needs well (Kleijnen et al., 2009).

How the barriers relate to active innovation resistance

According to the findings of the empirical study done by Joachim et al. (2018), all seventeen barriers have shown positive effects on AIR. Hence, we can definitively say that they have a positive association with the AIR. Furthermore, Joachim et al. (2018) also found that some barriers are more significant to AIR than others. The norm barrier, the value barrier, and the communicability barrier rank the highest, with respect to importance, among all seventeen barriers. Additionally, even the barrier with the least negative effect, such as the co-dependence barrier, show substantial influence on AIR. The results published by Joachim et al. (2018) can be seen in Table 2. To address the aforementioned problems this study proposes a simulation model that will use the AIR barrier proposed by Talke and Heidenreich (2014) to predict the AIR of its agents in response to an innovation.

Table 2. Structural model results of functional and psychological barriers

Barrier Type	Barrier Name	Path C ($\beta_{Prod.}$)
Psychological Barriers	Norm barriers	−0.676
	Economic risk barriers	−0.456
	Image barriers	−0.439
	Social risk barriers	−0.374
	Functional risk barriers	−0.348
	Personal risk barriers	−0.271
	Usage barriers	−0.225
	Information barriers	−0.186
Functional Barriers	Value barriers	−0.601
	Communicability barriers	−0.475
	Trialability barriers	−0.470
	Amenability barriers	−0.450
	Compatibility barriers	−0.349
	Complexity barriers	−0.283
	Visibility barriers	−0.276
	Realization barriers	−0.229
	Co-dependence barriers	−0.142

Note. Adapted from “Active innovation resistance: An empirical study on functional and psychological barriers to innovation adoption in different contexts”, by Joachim, V., Spieth, P., & Heidenreich, S., 2018, *Industrial Marketing Management, 71*, p. 104

SIMULATION FRAMEWORK

In this study, we plan on using the work of Joachim et al. (2018) as the basis of our simulation model. They provided empirical evidence for each barrier in the typology and its relational significance to AIR. The results published by Joachim et al. (2018) is produced by applying partial least squares structural equation modeling (PLS-SEM) to data they gathered through surveys. PLS-SEM is a statistical tool that can be used to model the relationship between variables. The results produced by Joachim et al. (2018) will provide a solid and practical basis for the simulation model. The use of this data will provide validity to the model. Validity is one of the most important aspects of simulation modeling. If a model is not validated by literature and real data, it is essentially useless. By using the published work of Joachim et al. (2018) as a basis for the simulation model, we have affirmed its validity. The innovations that were used by Joachim et al. (2018), for their survey of innovation rejection, were as follows: “the dual tablet Toshiba Libretto, the 3-D camera Fujifilm Finepix 3D and the holographic laser projector Blueoptics Light Touch" (Joachim et al., 2018, p. 103). However, we believe that using a fixed set of innovations would limit the generalizability and usefulness of the model. Therefore, instead of testing these innovations on the model, we will conduct simulation experiments by testing a combination of the attributes of innovations. Each test will have a combination of high, medium, and low values for innovation attributes, by doing this we can test the effect of these attributes on the AIR of the consumer. Functional barriers will be influenced by the functional attributes of the innovation. The most important aspect of the proposed model will be the behavior of the consumer (agent). The data provides us with a strong base for our model, but we also need a robust structure. The core logic of the agent behavior will act as our model's skeletal structure. "The PLS path model is the modeling anchor for the specification of agent behavior" (Schubring et al., 2016, p. 4607). Schubring et al. (2016) provide formulae for using the PLS-SEM path model to dictate agent behavior, they are as follows.

$$Active\ Innovation\ Resistance = \frac{\sum_{i=1}^{n} attribute\ score_i \times effect_i}{\sum_{i=1}^{n} \max(attribute\ score_i) \times effect_i} \quad (1)$$

The agent's actions are determined by the value of "*Probability (action)*". *Probability (action)* describes the activity of an agent. For example, based on the value of "*Probability (action)*" the agent will decide to either buy or not buy. In our study, this can be used to evaluate AIR, our dependent variable. Therefore, we will be able to use the PLS-SEM data published by Joachim et al. (2018) to define agent reasoning.

We provide an example to convey a better understanding of the aforementioned formulae and how we can measure our data. Each agent in the model will have a perception of the proposed innovation. We know, from the work of Talke and Heidenreich (2014), that this perception can be boiled down to seventeen barriers, when considering AIR. These barriers will act as our independent variables. Each of them will have an "*attribute value*" which will vary for each agent and be determined by the perception of the said agent in the simulation. The attribute value can range from 1 to 10. A high "*attribute value*" will mean that the agent perceives that barrier to be highly influential and vice versa. "*total effect*" will be the individual significance of each attribute, with respect to AIR. This will be determined from the data presented in Table 2. For AIR, let's take the example of *value barrier* and *norm barrier*, their attribute values can be 8 and 2 respectively. From table 2 we can see that their total effect is 0.601 for the value barrier and 0.676 for the norm barrier. We ignore the negative sign because we are calculating rejection rather than adoption. We also want to define that we are considering rejection by the consumer to be a result of the initial contact with the product. We are not considering any future contact that the customer will have with the product in question and how that might affect their decision. The latent variable score for value barrier and norm barrier will be 6.16, this value requires normalization, as a standard, we use the maximum value ($max\left(latent\ variable\ score_j\right)$). The maximum value in this example is 12.77 and thus, the value for AIR will be 48% (6.16/12.77). An AIR value of less than 50% will mean that the consumer will not reject the innovation.

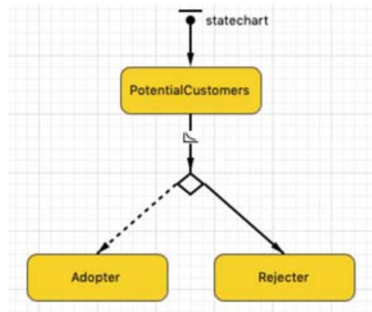


Figure 1. Decision model of agents

The decision model of an agent will have three distinct states. Based on the AIR value of each agent, the agents will move from the *PotentialCustomer* state to either the *Adopter* state or the *Rejecter* state. This decision model is represented in Figure 1. The number of agents that reject the innovation is the evaluation measure for the proposed innovation. In addition to the decision model, there is also a network that connects agents and dictates their interaction with each other. Like many ABS models we use a stylized random network (Trusov, Rand, & Joshi, 2013). This network will be used to simulate the social network of consumers. Some AIR barriers depend on the behavior of the consumer's social group, using the aforementioned networks we can simulate these social groups and get values for these barriers. These psychological barriers will be influenced by the properties of the agent and the abovementioned social network. We have provided sequence diagrams that provide a better understanding of the interaction and behavior of agents. Figure 2. shows the decision logic of the consumer in the model and the sequence of events that result in innovation rejection or adoption. We can see that the consumer makes a conscious decision to either reject or adopt the innovation after they have gained information about the functional traits of the innovation and evaluated it.

In figure 2, we can also see the sequence of events that lead to a *PotentialConsumer* rejecting the innovation due to advice from rejecters in their social group. If a rejecter feels that an innovation is too expensive, they will relay this message to people in their social group. Based on the established contact rate this message will result in other *PotentialConsumers* rejecting the innovation as well. We can see in figure 2, a rejecter will think that the innovation is too expensive if the perceived value of that innovation is less than 2 out of a possible 10.

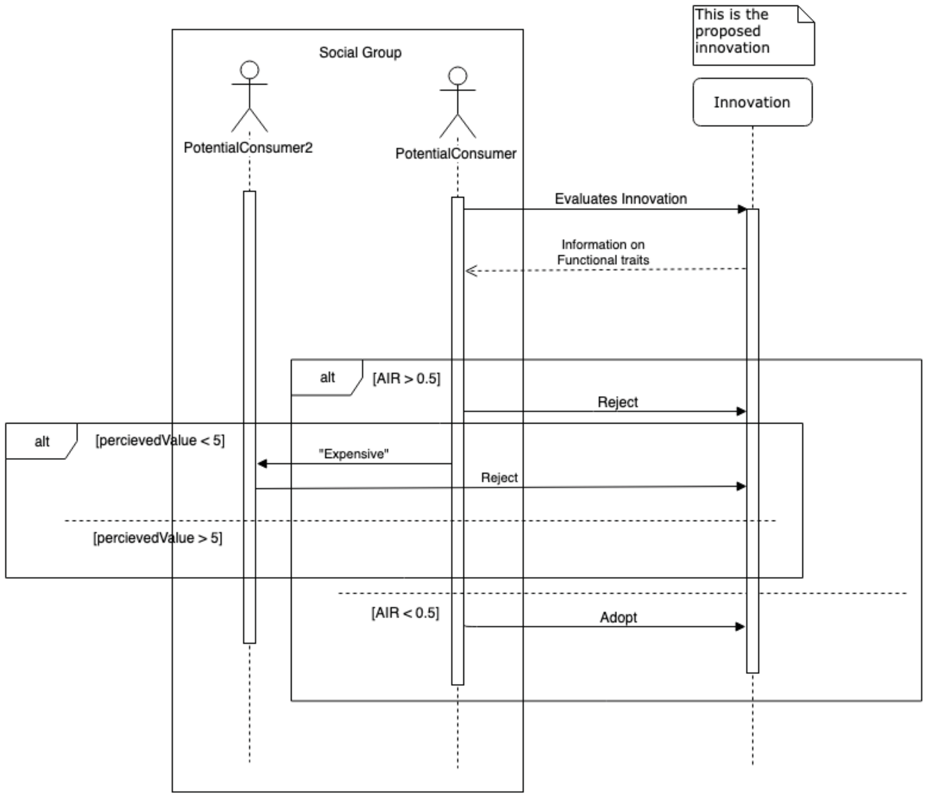


Figure 2. A sequence model that the decision and interaction model of agents (showing the sequence of actions that leads to rejection by agents and the interaction between agents)

Additionally, to provide a graphical visualization of the population of agents a rectangular space is used. Inside this rectangular space, we have our population of agents, rejecters are represented by red and adopters are represented by green. Furthermore, the lines between the agents represent the network they are connected with. We can see this visualization in Figure 3.

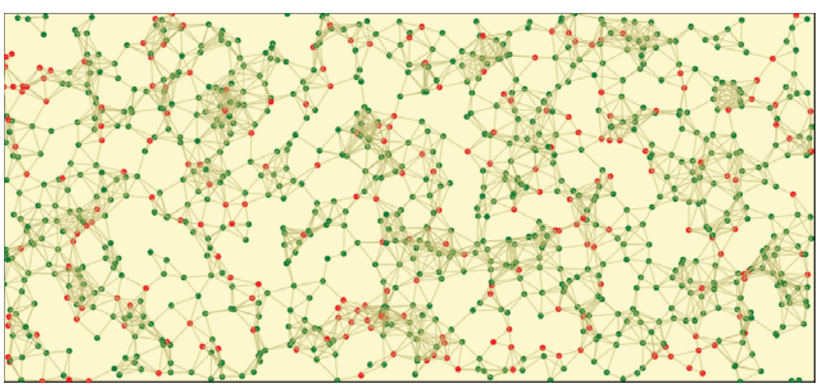


Figure 3. Visual representation of the population of agents and their network

We have decided to use Anylogic 8 to build the proposed model. Anylogic 8 is an industry standard tool for simulation modeling. It provides the depth and computation support necessary for this project. Anylogic 8 also provides integrated analysis tools to track the behavior of the simulated consumer base. Each agent will have behavior in accordance with the reasoning described in the previous section. Anylogic 8 will keep track of the most significant barriers that resulted in rejection. This will give us an insight on what aspect of the innovation causes influences the decision of the consumer. These insights can be used to improve future innovations and help mitigate the risk associated with launching innovations.

SIMULATION EXPERIMENTS AND RESULTS

The simulation experiments explore the consumer reaction towards multiple proposed innovations. The experiments will focus on the relationship between the innovation characteristics and the corresponding AIR of the consumer population. The results will help us better understand the model. The independent variable will be the functional traits of the innovation, derived from the functional barriers in the typology proposed by Talke and Heidenreich (2014). The notion is that an innovation can be represented by these attributes that affect the perception of the barriers in the mind of the consumer and it is this perception that leads to adoption or rejection. This experiment will vary the attributes of the innovation to high, medium and low values. Each innovation attribute can have a value ranging between 1 to 10, where a value of 10 has the maximum effect on AIR. We will use the values 2, 5, and 8 as the test values. This range of values is a conceptual definition. Varying these values will allow us to represent a number of different prospective innovations in the context of the model.

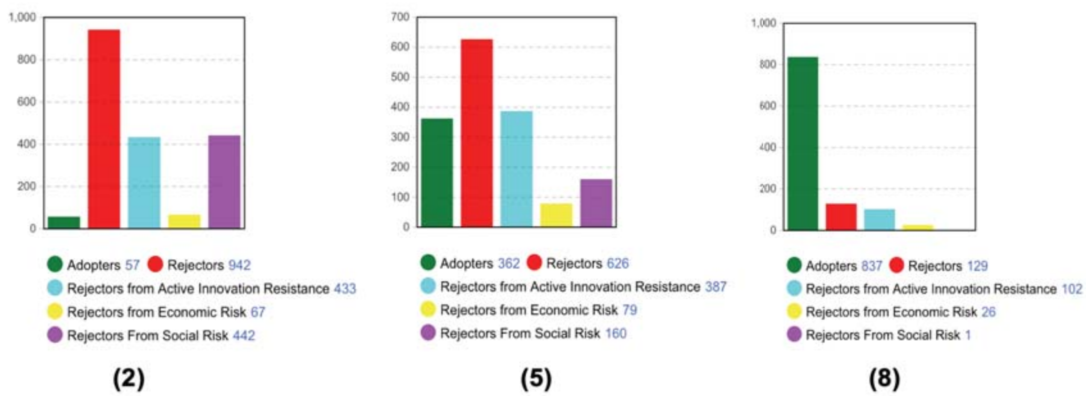


Figure 4. The bar charts show the comparison in the number of rejecters and adopters, in the agent population, for the innovation traits 2,5, and 8

In figure 4 we can see the change in the number of adopters and rejecters as we vary our independent variables. It can be concluded that, an innovation with higher values for its functional traits will result in lower values for the perceived barriers that are derived from those traits. Furthermore, these perceived barriers will result in less active innovation resistance and the proposed innovation will be largely adopted. Additionally, we can also see that our independent variables, not only affect rejection from active innovation resistance, but also have an effect on the number of rejecters from other sources. We believe that this effect is due to the initially higher number of rejecters from active innovation resistance caused by low functional traits. It seems that the initial rejecters spread their influence throughout the population. This seems to have a sort of domino effect on the entire population of agents. Which is apparent in the small number of adopter in the left most graph in figure 4. Figure 5. reaffirms these findings;

we can see that the rate of diffusion of rejecters increases as we increase our independent variables.

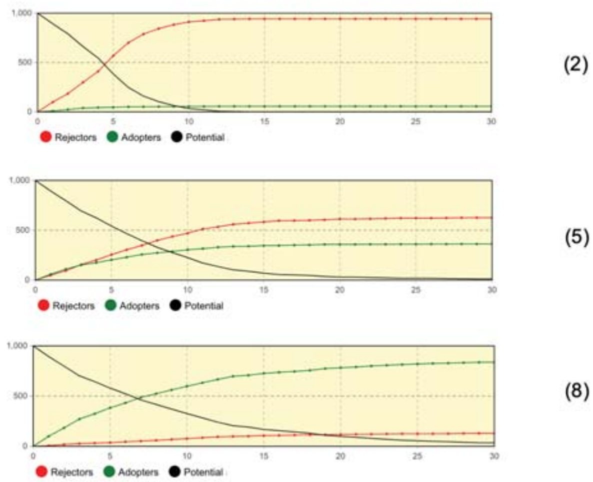


Figure 5. The diffusion of innovation with respect to time
(The x-axis shows the progression of days the model simulates and the y-axis shows the number of consumers, either adopted (green), rejected(red) or are potential consumers (black))

These results show how the change in the attribute of the innovation effects the behavior of the consumer base. Hence, by using the barriers of resistance we have defined attributes for an innovation that can directly affect the rejection it faces by consumers.

Figure 6. The attributes of innovation that are derived from the typology of active innovation resistance barriers

DISCUSSION AND CONCLUSIONS

This paper uses the predictive qualities of the data provided by Joachim et al. (2018) to construct a simulation model. This model has helped us reduce the unpredictability of innovation by providing a set of attributes that directly affect consumer behavior. We used simulation experiments to study the change in consumer behavior with respect to the change in the attributes of the innovation. We can conclude that the nine functional attributes in Figure 6 provide a good measure of success for prospective innovations that hope to succeed in the market. By using an agent network to represent a social group, this model can also successfully simulate the effect that a social group can have on a consumer’s rejection to an innovation.

Joachim et al. (2018) proposes a two-step procedure for companies to address the relevant barriers of innovation rejection. First, they propose a study that asses the influence of these barriers on intention to adopt of the targeted consumers. The model that has been proposed in this study can assist in this step of the procedure. By providing relevant insights on the significance of each barrier, this model can form the basis of the first step in the proposed two-step procedure. The second step in the procedure is to conduct a further study to reduce the identified barriers. The model that we have proposed can also be used to assess the effectiveness of the second step. By revealing the number of rejecters an innovation has the model can provide a base line for further improvements to the innovation to be measured against. Any effort made to reduce the identified barriers will have a quantifiable effect on the rejection of the innovation, this effect can be measured by using the model proposed in this study.

The study has successfully answered the research questions that were posed. Firstly, by providing a fair estimate of the rate of rejection, the model has successfully accomplished one of its goals. Secondly, the model also records the portion of the rejecters and adopters in response to a certain innovation. This is a vital statistic for managers that want to evaluate the performance of innovations with regards to their rejection ratio. Out of a target population of 1000 consumers the number of rejecters is a good indicator of the prospect of innovation rejection. Thirdly, the model records the reason for rejection of each consumer. It is not enough to know that a consumer has rejected the innovation, we must also be able to determine the reasons of rejection. Therefore, by providing the reasons of rejection to managers, we give them aspects of the innovation that they can improve upon to effectively reduce innovation resistance. By answering the posed research questions, the study has successfully produced a simulation model that managers and companies can use to effectively manage innovation resistance; thus, the study has managed to accomplish its research goals.

Further research can be done to improve the model. We recommend that any future research be aimed at introducing different types of consumers to the model. This will help grow the generalizability of the model and help it better represent a target market. Another possible extension would be to introduce multi-group analyses of SEM results so that the model can compare and contrast the results between groups of consumers and how the barriers vary with each group. Finally, improvements can be made to better represent innovations in the model. Currently, an innovation is only represented by nine functional attributes, this is an area which can improve the model. Utilizing more aspects of the innovation in the model will help the model better represent the innovation in the simulation.

REFERENCES

Castellion, G., & Markham, S. K. (2013). Perspective: New product failure rates: Influence of Argumentum ad Populum and Self-Interest. *Journal of Product Innovation Management, 30*(5), 976-979.

Davidson, R. S., & Walley, P. B. (1985). Computer fear and addiction. *Journal of Organizational Behavior Management, 6*(3-4), 37-52. doi:10.1300/J075v06n03_03

Eric A. Greenleaf, & Donald R. Lehmann. (1995). Reasons for substantial delay in consumer decision making. *Journal of Consumer Research, 22*(2), 186-199. doi:10.1086/209444

Garcia, R., & Jager, W. (2011). From the special issue editors: Agent-Based modeling of innovation diffusion. *Journal of Product Innovation Management, 28*(2), 148-151.

Gilbert, N. (2008). *Agent-based models* Sage.

Hansen, J. M., McDonald, R. E., & Mitchell, R. K. (2013). Competence resource specialization, causal ambiguity, and the creation and decay of competitiveness: The role of marketing strategy in new product performance and shareholder value. *Journal of the Academy of Marketing Science, 41*(3), 300-319.

Heidenrich, S., & Spieth, P. (2013). Why innovations fail: The case of passive and active innovation resistance. *International Journal of Innovation Management, 17*(5), 1-42. Retrieved from http://www.econis.eu/PPNSET?PPN=774662050

Henard, D. H., & Szymanski, D. M. (2001). Why some new products are more successful than others. *Journal of Marketing Research, 38*(3), 362-375.

Hui, C., Goldberg, M., Magdon-Ismail, M., & Wallace, W. A. (2010). Simulating the diffusion of information: An agent-based modeling approach. *International Journal of Agent Technologies and Systems (IJATS), 2*(3), 31-46.

Joachim, V., Spieth, P., & Heidenreich, S. (2018). Active innovation resistance: An empirical study on functional and psychological barriers to innovation adoption in different contexts. *Industrial Marketing Management, 71*, 95-107. doi:10.1016/j.indmarman.2017.12.011

Kleijnen, M., Lee, N., & Wetzels, M. (2009). An exploration of consumer resistance to innovation and its antecedents. *Journal of Economic Psychology, 30*(3), 344-357. doi:10.1016/j.joep.2009.02.004

Kuisma, T., Laukkanen, T., & Hiltunen, M. (2007). Mapping the reasons for resistance to internet banking: A means-end approach. *International Journal of Information Management, 27*(2), 75-85. doi:10.1016/j.ijinfomgt.2006.08.006

Laukkanen, T., Sinkkonen, S., & Laukkanen, P. (2009). Communication strategies to overcome functional and psychological resistance to internet banking. *International Journal of Information Management, 29*(2), 111-118.

Liao, S., Chou, C. Y., & Lin, T. (2015). Adverse behavioral and relational consequences of service innovation failure. *Journal of Business Research, 68*(4), 834-839.

Meyer, M., Lorscheid, I., & Troitzsch, K. (2009). The development of social simulation as reflected in the first ten years of JASSS: A citation and co-citation analysis. *Journal of Artificial Societies and Social Simulation, 12*(4) Retrieved from http://econpapers.repec.org/article/jasjasssj/2009-47-2.htm

Nabih, M. I., Bloem, S. G., & Poiesz, T. B. C. (1997). Conceptual issues in the study of innovation adoption behavior. *Advances in Consumer Research, 24*, 190. Retrieved from https://search.proquest.com/docview/1839785986

Ram, S., & Sheth, J. N. (1989). Consumer resistance to innovations: The marketing problem and its solutions. *Journal of Consumer Marketing, 6*(2), 5-14. doi:10.1108/EUM0000000002542

Repenning, N. (2002). A simulation-based approach to understanding the dynamics of innovation implementation. *Organization Science, 13*(2), 109-127.

Rogers, E. M. (2003). *Diffusion of innovations*. New York: Free Press.

Schubring, S., Lorscheid, I., Meyer, M., & Ringle, C. M. (2016). The PLS agent: Predictive modeling with PLS-SEM and agent-based simulation. *Journal of Business Research, 69*(10), 4604-4612.

Talke, K., & Heidenreich, S. (2014). How to overcome Pro-Change bias: Incorporating passive and active innovation resistance in innovation decision models. *Journal of Product Innovation Management, 31*(5), 894-907. doi:10.1111/jpim.12130

Trusov, M., Rand, W., & Joshi, Y. V. (2013). Improving prelaunch diffusion forecasts: Using synthetic networks as simulated priors. *Journal of Marketing Research, 50*(6), 675-690.

DECISION SCIENCES INSTITUTE
Managerial Decision Making: antecedent and consequent variables study

Larissa Alves Sincorá
Universidade Federal do Espírito Santo
Email: sincora.larissa@gmail.com

Lucas Poubel Timm do Carmo
Instituto Federal do Espírito Santo
Email: lucaspoubeladm@gmail.com

Marcos Paulo Valadares de Oliveira
Universidade Federal do Espírito Santo
Email: marcos.p.oliveira@ufes.br

Marcelo Moll Brandão
Universidade Federal do Espírito Santo
Email: mollmkt@gmail.com

Hélio Zanquetto Filho
Universidade Federal do Espírito Santo
Email: zanquetto@gmail.com

ABSTRACT

In order to examine the decisions rationally and intuitively processed, this paper reports the results of a research developed with 115 professionals working in managerial positions in Brazil. The data analysis used the structural equation modeling. The results revealed the extent to which the antecedent variables along with the decision process generate performance impacts.

KEYWORDS: Decision-making approaches, Situational antecedent, Personal antecedents, Information processing, Decision-making performance

INTRODUCTION

The rapid technological change that society has been experiencing in the last decades, known as the BigData era (Barbosa, Vicente, Ladeira, & Oliveira, 2017; Ghasemaghaei, Ebrahimi, & Hassanein, 2018; H. Wang, Xu, Fujita, & Liu, 2016), creates a problem of information processing, culminating in the increase of the load to be processed by the decision maker (Ensley, Pearce, & Hmieleski, 2006; McCormack & Trkman, 2014; Tushman, 1979). Weiss and Heide (1993) have shown that individuals respond to this increased information processing burden by intensifying cognitive effort in data search and analysis.

Besides the emerging context of high data and information production, world economies face greater dynamism and complexity (Luo, 2003), as well as environmental uncertainty (Duncan, 1972). In this scenario, decision-makers also tend to perceive more uncertainty at different levels and directions, thus requiring the use of decision strategies that are effective in dynamic and complex environments, affected by a multiplicity of organizational, situational, and personal variables (Carter, Kaufmann, & Wagner, 2017; Riedl, Kaufmann, Zimmermann, & Perols, 2013).

Given the large amount of data and information present in organizations and the advancement of technologies that support managerial activities, it is increasingly expected that today's managers make decisions based on rationality and analytical models (Appelbaum, Kogan, Vasarhelyi, & Yan, 2017; Vidgen, Shaw, & Grant, 2017).

It is believed that the use of formal logic results in the best available solution (Neumann, John Von Morgenstern, 1944). In fact, rational decision-making theory presupposes that the decision-maker possesses - and invariably uses - a logical strategy to produce valid inferences, on which an appropriate response option is made.

It is important to consider that a full set of tool-technology, as well as all information concerning a scenario or a problem situation, will not always be available to the decision maker. Some information available may be accurate, however, others may be incomplete or even confusing. In addition, the decision can be developed in contexts of multiple actors with organizational and environmental constraints that bring consequences that are often difficult to evaluate (Sincorá, Poubel, Souza, & Oliveira, 2018; Sincorá, Oliveira, Zanquetto-Filho, & Ladeira, 2018).

By considering such elements, which form a complex backdrop for the development of decision-making, it is difficult to articulate decision-making purely based on rationality and analytical models (Sayegh, Anthony, & Perrewe, 2004; Simon, 1987), as advocated by traditional Decison Making theorists. Thus, it is important to look at other elements that guide the decision-making process, besides the rational, for a more subjective and intuitive perspective. The relevance is to understand how they influence, complement and participate in managerial decision-making, and its results, which can bring relevant contributions to the advancement of both traditional and emerging approaches to Decision Making.

In addition to the discussion inherent to rational and intuitive decision-making processes, based on the literature review, it has been observed that it still embryonic the empirical research showing the effects and values generated by analytical capabilities to decisions (whether strategic, tactical or operational) and how these combined concepts are applied in practice (Bon & Broersen, 2017; Sharma, Mithas, & Kankanhalli, 2014; Shuradze & Wagner, 2016; Sincorá, Oliveira, Zanquetto-Filho, & Ladeira, 2018).

It was found that despite the existence of literature on time pressure in an experimental perspective (Gazdag, Haude, Hoegl, & Muethel, 2018; Lohse, Simon, & Konrad, 2018; Skerlavaj, Connelly, Cerne, & Dysvik, 2018; Szollos, 2009), the investigation of time pressure over decision-making in a quantitative survey-type approach is still relatively unexplored. The need for scales to measure this phenomenon is important to advance the understanding of how the decision-making process is configured and generates results in terms of performance in the presence of situational variables, specially those strongly present in our post-industrial society.

Another aspect is that there are few management studies bringing empirical evidences of the extent to which the individual's professional experience level may influence their decision-making process when intuitively configured (Dane & Pratt, 2007; Fisher, Chengalur-Smith, &

Ballou, 2003; Phillips, Fletcher, Marks, & Hine, 2016). However, there is also the question of how decision makers process information under uncertainty, which has received increasing interest in management literature, especially in behavioral operations. In addition, few studies investigate the information processing and decision-making strategies (rational and intuitive) that are used to minimize uncertainty in decisions (Kaufmann, Michel, & Carter, 2009).

Therefore, this article aims to fill these gaps through a quantitative approach, by analyzing what extent the manager's rational and intuitive load influence his decision results. Therefore, it proposes to answer the respective research problems: *To which extent do different situational and personal variables impact the use of rational and intuitive approaches to decision-making? And how do rational and intuitive approaches impact decision-making?*

DEVELOPMENT OF HYPOTHESES AND RESEARCH MODEL

The theoretical relationships that provide the basis for the proposition of the theoretical hypotheses of the research will be presented below. In addition, a brief conceptual argument is presented for the Total Decision-Making Cost (TDMC) construct, considered as a proxy of performance, representing an effort of this research to offer an alternative to understand the decision-making results.

Situational Antecedent

Tabatabaei (2002) highlights the important antecedent role of time pressure in the selection of decision strategies. Due to its relevance as a determinant of decision-making behavior (Szollos, 2009), time pressure is examined here as a situational antecedent of the decision-making process.

It is observed that, in certain contexts, the decision-making process may undergo the influence of the situational variable time pressure (Gazdag et al., 2018; Lohse et al., 2018; Skerlavaj et al., 2018). When the decision maker needs to articulate decisions under time pressure, which means being pressured to quickly make decisions, this situational characteristic can make rational/analytical decision strategies less attractive or even unfeasible (Ordonez & Benson III, 1997), influencing the selection of other decision methods. In particular, decision makers facing time pressure tend to rely more on intuitive decision strategies and less on analytical decision strategies (Vanharanta & Easton, 2010). Following this line of thinking Riedl et al. (2013) argue that the time pressure also restricts the search for appropriate alternatives and induces the individual to adopt more intuitive decision-making. Consequently, the time pressure is shown to be negatively related to the rational decision-making process. It is assumed that it will probably prevent the decision maker from collecting, processing and properly applying information in the delineation of its decision-making, indirectly and negatively impacting the results in terms of the total decision-making cost.

Still considering the emerging scenario marked by the high production and dissemination of data and information, and by the technological revolution, authors such as Dholakia et al. (1993) already pointed out at the end of the last century that time pressure in high technology environments - as experienced in today's society - would certainly lead people to adopt inadequate mechanisms of information collection and non-systematic and insufficient evaluation of alternatives. Following this line of thought, it is conjectured that time pressure not only restricts the search for appropriate alternatives and leads to intuitive decision strategies but increases the probability of the individual to generate choices based on biases and misperceptions (Kahneman, 2011). When the deadlines are tight, people are encouraged to speed up their deliberations since 'there isn’t much time', leading to a lower performance. In addition, when time becomes restricted and conservative, the decision-maker tends to "cut" his thinking, leading to a "cognitive closure" situation (Wright, 1974).

Therefore, it is worth considering that under high time pressure both individuals, the one that decides more rationally and other that decides more intuitively, will drastically reduce the amount of information processing, culminating in poor performance in both behavioral scenarios. According to this, the first two research proposals are raised:

H1a. Time pressure negatively impacts rational decision-making process performance;

H1b. Time pressure negatively impacts the intuitive decision-making process performance.

Personal Antecedents

In addition to situational antecedents, decisions are also influenced by personal variables (Ackerman & Thompson, 2017, Carter et al., 2017, Frederick, 2005, Jackson, Kleitman, Stankov, & Howie, 2017, Kirchler et al., 2017, Moritz , Hill, & L.Donohue, 2013, Narayanan & Moritz, 2015, Toplak, West, & Stanovich, 2014. Weinhardt, Hendijani, Harman, Steel, & Gonzalez, 2015). Previous research has argued that professional experience and analytical capabilities are important factors that affect decision-making behavior (Dane & Pratt, 2007; Sincorá, Oliveira, Zanquetto-Filho, & Ladeira, 2018). In this sense, it is pertinent to examine these factors as antecedents of both rational decision-making process and intuitive decision-making process.

It is recognized that professional or work experience interferes the decision strategy employed by the individual (Tsiros & Heilman, 2005). Beach and Mitchell (1978) have shown that past experience can help the decision maker regain specific decision strategies that have been successfully used previously and discard unsuccessful ones. In addition, Browne et al. (2007) argued that the level of experience determines whether a person uses a more rational or more intuitive decision-making approach. Therefore, studies suggest that experience is an important factor in trying to understand which decision strategy individuals’ use.

As far as the relationship between experience and rational decision-making is concerned, previous literature has shown that more experienced decision-makers consider more sources of information (Perkins & Rao, 1990), are more sensitive to information relevance, and process data in more complete form (Sanbonmatsu, Kardes, & Herr, 1992). This suggests that more experienced decision makers are more likely to use more sophisticated decision-making approaches, such as procedural rationality and decomposition (Riedl et al., 2013) - basic constituents of rational decision-making.

In this way, it is assumed that the decision-maker, by having significant professional experience, can complement his rational decision-making process by using intuitive judgments based on the recognition of situations-problems that have already been experienced and need to be resolved. Thus, the third hypothesis of the research is formulated:

H2a. Professional experience has a positive impact on rational decision-making process.

Consequently, it can be seen that the intuitive capacity and the expert judgment of the decision maker can be refined by the experience acquired by her over time in the exercise of his professional activity. According to Simon (1987), the intuition of the individual is formed by the storage of a large amount of knowledge in memory, which in turn, is obtained from formal trainings and experiences lived by the managers throughout their professional trajectory. Shanteau (1992), for example, when conducting a study analyzing the performance of intuition in different professions, found out, among other factors, that the amount of experience of the studied professionals was determinant for one group to perform well in relation to another that demonstrated low intuitive ability.

Therefore, the level of professional experience may act as an antecedent element, allowing decisions based on intuition to achieve less negative results, that is, minimizing possible negative influences coming from the decision-making strategy itself, allowing the use of a more intuitive and accurate ability to the detriment of biased impressions (Kahneman & Klein, 2009). Through the logical sequence, the fourth hypothesis of the research is presented:

H2b. Professional experience has a positive impact on the intuitive decision-making process.

In addition to professional experience, the analyst's level of analytical capabilities also impacts decision-making strategies. According to Sincorá et al. (2018), analytical capabilities can be subdivided into statistical capabilities, business capabilities, and information technology capabilities - perspectives related to human cognition - that are necessary for an individual to be able to take the advantage of analytics to support decision-making, thus enhancing the information processing.

The literature argues that analytical capabilities increase (i) the robustness of decision-making processes, since they are based on fact and data (Davenport & Harris, 2007; Sharma et al., 2014); (Bon & Broersen, 2017; Davenport & Harris, 2010; Y. Wang & Byrd, 2017), (ii) the amount of attention and effort that individuals invest in their decisions and (iii) the use of cognitively more demanding decision strategies (Delen & Demirkan, 2013, Gorman & Klimberg, 2014, Kahneman, 2011). However, empirical research that investigates the role of analytical capabilities in decision-making processes is still underdeveloped. Recent studies suggest that such capabilities can change strategies and entire decision-making philosophies, leading not only to increased cognitive effort and reflection, but also to the use of more analytical methods and tools in decision-making (Vale, Sincorá, & Milhomem, 2018).

Nonetheless, the use of analytical capabilities to transform data and information into decision-relevant knowledge makes decision-making tasks more robust and less uncertain. In addition, it facilitates the distinction between relevant and irrelevant information as well as the identification of attributes against which the available alternatives can be discriminated. Likewise, it is supposed that decision makers with an ability to understand the needs of the business, interpret analyzes carried out in large databases, and provide meaning to them in order to make decisions regarding problems and opportunities that emerge in the organization (Acito & Khatri, 2014, Delen & Demirkan, 2013, Wilder & Ozgur, 2015), can be more easily involved in more complex and structured decision-making such as procedural rationality (Riedl et al., 2013). Moreover, the ability to discriminate relevant data from irrelevant data is a key influencing factor in rational decision-making. Therefore, it is understood the appropriateness for the construction of the respective theoretical hypothesis:

H3a. Analytical capabilities positively impact rational decision-making process.

Consistently within this logic, but analyzing through the lens of intuitive decision-making processes, it is assumed that analytical capabilities also exert a positive influence on intuitive approaches. Such capabilities corroborate to make the decision maker more analytically oriented, allowing her to more easily identify the misperceptions generated by her cognition in information processing.

This supposition is justified by the assumptions of the Dual Process Theory (Evans, 2003; Kahneman, 2011) in recognizing that the individual, in facing a decision-making task, will activate 'System 1' (referring to the set of cognitive processes that are automatic, effortless, associative and fast) and will automatically generate intuitive proposals on how to solve the problem. 'System 2' (inherent in cognitive processes that are controlled, hard-working, deductive and slow) can be triggered to analyze intuitive proposals and decide whether to reject, approve or modify them. Final judgments are generally highly anchored in the initial impressions generated by System 1, which may in turn be 'contaminated' by heuristics and biases, culminating in unsatisfactory performance results.

Based on discussions about analytical capabilities, it is suspected that they are closely related to 'System 2', popularly known as 'reflexive reasoning' (Holyoak & Morrison, 2005; Kahneman, 2011). Such a relationship leads to the assumption that the more developed these capabilities are in the individual, the more analytically oriented their decisions will be. Thus, by accrediting it in intuitive decision-making, 'blocking' with more frequency and accuracy inconsistent impressions and judgments, commonly produced by 'System 1', thus increasing the probabilities of correctness and the overall quality of the decision. In this way, it is postulated:

H3b. Analytical capabilities positively impact the intuitive decision-making process.

Total Decision-Making Cost (TDMC): a performance proxy

The decision when analogously compared to a 'transaction' (Williamson, 1981) is subject to incurring costs before and after its occurrence. The way the decision-making process is constructed reflects on the results that are obtained. In this way, it is possible to measure the performance of the decision by mapping the costs incurred before and after the decision. These costs refer to tangible and intangible attributes that the decision maker can add to the decision-making scope.

Every decision has an inherent risk of something unexpected to happen (Knight, 1964), above all, it depends on the rationality limits of the decision maker (Simon, 1955), on the situational variables that influence the configuration of the decision process - , Goodie, Hall, & Wu, 2012) - as well as personal variables - such as professional experience and capabilities of the decision maker (Dane & Pratt, 2007; Sinanca, Oliveira, Zanquetto-Filho, & Ladeira, 2018), and of course, the probabilities of gain and loss with the choice of an alternative (Kahneman & Tversky, 1979). These factors, in turn, shape different behaviors in decision makers, who are likely to assume or minimize the harmful effects of risk and decision uncertainty, depending on the impact that such choices tend to produce on performance outcomes (Kahneman & Tversky, 1979).

Thus, individuals incur different types of "costs" to obtain positive decision outcomes, especially those related to information processing (Galbraith, 1974; McCormack & Trkman, 2014), since the knowledge acquired by the information processed contributes to reduce the uncertainty of decision making, increase the robustness of the probabilities and correct answers, identify needs for changes in resource allocation, timelines and criteria to be prioritized in decision making, thus favoring the results.

The theoretical concept of the decision making total cost is defined as the total cost of a decision considering the sum of the resources consumed prior to decision making (denominated costs ex-ante) and resources consumed after the decision (denominated costs ex-post). The main idea is that the more resources are spent before the decision, by reflecting, processing and reconsidering alternatives, the greater the chance of the decision being successful and contributing to the needs of individuals and companies. Whereas, the higher the ex-post cost incurred in a decision-making configuration, the greater the resource consumption to correct/adjust what went wrong, pointing to an unsatisfactory decision.

Rational Decision Making (RDM) and Total Decision-Making Cost (TDMC)

Commonly, people are highly vigilant and meticulous information processors when they need to justify their decisions to others (Wouters et al., 2009), especially when they hold management positions. They are also concerned with intensifying the search and overall acquisition of data and information (Doney & Armstrong, 1996; Siegel-Jacobs & Yates, 1996). In addition, they feel more responsible for the results produced by their choices, tending to use more analytical and to use complex decision-making strategies (Kaufmann et al., 2009), such as procedural rationality and decomposition in the execution of their tasks (Beach & Mitchell, 1978).

Considering that the rational decision-making process is that characteristically deliberative, structured and quantitative, supported by a wide range of statistical tools and expert system technology (Simon, 1987), it is assumed that this rational processing has a positive impact on the total decision-making cost, enabling the manager to generate satisfactory results. The positive relationship between rational decision-making processes and different performance dimensions has been found in several Decision Making studies (Carter, Kaufmann, & Michel, 2007; Miller, 2008; Mueller, Mone, & Barker, 2007; Phillips et al., Riedl et al., 2013; YI Wang, Highhouse, Lake, Petersen, & Rada, 2017).

This assumption is raised because it is precisely in the rational decision-making process that the manager is more likely to prioritize her decision, to consume more time and cognitive effort by collecting and analyzing information, managing complex tasks, determining a set of relevant criteria, and identifying priorities in the context of the decision (ex-ante costs) (Elbanna, 2006). These actions and activities corroborate that the result of the decision has greater feasibility to be satisfactory, reducing the need for adjustments and punctual corrections after decision making (ex-post costs). Thus, the theoretical hypothesis is postulated:

H4. Rational decision-making has a positive impact on the performance of the decision.

Intuitive Decision Making (IDM) and Total Decision-Making Cost (TDMC)

Although it is recognized in the literature a negative relationship between the intuitive processing and the decision performance result (Frederick, 2005; Kahneman & Frederick, 2002;

Lakeh & Ghaffarzadegan, 2015, 2016, Phillips et al. 2016; Weinhardt et al., 2015) there is no real decision-making option to make only rational and analytical decisions. This is because, due to increasing uncertainty and complexity that managers face in their work environments, it becomes almost impossible to rely solely on a rational decision-making process (Carter et al., 2017; Dane & Pratt, 2007), thus opening space for intuition.

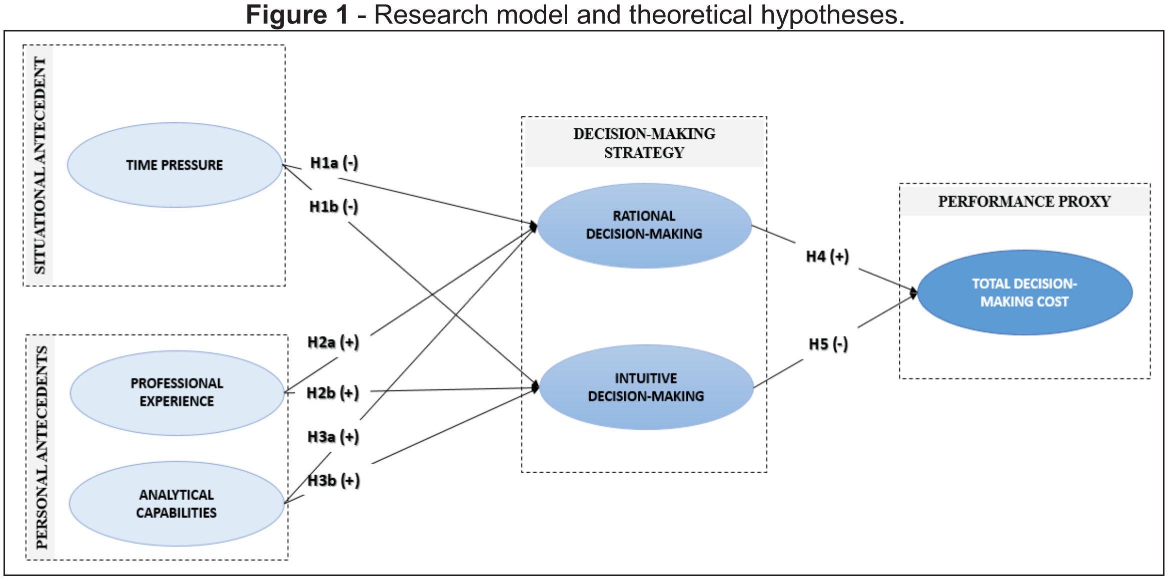
Taking into account that the intuitive decision-making process is characteristically qualitative and loosely structured (Simon, 1987), it is verified that the decision maker probably does not consume significant time for the construction of his decision, culminating in the use of little cognitive effort, low processing and analysis of information, inadequate evaluation of possible risks, insufficient identification of alternatives and development of decision steps (ex-ante costs). Undoubtedly, in this context, the probabilities of unsatisfactory results are high. Therefore, the following research hypothesis is formulated:

H5. Intuitive decision-making has a negative impact on the performance of the decision.

Research Model

The reflections obtained from the evaluation of the theoretical contents of rational decision-making (RDM), intuitive decision-making (IDM), analytical capabilities (AC), time pressure (TP), professional experience (PE) and total decision-making cost (TDMC), provided a fertile ground for the establishment of hypothetical relations amongst them. In this way, it is understood that the articulated literature confers conceptual justification for the proposition of the theoretical model (Figure 1).

Figure 1 - Research model and theoretical hypotheses.



Source: Prepared by the authors.

Figure 1 presents the nomological model of research, linking the rational decision-making (RDM) and the intuitive decision-making (IDM) to its antecedents and to the total decision-making cost (TDMC), considered as a consequent variable. The Appendix section of this manuscript presents in details the operationalization of each of the model's constructs.

METHODOLOGY

Scale Validadion and Construction of the Data Collection Instrument

The research questionnaire was elaborated based on a vast literature, which served as a theoretical basis for the formulation of 54 assertions - being 9 on the profile of the respondent/company and 45 on the constructs studied. In addition, the Likert scale was used from 1 (one) to 7 (seven) points, anchored at 1= Strongly Disagree up to 7 = Strongly Agree.

For the construction of the TP (time pressure), PE (professional experience) and TDMC (total decision-making cost) scales, a literature was mapped to identify a set of manifest variables - of higher incidence - that could measure the constructs studied, since no scales were found that were aligned with the research objectives. While the measurement of the RDM (rational decision-making), IDM (intuitive decision-making) and AC (analytical capabilities) constructs were inspired by previously developed scales, which have undergone occasional modifications due to linguistic adaptation and cultural validation. RDM, for example, was obtained from the work of Reidl et al. (2013), who investigated the procedural rationality of managers. IDM, was derived from the study by Carter et al. (2017), who proposed a unifying concept of intuition for the supply chain. Finally, the scales for AC emerged from the work of Sincorá et al. (2018) and validated with data from 288 companies in the industry, commerce and service sectors.

After the structuring of the questionnaire, the 30 assertions that had their scales constructed from the research effort (TP, PE and TDMC), received face validation (Mosier, 1947), in which 29 items obtained a correspondence of 80% for one same construct, based on the evaluation of 12 experts. After this step, all the items of the scale - 65 to the total - were evaluated by this same group of experts (teachers-doctors, academics and managers from different areas of the market) experienced in the conduction and application of surveys. Their validation contributed to the objectivity, clarity and coherence of the instrument, eliminating redundancies, ambiguities and content overlaps and allowing the common variance bias of the research instrument to be reduced.

The data used in the survey were collected from online questionnaires – through Google Docs technology - applied to managers affiliated with FINDES (Federation of Industries of the State of Espírito Santo), CRAES (Regional Council of State Administration of Espírito Santo), and those who had accounts on LinkedIn's social network (from all over Brazil). The data collection was performed between the last two months of 2017 and the second two months of 2018.

The respondents selected to participate in the research were those who held management positions in different areas (human resources, financial, accounting, production, logistics, marketing, sales, quality, research and development, supplies and others). In order to allow the use of the analysis technique of Structural Equation Modeling (SEM) in software Smart PLS-SEM 3.8.2 (Ringle, Wende, & Becker, 2014), which in turn is based on the Partial Least Squares (PLS) algorithm, it was identified the need to collect a minimum sample value of 60 respondents (Hair, Hult, Ringle, & Sarstedt, 2017). After performing a preliminary analysis to identify and treat possible problems with the data collected, the final sample consisted of 115 valid cases.

In order to evaluate the validity and reliability of the measurements, the PLS-SEM Smart software was started to perform tests for validation of the reflexive measurement models (composite reliability test, cronbach alpha, convergent validity, AVE and fornell-larcker test), formative measurement models (test of multicollinearity and significance and relevance), and structural model (test of multicollinearity, significance and relevance, determination of the coefficient of determination - R2 size of effect f2, predictive relevance Q2 and size of relative predictive relevance q2). All tests were performed according to the validation stages recommended by Hair et al. (2017).

RESULTS OF THE STRUCTURAL EQUATION MODELING

Validity and Reliability of the Measurement Model

When developing the validation tests of the reflective measurement models, from the PLS Algorithm method, it was observed that the values of Cronbach's Alpha and Composite Reliability attested that all the manifest indicators of the reflexive constructs were not measuring the same phenomenon and, thus, showing a high probability of being constituted in valid measures for the constructions. In relation to the Convergent Validity test, the values indicated that the reflective indicators have strong factorial loads, denoting that they really belong to their referred construct in the model, since they reveal to have much in common with each other. And as for the reflexive constructs, it turns out that they are responsible for explaining more than half the variation of their indicators. The exception lies only in the construct of ex-post costs which presented values of convergent validity and average variance extracted a little below the reference parameters, signaling the need for its measures to be reevaluated theoretically.

Finally, with regard to the Discriminant Validity test, the need to remove the 'EAC4' indicator belonging to the ex-ante costs construct was identified, in order to reduce the correlation between this construct and the Analytical Capabilities, and thus, minimize the likelihood that the indicators measure the same phenomenon in a similar way. After the withdrawal of the indicator that presented problems, all reflexive constructs were found to have a greater correlation with themselves than with any other construct of the model (Fornell-Larcker test).

Therefore, it is inferred from the results obtained in Table 1 that all existing relationships amongst reflexive constructs and their respective indicators are considered valid and reliable, within the quality criteria explained by Hair et al. (2017).

Table 1 - Values of the tests for validation of the reflexive measurement models.

	REFERENCE PARAMETERS (HAIR et al., 2017)				
	>0,60 e <0,90	>0,4	>0,708	>0,5	-
REFLECTIVE CONSTRUCTS	**Composite Reliability**	**Cronbach's Alpha**	**Convergent Validity**	**AVE**	**Fornell-Larcker**
Time Pressure	0,845	0,793	0,7275	0,529	0,728

Table 1 - Values of the tests for validation of the reflexive measurement models.

	REFERENCE PARAMETERS (HAIR et al., 2017)				
	>0,60 e <0,90	>0,4	>0,708	>0,5	-
REFLECTIVE CONSTRUCTS	**Composite Reliability**	**Cronbach's Alpha**	**Convergent Validity**	**AVE**	**Fornell-Larcker**
Procedural Rationality	0,811	0,690	0,7197	0,518	0,720
Decomposition	0,768	0,555	0,7266	0,528	0,727
Experience	0,829	0,690	0,7861	0,618	0,786
Emotion	0,794	0,662	0,7166	0,505	0,711
Automatic	0,823	0,676	0,7797	0,608	0,780
Ex-ante Cost	0,852	0,740	0,8111	0,658	0,811
Ex-post Cost	0,766	0,642	0,6855	0,464	0,681

Source: Elaborated by the authors based on the research data.

As for the formative measurement models, it has been identified through the Multicollinearity test (Table 2) that all the indicators of the constructs Statistical Capabilities, Business Capabilities, Information Technology Capabilities and Professional Experience are not convergent (equal) nor measured same part of the model. This indicates a good quality criterion to reveal the theoretical consistency of such indicators for the formation of the constructs to which they are related.

Therefore, from *Bootstrapping* it was possible to determine the significance and relevance of each set of formative indicators for its corresponding construct. The relevance test, then, was evidenced by the result of the statistic *t* that the indicators: SC1, ITC3, PE1, PE2, PE3, PE4 and EP5 presented values above the level of significance (0.05). However, such formative indicators were not removed from the model, since Hair et al. (2017) recommend that if the previous research and the theory consulted provide support to highlight the importance of the indicators to form the constructs to which they relate, then they must remain in the model. As for the other indicators, they all remained within the level of significance and statistical relevance.

Table 2 - Collinearity test for validation of formative models.

	REFERENCE PARAMETERS (Hair et al., 2017)			
	Magnitude: 0,90 or, at least, 0,80	Tol>0,2 and VIF<5	External Weights ≤ 1/√N and External Loads ≥ 0,5	Value p≤ 0,5
FORMATIVE CONSTRUCTS	**Convergent validity**	**Collinearity**	**Significance**	**Relevance**
Statistical Capabilities	there is no formative indicator of reflective character	The indicators did not present	Only the EP1 indicator did not meet the quality	Indicators with *p-value* >0,5: SC1, ITC3,

Table 2 - Collinearity test for validation of formative models.

	REFERENCE PARAMETERS (Hair et al., 2017)			
	Magnitude: 0,90 or, at least, 0,80	Tol>0,2 and VIF<5	External Weights ≤ 1/√N and External Loads ≥ 0,5	Value p≤ 0,5
FORMATIVE CONSTRUCTS	**Convergent validity**	**Collinearity**	**Significance**	**Relevance**
Business Capabilities	there is no formative indicator of reflective character	multicollinearity problems.	criteria. The others reached the reference parameter.	PE1, PE2, PE3, PE4 e PE5. The others reached the reference parameter.
Information Technology Capabilities	there is no formative indicator of reflective character			
Professional Experience	there is no formative indicator of reflective character			

Source: Elaborated by the authors based on the research data.

All the tests developed for the reflexive and formative models had the objective to identify if the quality of the model would be adequate and adjusted. Therefore, the results obtained, after adjustments, showed that the relations amongst indicators and constructs were considered valid. This demonstrates that the indicators selected for the study have robustness, reliability and validity to measure the first-order constructs to which they are related, demonstrating that they are good quality measures.

Structural Model Analysis

From the Multicollinearity test for the structural model it was not possible to verify any problem of high correlation between the constructs. This procedure is necessary to indicate that there is no conceptual shadowing between the investigated constructs, that is, that two or more latent variables are not equally measuring the same phenomenon.

The Significance and Relevance test of *t* distribution, with 114 degrees of freedom and 5% of significance level, was carried out by means of data extracted from *Bootstrapping* - Table 3 -, in which it was possible to identify that the hypotheses **H1a:** Time pressure negatively impacts the rational decision-making process; **H2a:** Professional experience has a positive impact on rational decision-making process; **H2b:** Professional experience positively impacts the intuitive decision-making process; and **H5:** The intuitive decision-making process has a negative impact on the performance of the decision, do not have statistical significance for the test. So, with 95% confidence, it is not possible to state that the antecedent variables of Time Pressure and Professional Experience have a relationship with rational and intuitive decision-making approaches for the sample researched. In addition, it was found that the Intuitive Decision Making (IDM) does not generate impacts on Total Decision-Making Cost (TDMC).

On the other hand, the other hypotheses **H1b:** The time pressure negatively impacts the intuitive decision-making process; **H3a:** Analytical capabilities positively impact rational decision-making process; **H3b:** Analytical capabilities positively impact the intuitive decision-making process; and **H4:** The rational decision-making process has a positive impact on the performance of the decision, were considered are significant to the theoretical model.

Table 3 - Significance test of the path coefficients for the structural model.

CONSTRUCTS	Original Sample	Media	Standard Deviation	Estat. *t*	*p-value**
H1a: Time Pressure -> Rational Decision Making	-0,027	-0,035	0,071	0,382	0,702
H1b: Time Pressure -> Intuitive Decision Making	0,343	0,342	0,118	2,913	**0,004**
H2a: Professional Experience -> Rational Decision Making	0,096	0,013	0,129	0,742	0,458
H2b: Professional Experience -> Intuitive Decision Making	0,099	0,058	0,146	0,680	0,497
H3a: Analytical Capabilities -> Rational Decision Making	0,728	0,736	0,057	12,870	**0,000**
H3b: Analytical Capabilities -> Intuitive Decision Making	0,395	0,386	0,171	2,312	**0,021**
H4: Rational Decision Making -> Total Decision-Making Cost	0,657	0,655	0,071	9,264	**0,000**
H5: Intuitive Decision Making -> Total Decision-Making Cost	0,131	0,135	0,100	1,309	0,190

* Considering as significant a p-value <0.05, when submitted to the *t* test with the *Bootstrapping* technique.
Source: Elaborated by the authors based on the research data.

One of the main questions of this research was to know to what extent the antecedent variables selected for the study impact on the rational and intuitive decision approaches. In this way, when analyzing the value of the Variance Determination Coefficient (R^2) for the Rational Decision Making ($R^2 = 0.575$), it can be stated that 57.5% of the variation occurring in this endogenous construct can be explained by the variation in behavior of the antecedent variables of the structural model (Time Pressure, Professional Experience and Capabilities Analytical) that influence it. On the other hand, it is worth noting that about 32.5% of the variation that happens in the RDM is not explained by the model. Considering that this is a value of moderate magnitude, it opens the possibility of a conceptual and empirical reflection on the need to identify other variables that are not contained in the model and that would help to clarify the behavior of the rational approach within the decision-making process. This gap should be incorporated into the model in future research.

Nevertheless, it was verified that the same antecedent variables are responsible for explaining the variation that happens in the behavior of the Intuitive Decision Making in approximately 27% (R^2). This reveals to be a 'weak' impact, since most of the variation (73%) cannot be justified by the model. Thus, it is reflected that there are other variables that are not articulated in the scope of work that could increase the explanatory power of the variation that happens in essentially intuitive decision-making. New investigations are then needed to uncover variables that contribute to a more holistic understanding of this decision strategy.

Also, concerning to the analysis of the value of R^2, it was verified that the endogenous Total Decision-Making Cost construct could have its variation explained in the order of 50% by means of the impact that it receives from the constructs RDM and IDM.

Now, in order to evaluate the size of the change in R^2 value, when a given exogenous construct is excluded from the structural model, it has been identified from the size of the f^2 - which evaluates how much each construct is 'useful' for the general adjustment of the model - that only the exogenous constructs Analytical Capabilities and Time Pressure, respectively, have large (1,17) and medium effect (0,16) in the endogenous RDM construct. With respect to the IDM endogenous construct, it is noted that only the Analytical Capabilities variable exerts an average effect (0.20) on the change in its coefficient of variance. Finally, when analyzing the TDMC, it is perceived that the RDM construct exerts great effect (0.79) and the IDM construct exerts a small effect (0,03) on its behavior. This assumes that all the constructs responsible for generating large and medium effect sizes functions as an important support to explain the behavior of the variables which are theoretically related.

Finally, in investigating the predictive relevance (Q^2) of the path model, referring to endogenous constructs of a reflexive nature (RDM, IDM and TDMC), we used *Blindfolding.* According to Hair et al. (2017) for predictive relevance (1-SSE / SSO) the values resulting from such a procedure must be > 0, and as shown in Table 4, it can be observed that the values of the tests for determination of Q^2 are different and larger that "0", thus demonstrating that the endogenous constructs have predictive relevance.

Table 4 - Values of the tests to determine the predictive relevance (Q2).

TOTAL	SSO	SSE	Q^2 (=1-SSE/SSO)
RDM	230,000	141,510	0,385
PRI	345,000	313,168	0,092
TDMC	805,000	696,381	0,135

Source: Elaborated by the authors based on the research data.

Within this same logic, from another dimension of analysis obtained through the *Blindfolding* procedure, it was possible to determine the size of the effect of the relative predictive relevance (q^2) of each exogenous variable in relation to its corresponding endogenous variable.

The results indicated that the omission of the Analytical Capabilities construct made the structural model practically incapable of predicting the variation that occurs in the RDM. Subsequently, the elimination of the RDM construct represented a medium effect of changing the R^2 value of the endogenous TDMC construct to which it relates. While the other constructs showed almost no change.

Thus, from the evaluation of the relative predictive relevance (q^2) and the size of the effect (f^2), it can be deduced that the analytical capability and the rational decision-making strategy are the variables responsible for verifying and contributing to the general adjustment of the structural model.

DISCUSSION AND MANAGERIAL IMPLICATIONS

The relationship between time pressure and rational decision-making was not shown to be significant (H1a). Research suggests that the choice of a decision maker under time pressure depends on the amount of pressure it receives (Ordonez & Benson III, 1997). More specifically, the low time pressure does not cause changes in decision strategies (Payne, Bettman, & Johnson, 1988), while the moderate time pressure leads to a more focused collection of relevant information, which is a component of rational decision-making, based on procedural rationality and decomposition of decision making (Riedl et al., 2013). Severe time pressure causes subjects to shift from more rational strategies to more intuitive strategies (Dane & Pratt, 2007). In view of the findings of previous studies, it is suspected that the interviewed managers may in their decision-making reality, not generally experience decision-making under moderate-time pressure, which would encourage the adoption of more rational and analytical strategies in the choice of alternatives, and explain the relationship between them. Thus, it is assumed that respondents are likely to 'suffer' relatively high pressure and, in these conditions, culminate in adopting other decision-making strategies.

On the other hand, a significant relationship between time pressure and intuitive decision-making (H1b) was identified. One explanation may be that survey respondents often deal with the physiological, cognitive, and emotional "stress" caused by the high time pressure in their work context, resulting in stress (Selye, 1983). The conceptual links with stress emphasize that time pressure is always an interaction between the person and the environment, and that variations of individual contexts always need to be taken into account (Szollos, 2009). Thus, in practical terms, it is suspected that the urgency to finish a given task or achieve a certain goal, culminates in decisions taken intuitively, above all, based on experience, emotion and automatic processing, generating many consequences in the processes of judgment of individuals. Thus, time pressure induces changes in the affective states of decision makers and their cognitive strategies (Ben Zur & Breznitz, 1981; Maule, Hockey, & Bdzola, 2000), imposing a series of limitations on them and, therefore, results in unsatisfactory performance.

According to H2a and H2b, it was not possible to find, respectively, a significant and positive relationship between professional experience and both rational and intuitive decision-making processes. The literature states that greater professional experience is likely to allow individuals to make more effective decisions with less cognitive effort, based on the recognition of situations and problems that have been experienced in the past, as a basis for approving choices in the present (Carter et al., 2017, Riedl et al., 2013, Simon, 1987). Under these conditions, the individual turns out to base her decision on tacit knowledge and expert judgment, characteristics of intuitive approaches. The adoption of more rational approaches, receives a divergent understanding in the scientific community. There are authors who believe that more experienced decision-makers tend to rely more on the principles of rationality, yet others understand that individuals with low to moderate work experience tend to adopt more rational decision-making strategies (Riedl et al., 2013). However, it was found that, although the managers who participated in the study had an average of 19 years of professional experience and 10 years of management occupation, this antecedent variable did not demonstrate any relationship with the principles of rationality or intuition in the decision-making process, which requires future research to understand the motivations.

Regarding the hypotheses H3a and H3b, it is understood that the benefit of using analytical capabilities in both rational and intuitive decision-making processes is significantly positive. It is argued that decision-makers having a certain level of analytical *expertise*, can exploit this potential as an additional benefit in their decision-making strategies, and can be indirectly rewarded with lower decision-making costs, which in turn act as indicators of positive performance outcomes. An explanation for the positive influence of analytical capabilities on decision-making processes may be that they are closely related to System 2, that is, to mental cognitive processes known to act in a controlled, laborious, deductive, and slow (serial) in situations of choice (Kahneman, 2011). Such a relationship leads to the assumption that the more robust such capacities are in the individual, the more developed will be his "System 2" and, therefore, the more analytically oriented his decisions will be. The decision maker will have the ability to transform data and information into knowledge useful and applied to the needs of rational decision-making; and on the other hand, she will have the ability, in intuitive decision-making, to 'block' more often incoherent impressions and judgments, commonly produced by its 'System 1', increasing the probabilities of correctness and improvement in the overall quality of the decision.

Previously, it has been argued that decision makers who use rational decision-making strategies, mainly based on procedural rationality, concentrate their search for relevant information (Browne et al., 2007) and thus gain a better understanding of their decision-making process (Ketchen , Snow, & Street, 2004), reducing decision uncertainty. By targeting their cognitive resources to more relevant information and having a better understanding of the decision process, decision makers are able to evaluate and project future performance more accurately and assertively. In this sense, it is possible to find explanations that justify the confirmation of the theoretical hypothesis H4, that rational decision processes generate positive impacts on performance. On the other hand, we identified the absence of statistical significance for the negative relationship between the intuitive decision-making process and the performance dimension, measured from the total decision-making cost. This may have occurred because of the managers who participated in the study, since they recognize that their decisions are predominantly taken analytically (from the constant use of results from data analysis) and not intuitively (based in their experience, tacit knowledge, instinct and emotions).

A curiosity observed in the study is that despite the average age of those who participated in the study being 42 years old, with μ=19 years of professional experience and μ=10 occupations in managerial positions, the respondents stated that they were more confident in strategies decisions that privilege the analysis of data and information, based on analytical capabilities, than involved in more intuitive decision strategies, which in turn privilege the professional experience acquired over the years. This counterintuitive result can be clarified by the combination of the current context in which organizations are overwhelmed with data and information produced daily by their ongoing operations. Furthermore, there is a growing popularization on analysis software and decision support systems that facilitate access to interpretation of information, motivating people to engage in more rational decision-making processes (Mortenson, Doherty, & Robinson, 2015).

Another explanation may lie in the fact that these managers simultaneously combine elements of rational and intuitive decision-making processes. The disciplines of psychology and management suggest that intuition can complement rationality as an effective approach to decision making (Carter et al., 2017; Dane & Pratt, 2007). Researchers who advocate a dual-process approach - which differentiates a rational system from an intuitive system - assume that these two systems work together in decision-making. Nonetheless, researchers of the subject present a variety of recommendations on how to use intuitions in combination with more rational decision-making. Simon (1987), for example, states that effective managers do not have the luxury of choosing between analysis and intuition - actual experience involves the use of both types of decision making. Hodgkinson and Sadler-Smith (2003) argue that the ability to switch between "mental habits" and "active thinking" is the ultimate skill of today's organizations.

Obviously, such clarifications are based on assumptions that merit more consistent exploration in order to identify characteristics and mechanisms that explain how decisions are being modeled in the current context.

CONCLUSIONS AND LIMITATIONS

This research provided several important *insights* to understand decision behavior when processed rationally and intuitively. By integrating the literature of the antecedent variables (TP, PE and CA) with the behavioral decision-making research, this study contributed to the *Decision Making* theory (i) developing a relatively comprehensive model of decision-making approaches; (ii) analyzing empirically what situational and personal characteristics influence the use of rational and intuitive decision-making processes; (iii) revealing the effectiveness of analytical capabilities as an additional resource for improved decision-making; and (iv) examining the idiosyncrasies of hypothetical model relationships from a sample of professionals working in managerial positions for at least 10 years. Moreover, research results point out to the need to deeply investigate the role of process rationality and decomposition as robust strategies for reducing uncertainty, and indirectly improving overall decision performance;

The empirical validation of the antecedent variables proved to be pertinent, since it provided a more complete understanding of the situational and personal characteristics that drive the rational and intuitive approaches of decision. In addition, revealing the extent to which rational and intuitive decision-making processes impact the performance. Although it lacks theoretical and operational adjustments, this study shows an effort in offering an alternative to the understanding of the behavior and results coming from management decisions. Therefore, these findings contribute to the important literature on antecedents and decision-making strategies that impact decision performance (Krause, Scannell, & Calantone, 2007; Modi & Mabert, 2007).

In summary, the purpose of the research served to answer the central problems of this study on the *extent to which different situational and personal variables impact the use of rational and intuitive decision-making approaches? And how do rational and intuitive approaches impact decision making?* The response obtained, then, was that PT negatively impacts the IDM, whereas the AC's positively influence the RDM and the IDM. Another relevant finding of the research lies in the RDM acting as a precedent of the TDMC generating positive impacts, as evidenced empirically. Therefore, half of the theoretical propositions of the work were confirmed (H1b, H3a, H3b and H4), and the other part was rejected (H1a, H2a, H2b and H5).

As a limitation of the study, among other points, it is important to emphasize that the research did not deal with a specific decision-making, but rather with a general evaluation of the decisions that are performed by the managers. Thus, it is understood that the use of unique respondents, and perceptions about general and non-specific decisions, probably added 'noise' to the study. However, it is believed that the methodological design of a single informant and the absence of a specific decision of the organizational context, do not violate the results obtained. In this way, future research is encouraged to collect responses from several respondents from the same company, who participate in specific decisions, such as in the area of operations management that involve supplier selection; deciding what, when and how much to buy; closing contracts; allocation of industrial plants; development of new products; production scheduling; project management; outsourcing; *etc.*, in order to validate the findings of this article.

Another possible extension of the research is to use longitudinal data, aiming to evidence stronger and more significant causal relationships. Future investigations should also consider participants from other countries and map moderating and mediating factors that provide more definite explanations as to why the theoretical relationships among variables exist. In addition, it is envisaged to discuss the interrelationship between rational and intuitive decision-making in normal working contexts and under the influence of situational variables - such as risks and uncertainties, from an experimental design approach.

Finally, it is important to highlight that articulating the analytical capabilities in the managerial decision-making process represents significant contours for the field. Considering that a few years ago the topic was effectively discussed within the organizational studies and the science of management, taking root and becoming rooted as a possibility of generation of teaching and research, since the publications are increasingly growing and popularizing, contributing for the evolution of the analytics movement.

APPENDIX

Figure 2 – Research constructs and indicators.

FORMATIVE CONSTRUCTS: SECOND-ORDER	FORMATIVE CONSTRUCTS: FIRST-ORDER	ITEM	INDICATORS
INTUITIVE DECISION-MAKING	**Experience-Based Processing**	EX1	Usually when I make decisions, I try to relate to similar situations in the past to decide what to do
		EX2	Usually when I make decisions, I look in other contexts for a reference to support my decision
		EX3	Usually when I make decisions, I rely on lessons learned by me considering my experience
	Emotional Processing	EM1	I always trust my feeling (feeling)
		EM2	When I'm not completely sure how to decide, I follow my feeling
		EM3	When I come across a problem for the first time, then I decide based on my feeling
		EM4	When my feeling says "something is wrong," then I reevaluate my decision
	Automatic Processing	AUT1	Usually when I make decisions, I know almost instantly what is the best course of action to follow
		AUT2	Usually when I make decisions, I know the right decision for the needs of the company without having to think too much
		AUT3	Usually when I make decisions, I always know exactly what to do
RATIONAL DECISION-MAKING	**Procedural Rationality**	PR1	Usually when I make decisions, I am rational
		PR2	Usually when I make decisions, I try to base my decisions on information
		PR3	Usually when making decisions, I use quantitative analyzes to identify what to do
		PR4	Usually when I make decisions, I analyze the relevant information extensively before reaching a conclusion

Figure 2 – Research constructs and indicators.

FORMATIVE CONSTRUCTS: SECOND-ORDER	FORMATIVE CONSTRUCTS: FIRST-ORDER	ITEM	INDICATORS
	Decomposing	DE1	Generally when I make decisions, I manage complex tasks, dividing them into subtasks, and dealing with each one at a time
		DE2	Usually when making decisions, we determine in advance a set of relevant criteria
		DE3	Usually when I make decisions, I identify what is priority in the context of the problem before requesting additional information that I need
ANALYTICAL CAPABILITIES	**Statistical Capabilities**	SC1	Data always inspires me with questions that I had never thought of before.
		SC2	I always use data to answer questions like "what happened and / or what is happening?"
		SC3	I always use data to answer questions like "what will happen and/or why will this happen?"
		SC4	I always use data to identify the best course of action to act on a particular situation.
	Business Capabilities	BC1	I know my work processes in an intensive way.
		BC2	I use logical reasoning when dealing with situations inherent in my context of work.
		BC3	I know the needs of the business in which I am inserted.
		BC4	I interpret the data to guide them in support of decision-making.
	Information Technology Capabilities	ITC1	I have the ability to organize data from different sources and formats in a logical way.
		ITC2	Work effectively with data (databases, reports, documents, sensors, maps, *click-streams,* etc.).
		ITC3	I have the ability to work with information technologies (computers, machines, systems, *software, hardware,* applications, computer modeling, etc.).
TOTAL DECISION-MAKING COST	**Ex-Ante Costs**	EAC1	Before making a decision, I usually collect and process information
		EAC2	Before making a decision, I usually review the information available carefully.
		EAC3	Before making a decision, I usually evaluate the possible consequences of the alternatives I have.
		EAC4	I usually spend a lot of time to make a decision.
	Ex-Post Costs	EPC1	After making a decision, I usually find myself rethinking if it was actually the most satisfactory.
		EPC2	After making a decision, I usually interrupt my activities to monitor the consequences of it.
		EPC3	After making a decision, I usually need to make corrections to what went wrong.
		EPC4	After making a decision, I often regret any thing.
	Time Pressure	TP1	The pressure of time causes me to decide as a matter of urgency.

Figure 2 – Research constructs and indicators.

FORMATIVE CONSTRUCTS: SECOND-ORDER	FORMATIVE CONSTRUCTS: FIRST-ORDER	ITEM	INDICATORS
		TP2	The pressure of time causes me to make decisions quickly, almost automatically.
		TP3	I usually have little time to decide.
		TP4	I usually have insufficient time to make a safe decision.
		TP5	The pressure to make an immediate decision is always very high.
	Professional Experience	PE1	What is your age?
		PE2	What is your time of professional experience?
		PE3	How many companies have you worked for?
		PE4	How many different departments / departments have you worked for in the companies you have worked with?
		PE5	How long have you held the management position?
		PE6	How many years of formal training do you have in the management area?

Source: Elaborated by the authors based on the research data.

REFERENCES

Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next? *Business Horizons*, *57*(5), 565–570. https://doi.org/10.1016/j.bushor.2014.06.001

Ackerman, R., & Thompson, V. A. (2017). Meta-Reasoning: Monitoring and Control of Thinking and Reasoning. *Trends in Cognitive Sciences*, *21*(8), 607–617.

Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, *25*(March), 29–44. https://doi.org/10.1016/j.accinf.2017.03.003

Barbosa, M. W., Vicente, A. de la C., Ladeira, M. B., & Oliveira, M. P. V. de. (2017). Managing supply chain resources with Big Data Analytics: a systematic review. *International Journal of Logistics Research and Applications*, *21*(3), 177–200.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review*, *3*(3), 439–449.

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, *47*, 89–104.

Bon, D., & Broersen, T. (2017). *Analytics in a Decision Service Context: Exploring analytical values for enhancing automated decision performance*. Lund University - School of Economics and Management.

Browne, G. J., Pitts, M. G., & Wetherbe, J. C. (2007). Cognitive stopping rules for terminating information search in online tasks. *MIS Quarterly*, *31*(1), 89–104.

Carter, C. R., Kaufmann, L., & Michel, A. (2007). Behavioral supply management: a taxonomy of judgment and decision ‐ making biases. *International Journal of Physical Distribution & Logistics Management*, *37*(8), 631–669. https://doi.org/10.1108/09600030710825694

Carter, C. R., Kaufmann, L., & Wagner, C. M. (2017). Reconceptualizing Intuition in Supply Chain Management. *Journal of Business Logistics*, *38*(2), 80–95.

Dane, E., & Pratt, M. G. (2007). Exploring intuition and its role in managerial decision making. *Academy of Management Review*, *32*(1), 33–54.

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning* (1° ed). Boston, MA, EUA: Harvard Business Review Press.

Davenport, T. H., & Harris, J. G. (2010). Leading the way towards better business insights. *Strategic HR Review*, *9*(4), 28–33. https://doi.org/10.1108/14754391011050388

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359–363. https://doi.org/10.1016/j.dss.2012.05.044

Dholakia, R. R., Johnson, J. L., Della Bitta, A. J., & Dholakia, N. (1993). Decision-making time in organizational buying behavior: an investigation of its antecedents. *Journal of the Academy of Marketing Science*, *21*(4), 281–292.

Doney, P. M., & Armstrong, G. M. (1996). Effects of accountability on symbolic information search and information analysis by organizational buyers. *Journal of the Academy of Marketing Science*, *24*(1), 57–65.

Duncan, R. B. (1972). Characteristics of organizational environments and perceived environmental uncertainty. *Administrative Science Quarterly*, *17*(3), 313–327.

Elbanna, S. (2006). Strategic Decision-Making: process perspectives. *International Journal of Management Reviews*, *8*(1), 1–20. https://doi.org/10.1111/j.1468-2370.2006.00118.x

Ensley, M. D., Pearce, C. L., & Hmieleski, K. M. (2006). The moderating effect of environmental dynamism on the relationship between entrepreneur leadership behavior and new venture performance. *Journal of Business Venturing*, *21*(2), 243–263.

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Fisher, C. W., Chengalur-Smith, I., & Ballou, D. P. (2003). The Impact of Experience and Time on the Use of Data Quality Information in Decision Making. *Information Systems Research*, *14*(2), 170–188. https://doi.org/10.1109/69.824597

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Galbraith, J. R. (1974). Organization Design: An Information Processing View. *Interfaces*, *4*(3), 28–36. https://doi.org/10.1287/inte.4.3.28

Gazdag, B. A., Haude, M., Hoegl, M., & Muethel, M. (2018). I Do Not Want to Trust You, but I Do: on the relationship between trust intent, trusting behavior, and time pressure. *Journal of Business and Psychology*. https://doi.org/10.1007/s10869-018-9597-y

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *Journal of Strategic Information Systems*, *27*, 101–113. https://doi.org/10.1016/j.jsis.2017.10.001

Gorman, M. F., & Klimberg, R. K. (2014). Benchmarking Academic Programs in Business Analytics. *Interfaces*, *44*(3), 329–341.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. (T. Oaks, Org.) (2ª). California: SAGE Publications, Inc.

Hodgkinson, G. P., & Sadler‐Smith, E. (2003). Complex or unitary? A critique and empirical reassessment of the Allinson's Cognitive Style Index. *Journal of Occupational and Organizational Psychology*, *76*(2), 243–268.

Holyoak, K. J., & Morrison, R. G. (2005). *The Cambridge Handbook of Thinking and Reasoning*. (K. J. Holyoak & R. G. Morrison, Orgs.) (1º ed). Cambridge, UK: Cambridge University Press.

Jackson, S. A., Kleitman, S., Stankov, L., & Howie, P. (2017). Individual Differences in Decision Making Depend on Cognitive Abilities, Monitoring and Control. *Journal of Behavioral Decision Making*, *30*(2), 209–223. https://doi.org/10.1002/bdm.1939

Kahneman, D. (2011). *Rápido e Devagar: duas formas de pensar*. (C. de A. Leite, Org.) (1º ed). Rio de Janeiro: Objetiva.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Orgs.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (p. 49–81). New York: Cambridge University Press.

Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise. *American Psychological Association*, *64*(6), 515–526. https://doi.org/10.1037/a0016755

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(March), 263–291. Recuperado de http://www.econometricsociety.org.

Kaufmann, L., Michel, A., & Carter, C. R. (2009). Debiasing strategies in supply management decision-making. *Journal of Business Logistics*, *30*(1), 85–106.

Ketchen, D. J., Snow, C. C., & Street, V. L. (2004). Improving firm performance by matching strategic decision-making processes to competitive dynamics. *Academy of Management Executive*, *18*(4), 29–43. https://doi.org/10.5465/ame.2004.15268671

Kirchler, M., Andersson, D., Bonn, C., Johannesson, M., Sørensen, E., Stefan, M., … Västfjäll, D. (2017). The effect of fast and slow decisions on risk taking. *Journal of Risk and Uncertainty*, *54*, 37–59. https://doi.org/10.1007/s11166-017-9252-4

Knight, F. H. (1964). *Risk, Uncertainty and Profit* (1º ed). New York: Sentry Press.

Krause, D. R., Scannell, T. V., & Calantone, R. J. (2007). A structural analysis of the effectiveness of buying firms' strategies to improve supplier performance. *Decision Sciences*, *31*(1), 33–55. https://doi.org/10.1111/j.1540-5915.2000.tb00923.x

Lakeh, A. B., & Ghaffarzadegan, N. (2015). Does analytical thinking improve understanding of accumulation? *System Dynamics Review*, *31*(1–2), 46–65. https://doi.org/10.1002/sdr.1528

Lakeh, A. B., & Ghaffarzadegan, N. (2016). The dual-process theory and understanding of stocks and flows. *System Dynamics Review*, *32*(3–4), 309–331.

Lohse, T., Simon, S. A., & Konrad, K. A. (2018). Deception under time pressure: conscious decision or a problem of awareness? *Journal of Economic Behavior and Organization*, *146*, 31–42. https://doi.org/10.1016/j.jebo.2017.11.026

Luo, Y. (2003). Market-seeking MNEs in an emerging market: how parent-subsidiary links shape overseas success. *Journal of International Business Studies*, *34*(3), 290–309.

Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy. *Acta Psychologica*. https://doi.org/10.1016/S0001-6918(00)00033-0

McCormack, K., & Trkman, P. (2014). The influence of information processing needs on the continuous use of business intelligence. *Information Research*, *19*(3), 135–150.

Miller, C. C. (2008). Decisional comprehensiveness and firm performance: Towards a more complete understanding. *Journal of Behavioral Decision Making*, *21*(5), 598–620.

Modi, S. B., & Mabert, V. A. (2007). Supplier development: Improving supplier performance through knowledge transfer. *Journal of Operations Management*, *25*(1), 42–64.

Moritz, B. B., Hill, A. V., & L.Donohue, K. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management*, *31*(1), 72–85. https://doi.org/https://doi.org/10.1016/j.jom.2012.11.006

Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research from Taylorism to Terabytes: A research agenda for the analytics age. *European Journal of Operational Research*, *241*(3), 583–595. https://doi.org/10.1016/j.ejor.2014.08.029

Mosier, C. I. (1947). A critical examination of the concepts of face validity. *Educational and Psychological Measurement*, *7*(2), 191–205.

Mueller, G. C., Mone, M. A., & Barker, V. L. (2007). Formal Strategic Analyses and Organizational Performance: Decomposing the Rational Model. *Organization Studies*, *28*(6), 853–883. https://doi.org/10.1177/0170840607075262

Narayanan, A., & Moritz, B. B. (2015). Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study. *Production and Operations Management*, *24*(8), 1216–1234. https://doi.org/10.1111/poms.12343

Neumann, John Von Morgenstern, O. (1944). *Theory of games and economic behavior*. (O. Morgenstern & J. von Neumann, Orgs.) (1º ed). Princeton University Press.

Ordonez, L., & Benson III, L. (1997). Decisions under time pressure: how time constraint affects risky decision making. *Organizational Behavior and Human Decision Processes*, *71*(2), 121–140.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 534–552.

Perkins, W. S., & Rao, R. C. (1990). The role of experience in information use and decision. *Journal of Marketing Research*, *27*(1), 1–10.

Phillips, W. J. T. styles and decision making: A. meta-analysis, Fletcher, J. M., Marks, A. D. G., & Hine, D. W. (2016). Thinking Styles and Decision Making: A Meta-Analysis. *Psychological Bulletin*, *142*(3), 260–290. https://doi.org/10.1037/bul0000027

Riedl, D. F., Kaufmann, L., Zimmermann, C., & Perols, J. L. (2013). Reducing uncertainty in supplier selection decisions: Antecedents and outcomes of procedural rationality. *Journal of Operations Management*, *31*(1–2), 24–36. https://doi.org/10.1016/j.jom.2012.10.003

Ringle, C. M., Wende, S., & Becker, J.-M. (2014). SmartPLS 3.0. Hamburg, Germany: SmartPLS.

Sanbonmatsu, D. M., Kardes, F. R., & Herr, P. M. (1992). The role of prior knowledge and missing information in multiattribute evaluation. *Organizational Behavior and Human Decision Processes*, *51*(1), 76–91.

Sayegh, L., Anthony, W. P., & Perrewe, P. L. (2004). Managerial decision-making under crisis: The role of emotion in an intuitive decision process. *Human Resource Management Review*, *14*, 179–199. https://doi.org/10.1016/j.hrmr.2004.05.002

Selye, H. (1983). The Stress Concept: Past, Present and Future'. In *Stress Research: Issues for the Eighties* (p. 1–20).

Shanteau, J. (1992). How much information does an expert use? Is It relevant? *Acta Psychologica*, *81*(1), 75–86.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433–441.

Shuradze, G., & Wagner, H. T. (2016). Towards a conceptualization of data analytics capabilities. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (Vol. 2016–March). https://doi.org/10.1109/HICSS.2016.626

Siegel-Jacobs, K., & Yates, J. F. (1996). Effects of procedural and outcome accountability on judgment quality. *Organizational Behavior and Human Decision Processes*, *65*(1), 1–17.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. Recuperado de http://www.jstor.org/stable/1884852

Simon, H. A. (1987). Making Management Decisions: the Role of Intuition and Emotion. *Academy of Management Executive*, *1*(1), 57–64.

Sincorá, L. A., Poubel, L., Souza, A. I. de, & Oliveira, M. P. V. de. (2018). The Decision Making Process of Productive Capacity in Port Environment. In *49th Annual Meeting of the Decision Sciences Institute* (p. 1–15). Chicago, IL (USA): Decision Science Institute.

Sincorá, L. A., Oliveira, M. P. V. de, Zanquetto-Filho, H., Carneiro, T. C. J., & Ladeira, M. B. (2018). The Role of Business Analytics on Organizational Resilience. In *49th Annual Meeting of the Decision Sciences Institute* (p. 1–18). Chicago, IL (USA): Decision Science Institute.

Sincorá, L. A., Oliveira, M. P. V. de, Zanquetto-Filho, H., & Ladeira, M. B. (2018). Business analytics leveraging resilience in organizational processes. *RAUSP management Journal*, *53*(3), 385–403. https://doi.org/10.1108/RAUSP-04-2018-002

Škerlavaj, M., Connelly, C. E., Cerne, M., & Dysvik, A. (2018). Tell me if you can: time pressure, prosocial motivation, perspective taking, and knowledge hiding. *Journal of Knowledge Management*, *22*(7), 1489–1509. https://doi.org/10.1108/JKM-05-2017-0179

Szollos, A. (2009). Toward a psychology of chronic time pressure: conceptual and methodological review. *Time & Society*, *18*(2), 332–350.

Tabatabaei, M. (2002). An experimental analysis of decision channeling by restrictive information display. *Journal of Behavioral Decision Making*, *15*(1), 419–432.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*(4), 1037–1048. https://doi.org/10.1037/a0034910

Tsiros, M., & Heilman, C. M. (2005). The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. *Journal of Marketing*, *69*(2), 114–129.

Tushman, M. L. (1979). Work characteristics and subunit communication structure: a contingency analysis. *Administrative Science Quarterly*, *24*(1), 82–98.

Vale, L. R., Sincorá, L. A., & Milhomem, L. D. S. (2018). The moderate effect of analytics capabilities on the service quality. *Journal of Operations and Supply Chain Management*, *11*(2), 101–113. https://doi.org/10.12660/joscmv11n2p101-113

Vanharanta, M., & Easton, G. (2010). Intuitive managerial thinking: the use of mental simulations in the industrial marketing context. *Industrial Marketing Management*, *39*(3), 425–436.

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639.

Wang, H., Xu, Z., Fujita, H., & Liu, S. (2016). Towards felicitous decision making: An overview on challenges and trends of Big Data. *Information Sciences*, *367–368*, 747–765.

Wang, Y., & Byrd, T. A. (2017). Business analytics-enabled decision-making effectiveness through knowledge absorptive capacity in health care. *Journal of Knowledge Management*, *21*(3), 517–539. https://doi.org/10.1108/JKM-08-2015-0301

Wang, Y. I., Highhouse, S., Lake, C. J., Petersen, N. L., & Rada, T. B. (2017). Meta-analytic Investigations of the Relation Between Intuition and Analysis. *Journal of Behavioral Decision Making*, *25*(2015), 15–25. https://doi.org/10.1002/bdm.1903

Weinhardt, J. M., Hendijani, R., Harman, J. L., Steel, P., & Gonzalez, C. (2015). How analytic reasoning style and global thinking relate to understanding stocks and flows. *Journal of Operations Management*, *39–40*, 23–30. https://doi.org/10.1016/j.jom.2015.07.003

Weiss, A. M., & Heide, J. B. (1993). The nature of organizational search in high technology markets. *Journal of Marketing Research*, *30*(2), 220–233.

Wilder, C. R., & Ozgur, C. O. (2015). Business Analytics Curriculum for Undergraduate Majors. *INFORMS Transactions on Education*, *15*(2), 180–187.

Williamson, O. E. (1981). The Economics of Organization: The Transaction Cost Approach. *The American Journal of Sociology*, *87*(3), 548–577.

Wouters, M., Anderson, J. C., Narus, J. A., & Wynstra, F. (2009). Improving sourcing decisions in NPD projects: monetary quantification of points of difference. *Journal of Operations Management*, *27*(1), 64–77.

Wright, P. (1974). The harassed decision maker: time pressures, distractions and the use of evidence. *Journal of Applied Psychology*, *59*(5), 555–561.

Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes*, *118*, 179–188. https://doi.org/10.1016/j.obhdp.2012.03.005

DECISION SCIENCES INSTITUTE
Implications of Organization Revenue for Green Practices and Performance

Susan M. L. Zee
Southeastern Louisiana University
susan.zee@selu.edu

Yun-Chen Morgan
Southeastern Louisiana University
ymorgan@selu.edu

Lillian Y. Fok
Western Washington University
fokl@wwu.edu

ABSTRACT

Organizational sustainability is heavily influenced by companies' ability to recognize the importance of green movement, and their ability to execute greener strategies. We examine the annual revenue levels of organizations and their connections to organizational green orientation and employee perceptions of impact, quality management maturity, cultural practices, and organizational performance.

Keywords: Green Movement; Quality Management; Organizational Culture.

INTRODUCTION

While most green movement research focus on green manufacturing, sustainability, green marketing, and green consumers, few studies discuss internalized characteristics of which may influence the success of companies' green initiatives. This study focus on organizations' revenue levels and explore the connections between revenue and the green movement within companies.

LITERATURE REVIEW

There have been significant interests in green movement research, such as studies focusing on sustainability based on green practices (Ahn, 2014; Lagas, 2016; Paul, Bhole, & Chaudhari, 2014) and green marketing (Maheshwari, 2014; Sujith, 2017; Yeow, Yazdanifard, 2014).

Some of the world's biggest companies – Walmart, GE, DuPont, among others – have incorporated green initiatives into their core strategies (Unruh & Ettenson, 2010).

Raluca and Cuza (2014) posited three stages to green marketing: finding solutions, cleaner technology, and quality outputs. They specifically differentiated the difference between green products and organic products. Yeow and Yazdanifard (2014) examined potential problems when failed to implement green marketing. They explained when customers are aware of production process contributing negatively to the environment; they can refuse to purchase those items. The authors also pointed out another advantage of green marketing are in the form of government assistance (i.e., grants or loans).

Total Quality Management (TQM), Sustainability, and Green Initiatives

Research examined employee involvement (Rapp & Eklund, 2002), human resource management and leadership (Daily & Huang, 2001), commitment (Matta, Davis, Mayer, & Conlon, 1996), and personality traits (Ahmad & Schroeder, 2002) all expressed possible connections between perceptions and attitudes. How employees perceive the effectiveness of TQM tools used in the workplace should have significant influences on employees' attitudes about going green.

Sustainable development is based on the perceived need to address environmental deterioration and to maintain vital functions for the well-being of all generations. It has been perceived as a new tool in company planning (Beatley & Manning, 1998) and a fundamentally important concept which should influence all policy developments within a firm (Loffler, 1998). Zairi and Liburd (2001) defined sustainability development as the ability of an organization to adapt to change in the business environment in order to deploy best contemporary methods to achieve and further maintain superior performance. Total Quality Management (TQM) represents the integrative approach in the pursuit of customer satisfaction (Chin, Pun, & Hua, 2001) and overall success. Facing intense pressure from global competition, organizations realize the need to incorporate sustainable development and TQM in order to reach higher level of improvement and ultimate profitability (Hitchcock & Willard, 2002; Jonker, 2000; McAdam & Leonard, 2003). The idea of expansion of TQM to include sustainability, and the expansion being fueled by pressure to ensure long-term survival or growth surrounded by the emphasis of globalization is thus born (Dervitsiotis, 2001; Wilkinson, Hill, & Gollan, 2001; Zairi, 2002).

Organizational Culture

Organizations incorporating green initiatives as part of their corporate culture could become the leaders in the green movement and reach higher level of sustainability and profitability (Acharya, Vadher, & Acharya, 2014). Organizational culture may impact employee perceptions of the green movement and its importance to the organization, and overall performance outcomes. We theorize that as organizations become greener, organizational culture will become more employee- and customer-centered. Employees may become more empowered and thus perform better.

Organization Size

Size is highly relevant to organizational strategy and structure, as mentioned in the literature of "Differences between Strategic Planning at Small Versus Large Firms" (Boundless, 2016). In an early research, Child (1972) highlighted the importance of *strategic choice* by organizational decision makers and its impact upon factors including organizational size and market focus. Other research has previously examined whether organizations of different size and market focus may differ on several variables related to the green movement and Quality Management in Jamaica and in the United States (Zee, Fok & Hartman, 2011). In other research, company size impacts its focus points, challenges and approaches in quality management strategies (Ahire & Golhar, 1996; Roberts, 2013). Studies can be found looking specifically into firm size versus implementation successes of quality management programs. For example, studies by Kober, Subraamanniam, and Watson , (2012), Assarlind and Gremyr (2014), as well as Melao and Guia (2015) all reinforced the differences in ability to successfully implement quality management programs between large firms and small firms. Finally, Cho (1994) and Isa, Zaroog, and Raju (2016) investigated the factors affecting successful implementation of TQM in large versus small organizations in the United States and in Malaysia, respectively. In turn, these choices ultimately impact organizational effectiveness (see especially Beckman, 2006; Kreitner & Kinicki, 2007).

RESEARCH QUESTION

In this study, we examine issues regarding the impact of size in terms of revenue to consider organizational culture, employee perceptions of organizational commitment to the green movement and QM maturity of the organization. Specifically, small revenue and large revenue organizations will have different levels of organizational green orientation, QM maturity, organizational culture, and report differing impacts of the green movement.

> Research Question: Small and large organizations in terms of annual revenue will have different levels of organizational green orientation, organizational culture, QM maturity, organizational performance, and report differing impacts of the green movement.

METHODOLOGY

Subjects of the Study

Subjects in the sample were approximately 331 full-time employees from a wide variety of industries in the South. The subjects were roughly 51.7 % male and 48% female with an average age of 41.49. The results are shown in Table 1.

Approximately 31.1% of the subjects are employed in a company which has over $1,000 million annual revenue, 4.2% of the subjects are employed in a company which has $501 to $1,000 million annual revenue, 3.9% of the subjects are employed in a company which has $251 to $500 million annual revenue, 4.2% of the subjects are employed in a company which has $101 to $250 million annual revenue, 4.5% of the subjects are employed in a company which has $51 to $100 million, 6.3% of the subjects are employed in a company which has $25 to 50 million annual revenue, and 34.1% of the subjects are employed in a company which has less than $25 million annual revenue. Subjects responded to a survey which asked about their perceptions and experiences about the green movement, quality management, and organizational culture in their own firms.

Table 1 Types of Industry and Annual Revenue of the Organizations

Industry		Frequency	Percent	Valid Percent	Cumulative Percent
	Manufacturing	42	12.7	12.7	12.7
	Financial Services	15	4.5	4.5	17.3
	Retail	73	22.1	22.1	39.4
	Utilities	9	2.7	2.7	42.1
	High Technology	7	2.1	2.1	44.2
	Education	18	5.4	5.5	49.7
	Health Care	69	20.8	20.9	70.6
	Government	5	1.5	1.5	72.1
	Other	92	27.8	27.9	100.0
	Total	330	99.7	100.0	
Missing	System	1	.3		
Total		331	100.0		

Annual revenue		Frequency	Percent	Valid Percent	Cumulative Percent
	Over $1,000 Million	103	31.1	35.2	35.2
	$501-$1,000 Million	14	4.2	4.8	39.9
	$251-$500 Million	13	3.9	4.4	44.4
	$101-$250 Million	14	4.2	4.8	49.1
	$51-$100 Million	15	4.5	5.1	54.3
	$25-$50 Million	21	6.3	7.2	61.4
	Less than $25 Million	113	34.1	38.6	100.0
	Total	293	88.5	100.0	
Missing	Unkmown	37	11.2		
	System	1	.3		
	Total	38	11.5		
Total		331	100.0		

Instrument

Organizational Green Orientation

Based on the previous research (Li, Hartman, and Zee, 2009), we measured the Organizational Green Movement by using the survey which provides twenty-one organizational green initiatives question items. The factor “Green Orientation” is a count of the total numbers of green initiatives implemented in an organization.

Quality Management (QM) Maturity

QM Maturity refers to the *degree* of QM implementation in an organization. We suggest, and previous research has shown (Ahire & Golhar, 1996; Flynn, Schroeder, & Sakakibara, 1994; Fok et al., 2000, 2001; Patti, Hartman, & Fok, 2001; Saraph, Benson, & Schroeder, 1989) that it can be measured by examining the perceived use of QM programs. In earlier research (Fok et al., 2000, 2001; Patti et al., 2001), we began the process of developing a measure of QM maturity. The instrument we developed dealt with perceived program *use* and asked respondents whether certain programs were in use in the organization, with a range from “not used” to “high usage.”

In this study, the QM maturity instrument was used to gauge QM maturity. We conducted a factor analysis to identify the underlying dimensionality. Two factors emerged from the “Usage” items. The first factor appeared to include all the traditional quality management programs and was termed “Basic Quality Improvement Tools.” The second factor was termed “Advanced Quality Improvement Tools” which includes programs like Six Sigma programs and Black Belt training. 67.39% of the variance was explained by these two factors. Table 2 below provides the items and shows the results of our factor analysis.

Table 2 Factor Analysis on Quality Programs Usage Items

Total Variance Explained

	Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %
1	2.275	37.915	37.915
2	1.768	29.475	67.390

Extraction Method: Principal Component Analysis.

Rotated Component Matrix

	Component	
	1	2
Quality management (QM) program	.646	.186
Employee suggestion channels	.714	.135
Employee quality training programs	.828	.051
Quality improvement seminars	.771	.181
Six sigma programs	.231	.903
Black belt training	.111	.930

Organizational Culture

Based on previous research (Fok et al., 2000, 2001; Harman, Fok, & Zee, 2009), we measure the Organizational Culture by constructing a series of paired opposite items which asked whether the organization's climate should be described as open vs. closed, soft vs. tough, competitive vs. collaborative, and the like. Table 3 below provides the items and shows the results of our factor analysis. We obtained a two-factor solution in case of the Organizational Culture items and have labeled Factor 1 as "Continuous Improvement Culture" and Factor 2 as "People Friendly Culture." 44.529% of the variance was explained by these two factors.

Table 3 Factor Analysis on Organizational Culture

Total Variance Explained

Component	Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %
1	2.128	26.596	26.596
2	1.435	17.933	44.529

Extraction Method: Principal Component Analysis.

Rotated Component Matrix[a]

	Component	
	1	2
Open	.759	.161
Tough	.300	.639
Formal	-.230	.656
Team-Oriented	.657	-.288
Impersonal	.376	.100
Centralized	-.097	.627
Participative	.666	-.114
Proactive	.618	-.268

Organizational Performance

The Organizational Performance items were primarily adapted from the Malcolm Baldridge National Quality Award outcome assessment measures. The Baldridge Awards are designed to identify organizations which are performing in an exceptional manner and include criteria for identifying excellence. We used the Baldridge criteria in the form of a scale which asks respondents to provide perceptions about their organizations along Baldridge lines. The resulting scale has been used and reported in previous work (Fok et al., 2000, 2001; Hartman, Fok, & Zee, 2009). Factor analysis in this study indicated that one factor was present. The results showed that 56.205% of the variance was explained by the factor and we named the factor as "Organizational Success."

Impact of Green Movement

The instruments included are items such as “Provide better products,” “Have better relationship with suppliers,” “Have better reputation,” “Increase profits,” “Reduce costs,” and “Improve productivity.” Factor analysis produced a single-factor solution and we named it “Impact of Green Movement.” 70.987% of the variance was explained by this factor.

Revenue Size

In this research, we extend the respondents to report the approximate annual revenue in their organizations. Those who reported with less than $25 million annual revenue are grouped into the small revenue group and those who reported with over $1,000 million annual revenue are grouped into the large revenue group. While those with between $25 million and $1,000 million annual revenues are dropped from the analysis.

RESULTS

Our research question suggested that small revenue (less than $25 million annual revenue) and large revenue (over $1,000 million annual revenue) organizations would have different levels of organizational green orientation, organizational culture, QM maturity, organizational performance, and impact of the green movement. As shown in Table 4, the MANOVA results are significant with p-value of .000, which implies that small revenue organizations were significantly different from large revenue organizations, and that respondents reported different levels of organizational culture, QM maturity, and impact of the green movement. Among the seven factors, we found that both “Use of Basic Quality Improvement Tools” and “Use of Advanced Quality Improvement Tools” are statistically significant at the levels of .000 and .000, respectively. For “Use of Basic Quality Improvement Tools”, the mean factor score of large revenue companies (0.255) is greater than that of small revenue companies (-0.309). For “Use of Advanced Quality Improvement Tools”, the mean factor score of large revenue companies (0.326) is greater than that of small revenue companies (-0.216). The results imply that large revenue organizations have higher levels of usage of both traditional TQM tools and advanced TQM tools than small revenue organizations. “People Friendly Culture” is significant at the level of .000. The mean factor score of small revenue organizations (0.298) is greater than that of large revenue organizations (-0.231). The results suggest that small revenue companies are more informal, more decentralized, and more personal than large revenue companies. The factor “Organizational Performance” is significant at the level of .000. The mean factor score of large revenue organizations (0.286) is greater than that of small revenue organizations (-0.221) which implies that large revenue organizations have higher levels of organizational success than that of small revenue organizations.

Table 4 Summary of MANOVA results – Revenue Size

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.944	486.196[b]	7.000	200.000	.000
	Wilks' Lambda	.056	486.196[b]	7.000	200.000	.000
	Hotelling's Trace	17.017	486.196[b]	7.000	200.000	.000
	Roy's Largest Root	17.017	486.196[b]	7.000	200.000	.000
Sales	Pillai's Trace	.201	7.198[b]	7.000	200.000	.000
	Wilks' Lambda	.799	7.198[b]	7.000	200.000	.000
	Hotelling's Trace	.252	7.198[b]	7.000	200.000	.000
	Roy's Largest Root	.252	7.198[b]	7.000	200.000	.000

Dependent Variable	Significance
Green Orientation	.148
Impact of Green Movement	.955
Continuous Improvement Culture	.781
People-Friendly Culture	.000**
Basic TQM Tools	.000**
Advanced TQM Tools	.000**
Organizational Performance	.000**

** F test is significant at the 0.05 level.

DISCUSSIONS AND CONCLUSIONS

In this research, we investigated whether companies with large annual revenue report differing perceptions of the company's impact on the environment compared to companies with small annual revenue. Results showed that smaller revenue companies are more likely to be informal employees are more likely to feel their work environment is a more people-oriented and friendlier. They also showed that participants from larger revenue organizations report higher usage of traditional TQM tools and advanced TQM tools compared to smaller revenue organizations. This can be explained by the availability of resources. Implementing TQM tools can be rather costly, and affordability may be a factor for smaller revenue companies. This study acknowledges the effort and movement to reduce environmentally harmful processes and the factors which influence the movement. Understanding the components related to the green movement should enable us to improve operations and skillsets, which could lead to bettered health and safety procedures and also employee morale.

REFERENCES

Acharya, S. G., Vadher, J. A. and Acharya, G. D. (2014). A review on evaluating green manufacturing for sustainable development in foundry industries. *International Journal of Emerging Technology and Advanced Engineering.* Vol. 4, Issue 1, pp. 232-237.

Ahire, S. L. and Golhar, D. Y. (1996) ‘Quality management in large versus small firms: An empirical investigation’, *Journal of Small Business Management*, Vol. 34, No. 2, pp. 1-13.

Ahmad, S. and Schroeder, R. G. (2002) ‘The importance of recruitment and selection process for sustainability of total quality management’, *International Journal of Quality & Reliability Management*, Vol. 19, No. 5, pp. 540-550.

Ahn, S.H. (2014) ‘An evaluation of green manufacturing technologies based on research databases’, *International Journal of Precision Engineering and Manufacturing Green Technology,* Vol. 1, No. 1, pp. 5-9.

Apple Rubber (2016) ‘5 important reasons to embrace sustainable manufacturing’, *Apple Rubber Blog.* Retrieved from https://www.applerubber.com/blog/5-important-reasons-to-embrace-sustainable-manufacturing/

Assarlind M. and Gremyr, I. (2014) ‘Critical factors for quality management initiatives in small- and medium-sized enterprises’, *Total Quality Management & Business Excellence*, Vol. 25, Issue 3-4, pp. 397-411.

Beatley, T. and Manning, K. (1998) *The ecology of place: planning for environment, economy and community*, Washington DC: Island Press.

Beckman C. M. (2006) ‘The influence of founding team company affiliations on firm behavior’, *Academy of Management Journal*, Vol. 49, No. 4, pp. 741-758.

Boundless (2016) ‘Differences between strategic planning at small versus large firms’, Boundless Management Boundless, 26 May. 2016. Retrieved 24 Aug. 2017 from https://www.boundless.com/management/textbooks/boundless-management-textbook/strategic-management-12/strategic-management-86/differences-between-strategic-planning-at-small-versus-large-firms-418-1407/

Child, J. (1972) ‘Organization structure, environment and performance: The role of strategic choice’, *Sociology: The Journal of the British Sociological Association*, Vol. 6, pp. 1-22.

Chin, K. S., Pun, K. F. and Hua, H. M. (2001) ‘Consolidation of China’s quality transformation efforts: a review’, *International Journal of Quality & Reliability Management*, Vol. 18, No. 8, pp. 836-853.

Cho, K. (1994) ‘Impact of Total Quality Management (TQM) on organizational performance in the U.S.: An empirical investigation of critical success factors’, ETD *collection for University of Nebraska - Lincoln*. AAI9425277.
http://digitalcommons.unl.edu/dissertations/AAI9425277

Daily, B. F. and Huang, S. (2001) ‘Achieving sustainability through attention to human resource factors in environmental management’, *International Journal of Operations& Production Manageement*, Vol. 21, Issue 12, pp. 1539-1552.

Deif, A. M. (2011) ‘A system model for green manufacturing’, *Journal of Cleaner Production*, Vol. 19, pp. 1553-1559.

Dervitsiotis, K. (2001) ‘Emerging elements of a world view for sustainable quality’, *Total Quality Management*, Vol. 12, No. 7, pp. 817-824.

Doszhanov, A. and Ahmad, Z. A. (2015) ‘Customers’ intention to use green products: The impact of green brand dimensions and green perceived value’, https://www.researchgate.net/publication/281733955.

Flynn, B. B., Schroeder, R. G. and Sakakibara, S. (1994) ‘A framework for quality management research and an associated measurement instrument’, *Journal of Operations Management*, Vol. 11, pp. 339-366.

Fok, L. Y., Fok, W. M. and Hartman, S. J. (2001) ‘Exploring the relationship between total quality management and information systems development’, *Information & Management*, Vol. 38, pp. 355-371.

Fok, L. Y., Hartman, S. J., Patti, A. L. and Razek, J. R. (2000) ‘The relationship between equity sensitivity, growth need strength, organizational citizenship behavior, and perceived outcomes in the quality environment: A study of accounting professionals’, *Journal of Social Behavior and Personality*, Vol. 15, No. 1, pp. 99-120.

Gan, C., Wee, H. Y., Ozanne, L. and Kao, T (2008) ‘Consumers’ purchasing behavior towards green products in New Zealand’, *Innovative Marketing*, Vol. 4, Issue 1, pp. 93-102.

Hartman, S. J., Fok, L. Y. and Zee, S. M. (2009) ‘Linkage among employee perceptions of organizational commitment to the green movement and organizational culture, and their perceived impacts upon outcomes’, *Journal of Organizational Culture, Communications and Conflict*, Vol. 13, No. 2, pp. 27-44.

Hayes, D. (1990) ‘The green decade’, *The Amicus Journal*, No. 12, pp. 24-29.

Hitchcock, D. and Willard, M. (2002) ‘Sustainability: enlarging quality's mission’, *Quality Progress*, Vol. 35, No. 2, pp. 43-47.

Isa, M. R., Zaroog, O. S. and Raju, S. (2016) ‘Survey on total quality management in small and medium manufacturing industries in Shah Alam’, *International Journal of Science and Research*, Vol. 5, No. 1, pp. 848-853.

Jonker, J. (2000) ‘Organizations as responsible contributors to society: Linking quality, sustainability and accountability’, *Total Quality Management*, Vol. 11, No. 4/5/6, pp. 741-746.

Kangovou (2017) ‘The importance of eco-friendly products’. Retrieved from https://www.kangovou.com/the-importance-of-eco-friendly-products/

Klebnikoff, S. D. (1996) ‘Eco-efficient leadership; the road to new business opportunities. *Proceeding of the APO World Conference on Green Productivity*, Manila, pp. 491-502.

Kober, R., Subraamanniam, T. and Watson, J. (2012) ‘The impact of total quality management adoption on small and medium enterprises’ financial performance’, *Accounting & Finance*, Vol. 52, Issue 2, pp. 421-438.

Kreitner, R. & Kinicki, A. (2007) *Organizational Behavior*, Boston: McGraw-Hill.

Lagas, B. (2016) ‘Five benefits of sustainability and green manufacturing’, *Quality Digest*.

Li, J., Hartman, S. J. and Zee, S. M. L. (2009) ‘A study of green movement perceptions and behavioral intentions’, *International Journal of Sustainable Economy*, Vol. 1, No. 2, pp. 133-143.

Loffler, P. (1998) ‘Sustainable development in Europe’, *A Cause for Regional Environment*, Vol. 8, pp. 113-120.

Ma, C., Liu, X., Zhang, H. and Wu, Y. (2016) ‘A green production strategies for carbon sensitive products with a carbon cap policy’, *Advances in Production Engineering and Management*, Vol. 11, No. 3, pp. 216-226.

Maheshwari, S. P. (2014) ‘Awareness of green marketing and its influence on buying behavior of consumers: Special reference to Madhya Pradesh, India’, *AIMA Journal of Management & Research*, Vol. 8, Issue 1/4, pp. 497-505.

Matta, K., Davis, J., Mayer, R. and Conlon, E. (1996) ‘Research questions on the implementation of total quality management’, *Total Quality Management*, Vol. 7, Issue 1, pp. 39-49.

McAdam, R. and Leonard, D. (2003) ‘Corporate social responsibility in a total quality management context: Opportunities for sustainable growth’, *Corporate Governance*, Vol. 3, Issue 4, pp. 36-45.

Melao, N. F. and Guia S. M. (2015) ‘Exploring the impacts of ISO 9001 on small- and medium-sized social service institutions: a multiple case study’, *Total Quality Management & Business Excellence*, Vol. 26, Issue 3-4, pp. 312-326.

Patti, A. L., Hartman, S. J. and Fok, L. Y. (2001) ‘Investigating organizational quality management maturity: An investment validation study’, *International Journal of Quality & Reliability Management*, Vol. 18, No. 9, pp. 882-899.

Paul, I. D., Bhole, G. P. and Chaudhari, J. R. (2014) ‘A review on green manufacturing: Its importance, methodology and its application’, *Procedia Materials Science,* Vol. 6, pp. 1644-1649.

Peattie, K. (2001) ‘Towards sustainability: The third age of green marketing’, *The Marketing Review*, Vol. 2, No. 2, pp. 129-146.

Rapp, C. and Eklund, J. (2002) ‘Sustainable development of improvement activities -the long-term operation of a suggestion scheme in a Swedish company’, *Total Quality Management*, Vol. 13, No. 7, pp. 945-969.

Roberts, M. (2013) ‘Large versus small business quality management strategies’, http://blog.lnsresearch.com/bid/179410/large-versus-small-business-quality-management-strategies-data

Sandu, R. (2014) ‘Green: Marketing, products, and consumers’, *SEA-Practical Application of Science*, Vol. 2, Issue 3(5), pp. 555-562.

Saraph, J. V., Benson, G. P. and Schroeder, R. G. (1989) ‘An instrument for measuring the critical factors of quality management’, *Decision Sciences*, Vol. 20 pp. 810-829.

Sen, P. K., Bohidar, S. K., Shrivas, Y., Sharma, C. and Modi, V. (2015) ‘Study on innovation, research and recent development in technology for green manufacturing’, *International Journal of Mechanical Engineering and Robotics Research*, Vol. 4, No. 1, pp. 185-194.

Sezen, B. and Cankaya, S. Y. (2013) ‘Effects of green manufacturing and eco-innovation on sustainability performance’, *Procedia – Social and behavioral Sciences,* Vol. 99, pp. 154-163.

Sujith, T.S. (2017) ‘Awareness of green marketing and its influence on buying behavior of consumers in Kerala’, *International Journal of Scientific Research and Management*, Vol. 5, No. 7, pp. 43.

Wilkinson, A., Hill, M. and Gollan, P. (2001) ‘The sustainability debate’, *International Journal of Operations & Production Management*, Vol. 21, No. 12, pp. 1492-1502.

World Commission on Environment and Development (WCED), 1987. Report: ‘Our Common Future’.

Yang, Y. C. (2017) ‘Consumer behavior towards green products’, *Journal of Economics, Business and Management*, Vol. 5, No. 4, pp. 160-167.

Yeow, K. Y. and Yazdanifard, R. (2014) ‘The concept of green marketing and green product development on consumer buying approach’, *Global Journal of Commerce & Management Perspective*, Vol. 3, pp. 33-38.

Zairi, M. (2002) ‘Beyond TQM implementation: the new paradigm of TQM sustainability’, *Total Quality Management*, Vol. 13, No. 8, pp. 1161-1172.

Zairi, M. and Liburd, I. M. (2001) ‘TQM sustainability – a roadmap for creating competitive advantage’, *Integrated Management, Proceeding of the 6th International Conference on ISO 9000 and TQM*, Paisley, Ayr, Scotland, 17-19 April, pp. 452-461.

Zee, S. M., Fok, L. Y., and Harman, S. J. (2011) ‘Exploring the relationships between organizational size and market focus and commitment to the green movement and impacts of organizational culture: A comparative study of Jamaica and the United States’, *International Journal of Business and Social Science*, Vol. 2, No. 22, pp. 19-34.

DECISION SCIENCES INSTITUTE

Investigating Firms' Strategies and Influencers of Energy Management in Green Manufacturing Practices

Madhavi Latha Nandi
University of Texas Rio Grande Valley
Email: madhavilatha.nandi01@utrgv.edu

Santosh Nandi
University of South Carolina
Email: nandisan@uscsumter.edu

Hiram Moya
University of Texas Rio Grande Valley
Email: hiram.moya@utrgv.edu

ABSTRACT

Manufacturing firms consume electrical energy in large quantities. Reducing energy consumption is a concern not only for governments, but also for firms. By reducing consumption of resources needed to produce electricity, firms' can significantly increase their productivity and conserve environment. In green manufacturing literature, the topic of energy conservation in manufacturing context has received less research attention when compared to other green practices such as eco-friendly product development and pollution reduction. In this study, we address this gap by investigating on what are the predominant green strategies for energy conservation implemented by US manufacturing firms in *Industrial Machinery and Equipment* sector. In addition, the study identifies the economic and organizational factors that influence firms' greening efforts of energy management. To achieve these two research objectives, energy assessment data comprising of 119 manufacturing firms is analyzed using descriptive statistics and Poisson regression technique. Findings of the study are expected provide insights to relevant stakeholders, such as manufacturing firms, governmental bodies, energy supplying companies, and universities, in collaborating for designing and emphasizing on appropriate greening interventions needed to promote green manufacturing practices pertaining to energy conservation.

KEYWORDS: Green manufacturing, Energy management, Green strategies, Energy conservation, Energy sustainability

INTRODUCTION

In recent past, manufacturing firms are forced to adopt an environmental perspective in planning and executing their operations due to several factors including concerns on deterioration of natural resources, customer preferences for green products/services, and regulatory pressure from government. Subsequently, 'green manufacturing' emerged as a topic of interest for global community including governments, firms, consumers and researchers. Green manufacturing refers to green strategies and techniques that firms employ in their production systems to become more eco-efficient (Deif, 2011). It reflects a manufacturing approach of the firms, that is aware of the impact of their production and product on environment. Some examples of green manufacturing practices include creating products with less material and energy consumption, using more eco-friendly materials, and renewable energy for production, reducing unwanted outputs and recycling the used products (Helu & Dornfeld, 2013).

Existing academic research on green manufacturing mainly comprises of studies that focus on developing frameworks and indices to assess, measure and monitor the greening efforts (Deif, 2011; Salem & Deif, 2013), impact of green manufacturing on firms' performance (Rusinko, 2007), green manufacturing practices pertaining to raw material usage and recycling in specific industries (Rusinko, 2007) and green energy recommendations (Anderson & Newell, 2004). Studying upon influencing factors, such as level of disruption, environment norms and local adopters, on implementation of a government recommendation is yet another underdeveloped area in green practices research (Muthulingam et al., 2013; Dowell & Muthulingam, 2017). In this vein, the present study focuses on green manufacturing practices pertaining to reducing energy consumption, which is a relatively under-researched area in comparison to other green manufacturing practices (Oh & Sun, 2016). According to a recent report on greenhouse gas (GHG) emissions by the United States Environmental Protection Agency (USEPA), electricity generation contributes the most to the total GHG emissions, accounting for 28% of the total GHG emissions (USEPA, 2016). Further, manufacturing sector is the third largest contributor to the total GHG emissions, accounting for 22% of the total GHG emissions. GHG emissions in manufacturing sector are produced directly from burning fuel for energy, chemical reactions, and from leakages of industrial processes or equipment. Manufacturing sector also contributes indirectly to the GHG emissions by the use of electricity that is produced by burning fossil fuels and other natural resources. In its attempt to reduce the amount of GHG emissions resulting from energy generation and consumption, the US government is emphasizing on energy conservation through various initiatives for manufacturing firms such as providing free energy assessment consultancy services, providing tax incentives for using renewable energy, and penalizing excessive energy consumption. As such, when manufacturing firms implement energy recommendations by the government, they compare costs of implementing a government recommendation with their energy savings (Anderson & Newell, 2004). In this background, the present study aims to extend the stretch of existing literature by providing an overall understanding of predominant green strategies implemented by firms in *Industrial Machinery and Equipment* sector.

The specific research questions addressed in this research study are as follows:

RQ1: What are the green strategies in energy management that manufacturing firms in *Industrial Machinery and Equipment* sector predominantly apply?

RQ2: What are the economic and organizational factors that influence manufacturing firms' greening efforts of energy management in *Industrial Machinery and Equipment* sector?

The findings are expected to provide deeper insights to scholars about the factors influencing the energy conservation efforts of firms in *Industrial Machinery and Equipment* sector. From a practitioner perspective, managers shall benefit from this study in designing appropriate greening interventions and in approaching energy conservation efforts more proficiently at least in *Industrial Machinery and Equipment* sector. Relevant stakeholders, including governmental bodies, energy companies and supply chain partners related with *Industrial Machinery and Equipment* sector., will additionally benefit from these findings in formulating useful collaborations to endorse for green manufacturing practices. The subsequent sections discuss the possible green strategies in energy management and methodology, followed by conclusion.

GREEN STRATEGIES IN ENERGY MANAGEMENT

Oh & Soh (2016) suggest that manufacturing firms can potentially deploy combinations of four different strategies in their energy management for achieving green manufacturing systems, which are peak demand management of electric power, production energy management, utility energy management, and use of renewable energy. Each of these green strategies in energy management are discussed in detail in the following sections.

Peak Demand Management

Peak demand management of electric power refers to a firm's efforts to reduce and manage load spikes in their manufacturing operations, thereby, reducing load on the power grid system. Achieving a perfect real-time balance between supply and demand is a critical factor for reliable operation of power grid systems (Albadi & El-Saadany, 2008). Industrial consumers such as manufacturing firms play a key role in achieving this supply-demand balance of power grid (Zhang & Grossman, 2016). Load spikes in manufacturing firms result when firms employ high short-term energy consumption processes within the same time-periods. Load spikes subsequently translate to high demand charges, which typically constitute major portion of the electricity costs for manufacturing firms. Peak demand management of electric power by manufacturing firms not only reduces the overall electricity costs of the firm, but also contributes to overall greening efforts at national level through reduction of social costs associated with generation of electricity and consistent operational performance of the power grid system (Oh & Sun, 2016). Firms manage peak demands using a different combination of methods such as use of alternate sources of energy supply such as power generators and solar energy to cater to peak demands, storing electric energy for later use to offset the load spikes, and make relevant changes in processes that can augment peak load reduction. The former two solutions for peak demand management of electric power involve high capital cost investments by the firms, while the latter solution involves significant implementation effort in terms of coordination among the different departments and adjustment to the changed processes and routines.

Production Energy Management

Production energy management refers to a firm's efforts to reduce the energy consumption in its production processes by upgrading the machinery, improving the utilization rate of production equipment, and exploring ways to reduce task time. Traditionally, manufacturing firms handled production and energy management separately, by determining production schedule first following by planning for electricity procurement (Zhang & Grossman, 2016). Such an approach typically led to suboptimal solutions by blindsiding firms on possible improvements that could be achieved on the production side (Zhang & Grossman, 2016). In the recent past, with mounting pressure to optimize production costs and adopt environment-friendly production systems, manufacturing firms are gradually shifting to an integrated approach for production and energy management (May et al., 2016). This approach however, is associated with significant capital costs and implementation effort in terms of identifying optimal schedule for operations, implementing process improvement programs, researching for better materials and method alternatives, integrating information flows, laying out preventive maintenance schedules, and investing on better machinery and equipment (May et al., 2017).

Utility Energy Management

Utility energy management refers to management of energy consumption by the utilities that essentially support core production processes such as heating, cooling, humidity control and lighting. Utility energy management is an easy target to improve energy efficiency, without affecting the core processes. Typically, utilities in a production system account for 30% of the energy consumption and therefore, is a hotspot for energy efficiency innovations and improvements (May et al., 2016). Continuous product innovations by utility equipment suppliers in bringing more eco-friendly products into market and development of smart devices further complement the potential for greening effort of firms in this area. These improvements typically do not entail large capital costs or implementation costs to companies.

Use of Renewable Energy

The strategy of using renewable energy refers to a firm's efforts to either generate or procure renewable energy for its consumption. Manufacturing firms are increasingly exploring the possibility of generating renewable energy, especially, solar energy to offset their energy requirements and support their objective of achieving green manufacturing systems, relatively lower levels of thermal efficiency, lack of reliability, and lack of sufficient market base to stimulate related technological innovations and standardization (Eleftheriadis & Anagnostopoulou, 2015; Luthra et al., 2015).

RESEARCH METHODOLOGY

Sampling and Analysis

A section of data of the Industrial Assessment Centers (IAC) program, an energy management initiative of the United States Department of Energy, is coded and analyzed. The IAC program partners with several reputed universities to establish IAC centers across the U.S. The IAC centers provide free consultancy services to medium and large sized manufacturing firms for energy and natural resource management and provide recommendations for improvement. Thereafter, the IAC centers follow-up with the consulted manufacturing firms on implementation of the recommendations and resulting savings from their implementations. In this study, we limit the scope to energy related assessments and recommendations of firms in one sector, namely, *Industrial Machinery and Equipment* sector. Accordingly, a random sample of 119 firms from the IAC database is extracted and carefully coded for analysis purpose. The details of the variables considered for the data analysis are presented in Appendix A.

In our first analysis (Analysis 1), we identified predominant green strategies of energy management that the sample firms' have implemented based upon IAC recommendations. The predominant strategies are identified by analyzing the total number of recommendations given to the firms in the data sample and total number of recommendations implemented by the firms for each green energy management strategies. In our second analysis (Analysis 2), we identified the economic and organizational factors that influence the sample firms' greening efforts of energy management using Poisson regression technique. The greening effort is represented by the total number of recommendations in energy management that have been implemented by the firms. Organizational and economic factors that are considered for the analysis include level of demand management recommendations, level of production management recommendations, level of utility management recommendations, annual sales of the firm, total employees of the firm, plant area, annual production hours, potential savings expected from the recommendations, annual electrical costs, and total costs associated with the recommendation implementation. Poisson regression technique is apt for the present research model since the dependent variable is a count variable and approximates a Poisson distribution (Nussbaum et al., 2011; Dunteman & Ho, 2011). A variable can be approximated to follow a Poisson distribution if it has similar values of mean and variance. In the present case, the dependent variable has a mean of 5.5 and the variance of 5.79 as shown in Table 1. Also, some of the independent variables in the research model are categorical, while others are continuous. Poisson regression model accommodates both types of independent variables (Nussbaum et al., 2011; Dunteman & Ho, 2011). Poisson regression is conducted using IBM-SPSS software package for this study.

Table 1: Mean and Variance of the Dependent Variable

Dependent Variable	Mean	Standard Deviation	Variance
TotRecImp	5.55	2.407	5.794

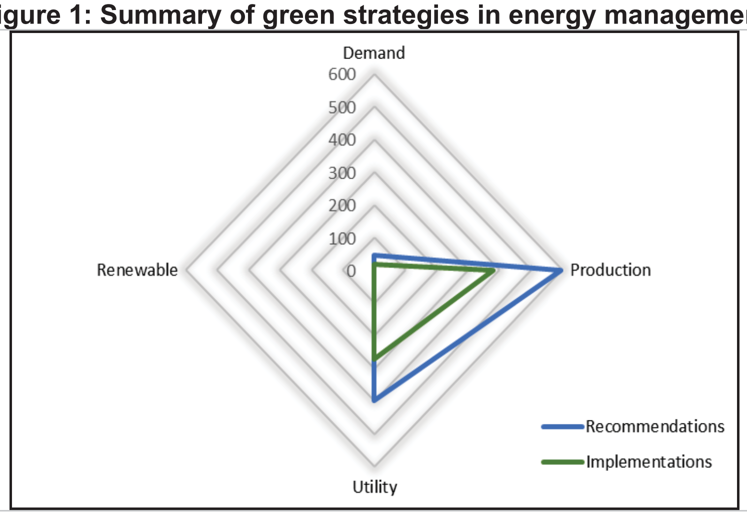
RESULTS

The results for Analysis 1 are presented in Table 2 and Figure 1. The results revealed that the production energy management and utility energy management predominantly implemented by the firms. These two strategies accounted for 96% of the recommendations indicating a greater potential for bringing in energy efficiencies and subsequently, contribute to the greenness of the production system. Demand management recommendations accounted for 4% of the recommendations, indicating that the firms in this sector are doing well with management of electricity loads in their operation. There were no recommendations offered to the manufacturing firms by the IAC centers related to use of renewable energy. One of the possible reasons for the lack of recommendations in this area could be that the recommendations related to use of renewable energy are not viable for the huge and critical power requirements on the firms in this sector. Also, the large capital costs associated with laying out the infrastructure for generating renewable energy do not necessarily justify the expected benefits from these recommendations.

Table 2: Results of Analysis 1

Strategy	Recommendations	Implementations
Demand	45	17
Production	595	377
Utility	399	271
Renewable	0	0
Total	1039	665

Figure 1: Summary of green strategies in energy management



The results of Analysis 2 are presented in Table 3, Table 4 and Table 5. The independent variable representing recommendations related to use of renewable, RenRecLevel, is dropped from the analysis, since there are no recommendations given in this category. Table 3 presents the results of goodness-of-fit test that assesses how well the regression model fits the data. The parameter "Value/df" for Pearson Chi-Square needs to close to 1 confirming the assumption of equi-dispersion (Bruin, 2016; Laerd Statistics, 2018). For the study's model the value for this parameter is 0.737 indicating slight but acceptable level of under-dispersion. Overall, the regression model can be assumed to fit the study data. The results for Omnibus Test, which is a likelihood ratio test that compares the full regression model with all the independent variables with the intercept-only model (Bruin, 2016; Laerd Statistics, 2018). The p-value of .000 in the present analysis indicates a statistically significant overall model. The results for the "Tests of Model Effects" indicating the statistical significance of each independent variable is presented in Table 5. The results indicate that the variables ProdRecLevel, UtilRecLevel, Sales, Employees and TotSav have statistically significant influence on the dependent variable TotRecImp. The parameter estimates of the independent variables are presented in Table 6. The regression model for the study is presented in the equation below.

$$\begin{gathered} logTotRecImp = 1.875 + exp(-0.583)*ProdRecLevel1 \\ + exp(-0.393)*ProdRecLevel2 \\ + exp\ (-0.391)UtilRecLevel1 \\ + exp(-0.002)Sales \\ + exp(-0.001)Employees \\ + exp(3.47*10^{-6})TotSav \\ + exp(-2.465*10^{-7})ElecCost \end{gathered}$$

Table 3: Goodness of Fit

	Value	**df**	**Value/df**
Deviance	78.326	106	0.739
Scaled Deviance	78.326	106	
Pearson Chi-Square	78.159	106	0.737
Scaled Pearson Chi-Square	78.159	106	
Log Likelihood[a]	-246.262		
Akaike' Information Criterion (AIC)	518.523		
Finite Sample Corrected AIC (AICC)	521.990		
Bayesian Information Criterion (BIC)	554.652		
Consistent AIC (CAIC)	567.652		

DependentVariable: TotRecImp
Model: (Intercept), DemRecLevel, ProdRecLevel, UtiRecLevel, Sales, Employees, PlantArea, ProdHours, TotSav, ElecCost, TotCost

a. The full log likelihood function is displayed and used in computing information criteria

Table 4: Omnibus Test[a]		
Likelihood Ratio Chi Square	**df**	**Sig.**
48.629	12	0.000

a. Compares the fitted model against the intercept-only model.

Table 5: Tests of Model Effects and Parameter Estimates

Source	Wald Chi Square	df	Wald Chi Square Sig.	B	B Sig.
(Intercept)	127.257	1	**0.000**	**1.875**	**0.000**
DemRecLevel	0.001	1	0.981		
DemRecLevel1				0.002	0.981
ProdRecLevel	21.271***	2	**0.000**		
ProdRecLevel1				**-0.583**	**0.034**
ProdRecLevel2				**-0.393**	**0.000**
UtiRecLevel	4.749**	2	**0.093**		
UtiRecLevel1				**-0.391**	**0.035**
UtiRecLevel2				-0.058	0.582
Sales	4.944**	1	**0.026**	**-0.002**	**0.026**
Employees	3.237*	1	**0.072**	**0.001**	**0.072**
PlantArea	0.793	1	0.373	4.286E-9	0.373
ProdHours	0.010	1	0.920	8.568E-7	0.920
TotSav	8.175***	1	**0.004**	**3.470E-6**	**0.004**
ElecCost	2.464	1	0.116	-2.465E-7	0.116
TotCost	0.551	1	0.458	-4.146E-7	0.458

Note: Significance value: * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

CONCLUSION

The present study seeks to investigate the predominant green strategies for energy conservation adopted by the US firms in Industrial Machinery and Equipment sector and identify the economic and organizational factors influencing the greening effort in energy management of these firms. The study results indicate that manufacturing firms in the focus sector predominantly implemented production energy management and utility energy management strategies in their greening efforts. The economic and organizational factors that influence the greening effort include level of production recommendations, level of utility management recommendations, annual sales of the firm, number of employees, potential savings resulting from implementation of the recommendations. Surprisingly, annual cost of electricity has a positive but somewhat weak influence on total recommendations implemented by the firm. The findings of the study point out that the costs associated with recommendation implementations do not seem to affect the firms' greening effort, which indicates firms' positivity in committing to green manufacturing practices in Industrial Machinery and Equipment sector.

The study adds new perspectives to the existing literature on implementation of environment focused energy-saving initiatives in at least two ways. First, our study replicates earlier studies on energy-saving initiatives. When we compared our study findings based on recent IAC energy data for the period 2000 – 2017 with that of an earlier study by Anderson & Newell (2004) based on IAC energy data for the period 1981 – 2000, both studies appeared consistent about adopting production and utility energy-saving recommendations. Similar to Anderson & Newell (2004), we find that firms are likely to adopt production and utility energy savings recommendations when such adoptions can lead firms to save of energy costs. Counterintuitively, this finding suggests that firms may be less likely to commit to invest in environment focused energy savings measures when there is lesser scope for immediate returns. However, this is not the case in our data sample of randomly selected 150 firms representing Industrial Machinery and Equipment sector. Our results show that the effect of implementation costs is not significant on the number of energy related recommendations that a firm has implemented. As such, this finding contradicts with earlier studies, such as Dowell & Muthulingam (2017), that suggested implementation costs as a major factor. We argue that this inconsistent finding is an outcome of the industry effect (i.e., specific to Industrial Machinery and Equipment sector), or the effect of type of recommendations (i.e., energy related). Second, our study finds that firm characteristics, such as annual sales of the firms, number of employees, annual electricity costs, act as significant proxies to indicate for firms' likelihood of adopting energy-saving recommendations. Our finding is based on our independent treatment of each attribute of firm characteristics as an orthogonal statistical approach to prior studies, e.g., Dowell & Muthulingam (2017), that controlled for firm characteristics.

The study is expected to provide deeper insights to researchers and managers about factors influencing the energy conservation efforts in Industrial Machinery and Equipment sector. As such, researchers at national level energy guiding agencies, such as IAC, can outrightly use several of the study findings to support their ongoing consulting efforts to manufacturing firms. The study offers a model that can help managers in predicting the number of energy management recommendations that are likely to be implemented by the firms in this sector based on the significant factors found in the analysis. Future research studies can test this model can for manufacturing firms in other sectors to gain sector-specific insights. The study has some limitations as well. One main limitation relates to the under-dispersion of the dependent variable. Even though the under-dispersion of the dependent variable is expected to not hinder the robustness of the study results, further in-depth examination can offer rich insights on effect of under-dispersion of the dependent variable in a Poisson regression model.

REFERENCES

Albadi, M.H., & El-Saadany, E.F. (2008). A summary of demand response in electricity markets. *Electric power systems research*, *78*(11), 1989-1996.

Anderson, S. T., & Newell, R. G. (2004). Information programs for technology adoption: the case of energy-efficiency audits. *Resource and Energy Economics*, *26*(1), 27-50.

Deif, A.M. (2011). A system model for green manufacturing. *Journal of Cleaner Production*, *19*(14), 1553-1559.

Dowell, G. W., & Muthulingam, S. (2017). Will firms go green if it pays? The impact of disruption, cost, and external factors on the adoption of environmental initiatives. *Strategic Management Journal*, *38*(6), 1287-1304.

Dunteman, G.H., & Ho, M.H.R. (2011). *An Introduction to Generalized Linear Models: Poisson Regression.* Sage Publications, Thousand Oaks, 44-52.

Eleftheriadis, I.M., & Anagnostopoulou, E.G. (2015). Identifying barriers in the diffusion of renewable energy sources. *Energy Policy*, *80*, 153-164.

Helu, M., & Dornfeld, D. (2013). Principles of green manufacturing. In *Green Manufacturing* (pp. 107-115). Springer, Boston, MA.

Laerd Statistics, 2018. *Poisson Regression Analysis using SPSS Statistics*. Lund Research Ltd. Retrieved from: https://statistics.laerd.com/spss-tutorials/poisson-regression-using-spss-statistics.php/.

Luthra, S., Kumar, S., Garg, D., & Haleem, A. (2015). Barriers to renewable/sustainable energy technologies adoption. *Renewable and sustainable energy reviews*, *41*, 762-776.

May, G., Stahl, B., & Taisch, M. (2016). Energy management in manufacturing: Toward eco-factories of the future–A focus group study. *Applied Energy*, *164*, 628-638.

May, G., Stahl, B., Taisch, M., & Kiritsis, D. (2017). Energy management in manufacturing: From literature review to a conceptual framework. *Journal of cleaner production*, *167*, 1464-1489.

Nussbaum, E.M., Elsadat, S., & Khago, A. H. (2008). Best practices in analyzing count data: Poisson regression. *Best practices in quantitative methods*, 306-323.

Oh, E., & Son, S.Y. (2016). Toward dynamic energy management for green manufacturing systems. *IEEE Communications Magazine*, *54*(10), 74-79.

Rusinko, C. (2007). Green manufacturing. *IEEE Transactions on Engineering Management*, *54*(3), 445-454.

Salem, A.H., & Deif, A.M. (2014). An integrated approach to assess manufacturing greenness level. *Procedia CIRP*, *17*, 541-546.

USEPA. (2016). *Sources of Green House Gas Emissions.* United States Environmental Protection Agency. Retrieved from: https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions/.

Zhang, Q., & Grossmann, I.E. (2016). Enterprise-wide optimization for industrial demand side management: Fundamentals, advances, and perspectives. *Chemical Engineering Research and Design*, *116*, 114-131.

APPENDIX A

Variable Name	Variable Description
TotRecImp	Total energy management recommendations implemented
DemRecLevel	Level of demand management recommendations Categorical variable: 1: 0 recommendations 2: 1 to 3 recommendations 3: 4 and above recommendations
ProdRecLevel	Level of production management recommendations Categorical variable: 1: 0 recommendations 2: 1 to 3 recommendations 3: 4 and above recommendations
UtilRecLevel	Level of utility management recommendations Categorical variable: 1: 0 recommendations 2: 1 to 3 recommendations 3: 4 and above recommendations
RenRecLevel	Level of renewable energy use recommendations Categorical variable: 1: 0 recommendations 2: 1 to 3 recommendations 3: 4 and above recommendations
TotCost	Expected cost of energy management recommendation implementation
TotSav	Potential savings resulting from implementation of energy management recommendations
ElecCost	Annual electricity consumption cost
Sales	Annual sales of the firm
Employees	Number of employees working in the firm
Plant Area	Built-up area of the production system
ProdHours	Annual hours of production operations

DECISION SCIENCES INSTITUTE

Is Big Data Analytics Half Embraced by Construction Industry?

Niharika Dayyalal
Illinois State University
Email: ndayyal@ilstu.edu

Sundeep Inti
Purdue University Northwest
Email: sinti@pnw.edu

Shahrukh Khan
Purdue University Northwest
Email: khan270@pnw.edu

ABSTRACT

Big Data Analytics (BDA) supports industries by providing insights from data. A significant amount of previous research documented that the construction industry, embraced BDA poorly. This study explores the dispersion of BDA in various sectors of construction using a systematic literature search. The study identifies that BDA penetrated well into transportation sectors such as highways and railways compared to building construction. The reasons for uneven absorption are deficiency in data science education in civil engineering curriculum, lack of simple BDA tools, resistance to embrace BDA, etc. The study also presents recommendations for effective diffusion of BDA in construction.

KEYWORDS: Big Data, Construction, Transportation Sector, Building Sector, Data Analytics

DECISION SCIENCES INSTITUTE

Improving Model Performance in Healthcare Predictive Analytics

Piyush Jain
Florida Atlantic University
pjain2012@fau.edu

C. Derrick Huang
Florida Atlantic University
dhuang@fau.edu

Ravi S. Behara
Florida Atlantic University
rbehara@fau.edu

Ankur Agarwal
Florida Atlantic University
aagarwa2@fau.edu

ABSTRACT

Healthcare data used in predictive analytics are typically highly imbalanced with high dimensionality, which affects classification performance negatively. This study focuses on improving model performance in healthcare analytics under these constraints. It uses data sampling to reduce imbalance and feature selection to reduce high dimensionality, thereby improving overall model performance.

KEYWORDS: Hospital Readmission, High Imbalance, High dimensionality, Data Sampling, Feature Selection

INTRODUCTION

Electronic health record (HER) systems produce high volumes of healthcare data which can be used to develop machine learning and classification algorithms to analyze and predict various healthcare challenges. Healthcare data including patient discharge summaries are of high predictive value (Isaac et al., 2015; Joseph et al., 2015). But the discharge summary datasets used in prior studies are highly imbalanced and have high dimensionality, which is a common characteristic of all healthcare datasets because the minority (positive) class is often the class of interest. In addition, the feature set extracted from discharge summaries generally has high cardinality (Shanab et al., 2011).

High imbalance in healthcare datasets is caused due to uneven distribution between a minority of positive cases and a majority of negative cases. Classification algorithms tend to assign the predicted label to the majority class in order to reduce the overall misclassification rate (Shanab et al., 2011), resulting in poor classification performance, higher false negatives, and higher false positives. These usually result in overfitting the data (Soltani et al., 2011; Sandhan et al., 2014).

In addition, a wide variety of data including data from various hospitals, various health practitioners, and their varied methods of writing discharge summaries in the case of text data, also result in high dimensionality datasets. In this study, we used a bag-of-words model to extract features from all discharge summaries which resulted in a large number of features. The high dimensional dataset shows problems such as overfitting, more space and memory needs, and requires high processing power for data analysis. The solution is to filter the features and use these filtered datasets to measure the classification performance. The basic assumption behind selecting optimal feature sets is that they would provide higher classification accuracy, since the subset captures all basic information about the dataset and excludes features which are noisy or irrelevant (Lusa and Blagus, 2012; Taşkın et al., 2017).

The dataset used in this study shows high imbalance and high dimensionality with a total number of 191,205 patient discharge summaries with 8,204 records for positive (minority) cases and 183,001 negative (majority) cases. This study used a data sampling technique to reduce the imbalance and achieve better classification performance. A total of 391 features were extracted by using bag-of-words model with a minimum document frequency of 0.05. The high dimensionality of the resulting dataset is then reduced by using a feature selection technique.

The Naive Bayes classification algorithm was used in this study. A Naive Bayes implementation allows the system to be highly scalable and non-sensitive to irrelevant data. In addition, newly acquired patient data can be incrementally utilized to train the model which will allows the model to be relevant over time. The model was designed on the High Performance Computing Cluster (HPCC), a big data platform. This allows the model to be used in real life scenarios where new discharge summaries can be continuously added and utilized to improve the current model.

This study addresses the hospital readmission problem for patients with Chronic Obstructive Pulmonary Disease (COPD) because it is a major challenge, with almost 20% of Medicare beneficiaries hospitalized for COPD getting readmitted within a 30-day period (Surya et al., 2017). Further, a significant percentage of these readmissions are avoidable. Hence there are many ongoing efforts to develop effective predictive models to reduce this.

LITERATURE REVIEW

The growth of healthcare related information has led to an increase in implementation of various healthcare related machine learning tools and platforms to support decision making, such as identifying high-cost patients, analyzing and predicting hospital readmissions (Carolyn, 2012) and improving patient triage where the risk of complications is estimated (Xiaoqian et al., 2012).

The focus of this study is to predict and analyze hospital readmissions for COPD within a 30-day period of initial discharge. This study uses discharge summaries as opposed to claims data (Danning et al., 2014) because discharge summaries are available in a timely fashion when the patient is being discharged. They also summarize all the information about patient's condition at the time of discharge. In contrast, claims data takes significant time to get generated and processed. Since hospital readmissions must be predicted within 30-days of initial discharge from hospital, discharge summaries provide a better option to generate true predictions in a timely fashion. This would support timely and effective preventative interventions to help mitigate the risk of avoidable patient readmissions.

As indicated earlier, since the discharge summary dataset is produced by various hospitals and healthcare practitioners, the data has many features and is highly imbalanced. According to the study performed by (Van Hulse et al., 2011), sampling techniques improve the performance of various classification algorithms where the dataset is highly imbalanced. Their study generated a total of 1,232,000 classifiers using 35 different benchmark datasets, 11 learning algorithms and 7 different sampling techniques. It further showed that Random Under-Sampling (RUS) performed better than other intelligent sampling techniques across different learners. In another study (Seiffert et al., 2009), RUS was also shown to perform better than other sampling techniques, for software quality prediction dataset. Joffrey et al. (2018) also show that the Random Under-Sampling technique improves the overall classification performance for a range of learners, datasets and environments. Khoshgoftaar et al. (2014), shows that RUS in conjunction with other performance improvement techniques achieve good results for six different highly imbalanced bioinformatics datasets.

To reduce high dimensionality in the data, feature selection techniques work best as compared to other dimensionality reduction techniques. This improves overall performance, and reduces the memory and processing power requirements. According to Gao et al. (2011), feature selection plays a very important role in reducing dimensionality. Kira and Rendell (1992) developed the first Relief algorithm as a feature selection method based on evaluating individual features (Figure 1). The algorithm calculates a feature weight (also referred to as feature score) for each feature that can be used to estimate its relevance to the predicting value. It ranges from -1 (worst) to +1 (best). The original algorithm was limited to binary classification problems with many limitations including the lack of ability to handle missing data. So a variety of methods have since been developed to extend the Relief algorithm. Urbanowicz et al. (2018) review relief-based feature selection research in further detail.

Require: for each training instance a vector of feature values and the class value
$n \leftarrow$ number of training instances
$a \leftarrow$ number of features (i.e. attributes)
Parameter: $m \leftarrow$ number of random training instances out of n used to update W

initialize all feature weights $W[A] := 0.0$
for i:=1 **to** m **do**
randomly select a 'target' instance R_i
find a nearest hit 'H' and nearest miss 'M' (instances)
for A:= 1 **to** a **do**
$W[A] := W[A] - diff(A, R_i, H)/m + diff(A, R_i, M)/m$
end for
end for
return the vector W of feature scores that estimate the quality of features

Figure 1: Pseudo-code for the original Relief algorithm (source: Urbanowicz, R.J et al., 2018)

Their study evaluated various feature selection techniques and measured the stability. ReliefF was shown to be most stable and achieves the best classification performance. It is the most used Relief-based algorithm, and the 'F' in ReliefF refers to the sixth algorithm variation (from A to F) proposed by Kononenko (1994). Visalakshi and Radha (2014), evaluates different feature selection techniques and their performance measures in various datasets and shows that feature selection improves the performance. According to Khoshgoftaar et al. (2011) ReliefF performs best when compared with other techniques. The algorithm continues to be improved in some recent applications (Wang et al., 2016). So we find that prior studies performed in various application areas on high imbalance and high dimensionality show that sampling and feature selection techniques work best in varying scenarios and provide better classification performance. However our literature review indicates that these techniques have not been used for the COPD hospital readmission problem.

Shanab et al. (2011) evaluated various sampling and feature selection techniques. The best approach takes the original data and applies data sampling techniques. The sampled data set is then used for feature ranking, and the features with high ranks are then selected. These selected feature values from the sampled dataset are used for analysis. This approach was adopted in the current study.

The Naive Bayes implementation in this study allows the framework to be highly scalable and non-sensitive to irrelevant data. According to Bingwei et al. (2013), Naive Bayes tend to perform better when the dataset increases and suggests that Naive Bayes provides simple data fusion architecture which allows the algorithm to be flexible and elastic. In a real-world setting, newly acquired and sampled patient data can be incrementally utilized to train the model which will allows the model to be relevant over time. The model in this study was designed on a High Performance Computing Cluster (HPCC) big data platform which allows the model to be used in real-world scenarios where large number of new discharge summaries exist.

DATA AND METHODS

COPD Data

Discharge summaries and clinical notes were extracted from 191,205 hospital admissions and readmissions between 2009 and 2015 from various hospitals across the region in which the authors work. The data was gathered from a variety of care centers including but not limited to Skilled nursing facility, rehabilitation centers, independent hospitals, hospital associations, nursing homes and long-term care facilities. A variety of datasets were utilized in order to capture different styles and forms of written discharge summaries. The data was labelled with a Boolean variable (readmit) showing readmission status for all the COPD related readmissions. This study was conducted using de-identified patient health records including discharge summaries and readmit status, taken from multiple hospital EHR systems. This study and the dataset used in this study were approved by the authors’ Institutional Review Board.

Data preprocessing was used to remove special characteristics and stop words. The extracted discharge summaries were used to generate a bag-of-words from all the discharge summaries, and then a count vector was created by applying minimum document frequency (Min-df) of 0.05 and Maximum document frequency (Max-df) of 1.0. This allowed removing of corpus specific stop words and removing words that appeared too infrequently. A total of 391 distinct features were extracted from the dataset. Table 1 below shows a sample of categories of different extracted features.

Category	Count	Sample Features
Encounter Reason	58	Headache, Infection, lymphadenopathy
Location	44	Abdomen, kidney, chest
Medications	27	Creatinine, Lasix, Albuterol
Vitals	22	Vitamin, Weight, Glucose
Health History	8	Alcohol, Allergy, Smoking
Demographics	2	Female, male

Table 1: Sample Categories and Extracted Features

Model Improvement Methods

The Data Preprocessing component included two main subcomponents, Python Engine and Sampling and Feature Selection using WEKA. The Python Engine was used to create the bag-of-words for all the words in the entire discharge summary dataset. On the extracted bag-of-words, scikit-learn Python package was used to create a count vector for all words with the count of number of times a word appears in a specific instance of the discharge summary. The dataset was further filtered to reduce dimensionality and remove words that appear too infrequently and increase false positives. The filtering was done by limiting and removing the words with document frequency lower than 0.05 and resulted in a total of 391 features with a total of 191,205 instances as shown in Table 2. The dataset is also highly imbalanced with only 4.4% instances of positive class in the original dataset.

The filtered dataset was still highly imbalanced and with high dimensionality. This was resolved by implementing the Random Under sampling technique and ReliefF Feature Selection technique based on the discussions in Literature Review section. The Random Under-sampling was implemented using WEKA and three different under sampled datasets were created by changing the distribution spread between majority (negative) and minority (positive) cases or classes. Table 2 shows how the distribution spread was used to Randomly Under-sample the majority class and reduce imbalance in the dataset.

Dataset	# of Instances	# Positive Class	# Negative Class	%Positive	# Features
No Sampling	191,205	8204	183,001	4.4%	391
RUS (50:50)	16,408	8,204	8,204	50%	391
RUS (65:35)	23,438	8,204	15,234	35%	391
RUS (80:20)	41,020	8,204	32,816	20%	391

Table 2 Dataset description showing High Imbalance improvement with Random Under-Sampling (RUS)

For reducing the dimensionality of the data, ReliefF feature selection technique was implemented in WEKA on the four datasets (sampled and non-sampled). This technique calculates a feature score for all features based on the determination of feature value differences between nearest neighbor instance pairs. If the observed difference is between feature pairs of same class, the score decreases, but if feature pairs are from different classes, the score increases. The feature score can then be applied to rank and select top scoring features for feature selection. In this experiment, we extracted three different feature set sizes from original dataset and three datasets were extracted for each sample. For each dataset, we extracted 300 features and also used all the features, and generated multiple datasets including one dataset with no sampling and no Feature selection as shown in Table 3 in the next section.

DISCUSSION AND CONCLUSIONS

The experiments performed in this study used area under ROC (Receiver Operating Characteristic) curve, recall, and precision as performance metrics. The AUC denotes a trade-off between sensitivity and fall-out rate. The higher the value of AUC, better the classification performance. AUC is an important classification performance indicator in healthcare as it denotes higher TP (true positive) recall and lower FP (false positive) recall. This experiment also uses recall which denotes what proportion of actual positives was identified correctly [TP/(TP+FN)], where FN is false negative. This determines the positively identified actual readmission. Precision denotes what proportion of positive identifications was actually correct. This determines the correct rate of positively predicted values [TP/(TP+FP)].

Table 3 below shows mean values of all performance metrics across various distribution sizes. Based on this, it shows that the sampled datasets performed better than the base imbalanced dataset on which RUS was not implemented. Among the three sampled datasets, the dataset with distribution spread of 65:35, which is also the median distribution spread selected under this study, performed the best.

Distribution Spread using RUS	Mean AUC	Mean Precision	Mean Recall
ALL	0.54	0.90	0.50
50:50	0.56	0.71	0.49
65:35	0.56	0.88	0.60
80:20	0.54	0.85	0.60

Table 3: Mean Performance Metrics across various Distribution Spreads

Table 4 below shows the implementation of feature selection has evidently improved classification performance. The dataset with maximum number of features selected has shown highest mean performance.

Feature Selection	Mean AUC	Mean Precision	Mean Recall
No	0.54	0.90	0.50
ReliefF-300	0.57	0.91	0.59

Table 4: Mean Performance Metrics for Feature set size of 300

Distribution Spread using RUS	Feature Selection	Number of Features	AUC	Precision	Recall
ALL	No	ALL	0.54	0.9	0.5
50:50	No	ALL	0.56	0.89	0.46
50:50	ReliefF	300	0.6	0.91	0.64
65:35	No	ALL	0.54	0.89	0.48
65:35	ReliefF	300	0.58	0.95	0.6
80:20	No	ALL	0.55	0.89	0.45
80:20	ReliefF	300	0.54	0.88	0.53

Table 5: Performance metrics for different Distribution samples and feature sets

Table 5 above shows the results for each sampled dataset in terms of AUC, precision, and recall. Recall and precision are the most important measure as recall shows the ability of this model to identify all the relevant cases within a dataset and precision, since it identifies the proportion of data points which are flagged as relevant, as actually being relevant. The dataset with even distribution and 300 extracted features shows best precision value of 0.91 and the recall value 0.64 which is higher than the other datasets. This is a positive indicator showing that data with even distribution and feature selection implemented shows the best performance for our COPD data.

The results show that data sampling and feature selection have positively impacted the classification performance. This paper presents an empirical study of supervised classification of COPD-related hospital readmission prediction models when the dataset is highly imbalanced and has high dimensionality. We address the issues of negative impact of imbalanced and high dimensional data on classification performance. This study, to best of our knowledge, is the first to use sampling and feature selection technique, and achieve positive classification performance for COPD readmission data. Future work will include implementing the improved model in a high performing computing cluster, and develop a real time clinical decision support system.

REFERENCES

Bingwei L, Erik B, Yu C, Dan S and Genshe C. Scalable sentiment classification for Big Data analysis using Naïve Bayes Classifier. 2013 IEEE International Conference on Big Data, Silicon Valley, CA, 2013, pp. 99-104. doi: 10.1109/BigData.2013.6691740

Carolyn MC. Commentary: reducing hospital readmissions: aligning financial and quality incentives. Am J Med Qual . 2012 ; 27 (5): 441 3 .

Danning H, Simon CM, Anthony NK, Susan H. Mining high-dimensional administrative claims data to predict early hospital readmissions. Journal of the American Medical Informatics Association. 2014; 21 (2):2729. https://doi.org/10.1136/amiajnl-2013-002151 PMID: 24076748

Gao K., T. M. Khoshgoftaar and A. Napolitano, "Impact of Data Sampling on Stability of Feature Selection for Software Measurement Data," 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence, Boca Raton, FL, 2011, pp. 1004-1011. doi: 10.1109/ICTAI.2011.172

Isaac S., Saeede A, Kai Y. A predictive analytics approach to reducing 30- day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD. Health care management science. 2015; 18(1):1934. https://doi.org/10.1007/s10729-014-9278-y PMID: 24792081

Joffrey L., T. M. Khoshgoftaar, Richard B. A survey on addressing high-class imbalance in big data. J Big Data (2018) 5: 42. https://doi.org/10.1186/s40537-018-0151-6

Joseph F, Jonathan M, Joseph L. A comparison of models for predicting early hospital readmissions. Journal of biomedical informatics. 2015; 56:22938. https://doi.org/10.1016/j.jbi.2015.05.016 PMID: 26044081

Khoshgoftaar T.M., A. Fazelpour, D. J. Dittman and A. Napolitano, "Effects of the Use of Boosting on Classification Performance of Imbalanced Bioinformatics Datasets," 2014 IEEE International Conference on Bioinformatics and Bioengineering, Boca Raton, FL, 2014, pp. 420-426. doi: 10.1109/BIBE.2014.68

Khoshgoftaar T.M., H. Wang and Q. Liang, "Stability and Classification Performance of Feature Selection Techniques," Machine Learning and Applications, Fourth International Conference on (ICMLA), Honolulu, Hawaii USA, 2011, pp. 151-156. doi:10.1109/ICMLA.2011.133

Kira K., L.A. Rendell, The feature selection problem: traditional methods and a new algorithm, AAAI, vol. 2, 1992, pp. 129–134.

Kononenko I., Estimating attributes: analysis and extensions of relief, European Conference on Machine Learning, Springer, 1994, pp. 171–182.

Lusa L. and R. Blagus, "The Class-Imbalance Problem for High-Dimensional Class Prediction," 2012 11th International Conference on Machine Learning and Applications, Boca Raton, FL, 2012, pp. 123-126. doi: 10.1109/ICMLA.2012.223

Surya PB, J Michael W. et al. Results of a Medicare Bundled Payments for Care Improvement Initiative for Chronic Obstructive Pulmonary Disease Readmissions. Ann Am Thorac Soc. 2017 May;14(5):643-648. doi: 10.1513/AnnalsATS.201610-775BC.

Shanab A.A., T. M. Khoshgoftaar, R. Wald and J. Van Hulse, "Comparison of approaches to alleviate problems with high-dimensional and class-imbalanced data," 2011 IEEE International Conference on Information Reuse & Integration, Las Vegas, NV, 2011, pp. 234-239. doi: 10.1109/IRI.2011.6009552

Soltani S., J. Sadri and H. A. Torshizi, "Feature selection and Ensemble Hierarchical Cluster-based Under-sampling approach for extremely imbalanced datasets: Application to gene classification," 2011 1st International eConference on Computer and Knowledge Engineering (ICCKE), Mashhad, 2011, pp. 166-171. doi: 10.1109/ICCKE.2011.6413345

Sandhan T. and J. Y. Choi, "Handling Imbalanced Datasets by Partially Guided Hybrid Sampling for Pattern Recognition," 2014 22nd International Conference on Pattern Recognition, Stockholm, 2014, pp. 1449-1453. doi: 10.1109/ICPR.2014.258

Seiffert C., T. M. Khoshgoftaar and J. Van Hulse, "Improving Software-Quality Predictions With Data Sampling and Boosting," in IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 39, no. 6, pp. 1283-1294, Nov. 2009. doi: 10.1109/TSMCA.2009.2027131

Taşkın G., H. Kaya and L. Bruzzone, "Feature Selection Based on High Dimensional Model Representation for Hyperspectral Images," in IEEE Transactions on Image Processing, vol. 26, no. 6, pp. 2918-2928, June 2017. doi: 10.1109/TIP.2017.2687128

Van Hulse J., T. M. Khoshgoftaar and A. Napolitano, "A comparative evaluation of feature ranking methods for high dimensional bioinformatics data," 2011 IEEE International Conference on Information Reuse & Integration, Las Vegas, NV, 2011, pp. 315-320. doi: 10.1109/IRI.2011.6009566

Visalakshi S. and V. Radha, "A literature review of feature selection techniques and applications: Review of feature selection in data mining," 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, 2014, pp. 1-6. doi: 10.1109/ICCIC.2014.7238499

Xiaoqian J, Aziz AB, Robert E, Kim J , Lucila O. A patient-driven adaptive prediction technique to improve personalized risk estimation for clinical decision support. J Am Med Inform Assoc. 2012 ; 19 (e1): e36 42

Wang Z. et al., "Application of ReliefF algorithm to selecting feature sets for classification of high resolution remote sensing image," *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, 2016, pp. 755-758.
doi: 10.1109/IGARSS.2016.7729190

DECISION SCIENCES INSTITUTE

Project-based service-learning integration in data analytics and statistics courses

Gokhan Egilmez, Ph.D.
Department of Mechanical, Industrial, and Biomedical Engineering,
University of New Haven
21 Ruden St. Room: 204, West Haven, CT 06516
Tel: +1 (203) 479-4196 Email: GEgilmez@NewHaven.Edu

ABSTRACT

In this paper, a service-learning integration is introduced to a newly developed Data Analytics and an existing Statistics courses. Both courses were taught in graduate levels in three terms (Spring, 2016, 2017 and 2018). The service-learning integration was made possible with local collaborations with 3 non-profit organizations. A total of 19 projects were completed by 60 graduate students (study sample). The impacts of SLPs on students learning, educational growth, civic engagement, community awareness, personal and professional growth were surveyed with a response rate of approximately 75%. Majority of the survey categories were found to be positively impacted by the SLPs, which was also supported with instructor assessment and course evaluations.

KEYWORDS: Service learning, Learning effectiveness, Statistics, Data Analytics, and Curriculum development

INTRODUCTION

Service learning (SL) is a pedagogical approach used in higher education and K12 levels which focuses on providing students a practical exposure to preferably local or global community issues where they could utilize the knowledge that they acquired in the course towards community contexts (Bringle & Hatcher, 1996). SL has long time been used and implemented in various higher education settings in various majors including social work (Lalayants, 2012), anthropology (Fletcher & Cambre, 2016), mechanical engineering (Tsang, Haneghan, Johnson, Newman, & Eck, 2001), supply chain management (Tsang et al., 2001) (Schoenherr, 2015), humanitarian logistics (Tsang et al., 2001), statistics (Anderson & Sungur, 1999; Lalayants, 2012), and at the institutional levels (Bennett, Sunderland, Bartleet, & Power, 2016).

SL aims to establish a mutually beneficial environment among the student, instructor and the non-profit. The students are expected to have practical exposure to the local community issues which fosters civic engagement as well as utilization of the theoretical knowledge in non-structured problem settings (Schoenherr, 2015). Therefore, it is a demanding educational development and growth process both from instructors' and students' perspective. The instructor needs to identify a SL partner from the local community and establish a framework which enables the local non-profit to receive technical and professional domain-specific support (e.g. statistics, procurement, staff planning, project management, etc.) from the higher education institution. In the majority of regions, local nonprofits lack the technical resources in terms of personnel (engineers, statisticians, managers, business and marketing people, etc.), budget, hardware, logistics, software, etc. to be able to carry out such projects, and in an ideal situation, the instructor and students team up with the nonprofit to fill this gap. SL can be implemented in a course in numerous ways depending on the context and scope of SL activity. For instance, students could be assigned a week-long HW or mini project with reflection after an pre-arranged visit to SL facility where they could provide assistance and then are asked for reflection on their learning. Furthermore, SL could be implemented as a semester long project which requires more intense collaboration among the students, instructor, and SL partner. This could also be termed as Project-based Service Learning (PBSL) since SL is integrated in a project-based learning structure. The students are expected to put more efforts in an organized way as a team, and are expected to be more pro-active in terms of defining the problem and working on a potential analytical solution unless the problem and data are given to them in advance by the instructor and/or the SL partner (Schoenherr, 2015).

LITERATURE REVIEW

Statistics and business analytics are among the most convenient courses where SL could be applied effectively since many community organizations have data in need of summary and analysis (Anderson & Sungur, 1999). As a matter of fact, both of the courses include learning outcomes that generally include the gaining practical knowledge of collecting, preparing, cleaning, and analyzing data from various perspectives and for various objectives. The state of art indicates that SL is beneficial to the students in many ways including students' engagement, willingness to explore nuances, and understanding of the material, enjoyment, social responsibility, and personal growth (Root & Thorme, 2002; Nordmoe, 2007; Schoenherr, 2015; Hiedemann & Jones, 2010). In contrast to many benefits, Hiedemann and Jones found that there was no statistical difference between integrating SL vs. using case study in statistics courses in terms of "mastery of course content" and " appreciation of the relevance of statistics" (Hiedemann & Jones, 2010). Even though there are many advantages of integrating SL in statistics course, these aforementioned benefits purely relies on successful integration of SL in a project-based environment where instructor, students, and the non-profit partner receives the utmost benefit from their sides (Schoenherr, 2015). Compared to statistics course, business analytics gained less attention in terms of SL in the literature. Even though previous works addressed the advantages and challenges, and effectiveness of SL in undergraduate course settings, it is important to address graduate level courses. In addition, it is also beneficial to compare statistics and business analytics courses from SL effectiveness perspective. In this study, a PBSL framework is introduced and students' perceptions of reaching learning outcomes, and other impacts of SL from civic engagement and professional growth perspectives are studied. The in investigation was focused on application of PBSL in statistics and business analytics courses that were taught at the graduate level at a mid-size higher education institution is conducted. The investigation covered both quantitative (empirical surveys, tertial data, instructor's assessment) and qualitative feedback, and results are analyzed graphically and statistically.

METHODS

At the authors' institution, service learning (SL) implementation requires an institutional process where the faculty who are interested in implementing SL are required to place an application fo SL implementation. If approved, the letter S is added to the end of the corresponding course code ID, which is also termed as S-designation. However, prior to the SL application, the instructor needs to attend the service-learning partners' speed dating event, which takes place every semester and the instructor needs to attend either the prior year or the prior semester to the term the course is to be taught. The speed dating event gathers over 50 local non-profit organizations (potential SL partners) and faculty who are interested in integrating SL into an undergraduate and/or graduate level courses. Following step-by-step procedure is employed during the deployment of PBSL in statistics and business analytics courses at the authors' institution. The courses are taught in Spring 2016, Spring 2017, and Spring, 2018 semesters to a total of 60 students, and a total of 19 Service-learning projects (SLPs) were completed.

Service-learning course development process

1. Identification of a Service-learning (SL) Partner: SL partner was identified during the interviews conducted at the Community Partner Speed Dating event hosted by the University's Service-Learning Center in the previous semesters.
2. S-designation application: S-Designation application does not require an assigned SL partner but encourages during the application phase, which takes place in the end of the 2nd month of the semester prior to the scheduled semester of the course. The S-designation application requires approval of the department chair and the program coordinator as well as the Institution's Service-Learning Center Director.
3. Service-learning integration into the Course Syllabus: The integration of SL into the courses was made as a project-based learning framework. The instructor arranged multiple online and in-person meetings with the SL partner to frame the scope and objectives of the Service-Learning Projects (SLPs).
4. Introducing SL partner to the students: The introduction of the SL partner was made in the first week and at least one representative from the SL partner was present and gave an overview about the partnering organization.
5. SL project team up: Students were asked to team up to groups whose size ranges from 2 to 4 students depending on the enrollment size (e.g. 12 students were enrolled in Spring 2018 course, which required 2 students per team, whereas 24 students were enrolled in Spring 2017, which required 3-4 students per team). The group size decision was also affected by the number of SLPs arranged with the SL partner.
6. SLP meetings and visits
7. SLP presentations and reports
8. Assessment and wrap-up

Descriptions of the Deployed SLPs

Following section introduces the descriptions of SLP deployed in Spring 2016, 2017, and 2018 semesters in 3 distinct service-learning partners.

<u>SLPs of the Data Analytics taught in Spring 2016 and Spring 2017</u>

The term project will function as the service (project-based) learning activity which will require students' participation in meetings, individual and collective work with team mates and the organization officer on a predefined problem related to organization's mission.

Definition: Service learning is defined as "Service learning is a credit-bearing, educational, experience in which students participate in an organized service activity that meets identified community needs and reflects on the service activity in such a way as to gain further understanding of course content, a broader appreciation of the discipline, and an enhanced sense of civic responsibility (Bringle & Hatcher,1995).

Justification: The service-learning experience is highly related to the course because of the following reasons: 1) Students will be collecting data about a previously defined research question that is important for the regional, national, and/or global socio-economic and environmental sustainability-related development. 2) Various data analytics methods and model validation techniques will be heavily employed in the project. 3) All in all, students will be using the acquired knowledge in the class systematically for the project and the project tasks and meetings with mentors will enable them to develop a variety of professional skill sets and disseminate the findings of the projects with the specific community through online and onsite presentations.

Reflection: Reflection is critical part of service learning. Therefore, specific attention will be paid in this course towards making sure that students are reflecting the abovementioned knowledge and skills gained this course at the end and they will utilize from now on in their personal and professional life. To do so, there will be frequent meetings between project groups and mentors where students will be asked about how they can utilize the knowledge and skills gained for the projects. So, the interactive meetings will be playing a critical role for initiating the reflection. Discussions, presentations, Q&A sessions, the project report, and the online published content will be predominantly used as the strategies for reflection.

Length of Service: The students will spend about 4 weeks and 5 hours per week on the service-learning project. Typical project steps and timeline are also included in the syllabus.

SLP Summary: In the service-learning project, students will work on cleaning, preparing the previously funded projects by Service-Learning Partner 1& 2. Service-Learning Partner 1&2 expects the project groups to figure out ways to assess their impacts on the community. They will be providing their project information and report files, which will be used by students to carry out the SLP. SLP will take about 6 weeks, in the beginning phase of the semester. A total of 25 hours will be required to conduct the literature survey about the United Nation's Climate Change and Environmental Sustainability goals and local community guides. Then, students will start creating a database of projects, and developing a method to assess the impact of the agency on the community by using data analytics methods.

SLP of the Applied Statistics, Quality, and Engineering Management taught in Spring 2018:

Assessing Community Well-being in the State of Connecticut:

Community well-being is a vital phenomenon that needs to be studied and assessed frequently to inform the policy makers and politicians about the area of improvements within local communities. In this context, Service-Learning Partner 3, a local nonprofit organization has conducted a community wellbeing survey for the last decade and recently developed a visual assessment tool. The 2016 dataset consists of 12 main variables 189 rows, which are aggregated based on state house district and chamber. The dataset can be accessed with the link given on BB. Following are main category variables given in a tabular format:

There were 12 students in class, which made up 6 groups of 2. Each group was asked to pick 1 of the main variable categories including: Community Vitality, Economic Security, Community Health: Adults, Households, Education, Income

Assessment

The total sample size of the study was 60 students (Spring 2018: 12, Spring 2017 and Spring 2016: 24 students each). The assessment focused on instructors' assessment, student evaluations of the courses, and an empirical survey that focuses on assessing the students learning effectiveness in SLP.

Empirical survey

Since the course evaluations conduct a general assessment of the course, which indirectly assesses SL, it is important to survey students to have them provide quantitative and qualitative feedback about their learning experience specific to the service-learning project. To do so, an empirical survey is developed and deployed to the students. The survey consisted of 5 sections:

1) Demographics (Self-identification, Race/ethnic heritage, Job status, and Major)
2) Community needs addressed with the SLP
3) Impact of SL on Learning and Professional Growth
4) Impact of SL on Community Awareness and Civic Engagement
5) Impact of SL on Personal and Professional Growth
6) Qualitative feedback

A Likert Scale was used to organize and standardize students' responses, which consisted of the following levels: 5: Strongly agree (SA), 4: Agree (A), 3: Neutral (N), 2: Disagree (D), 1: Strongly Disagree (SD).

Instructor assessment

The instructor assessment consisted of Service-learning Project Report, Service-learning Project Presentation, pre and post-surveys (Deployed in Data Analytics course since it was a newly developed special-topics course).

Course evaluations

Student evaluations are also used as part of assessment. Following questions are surveyed to the students enrolled in the course at the end of the enrollment term. Students were asked to provide answer over the same Likert scale used in empirical survey.

1. The objectives of the course were clear
2. The course materials (as listed on the syllabus) contributed to my learning.
3. Assignments and other graded activities gave me an opportunity to demonstrate what I learned.
4. The grading system for the course was clear.
5. The instructor was prepared for each class.
6. The instructor's presentations were understandable.
7. The instructor provided helpful feedback.
8. The instructor used class time effectively.
9. My interest in the subject matter was enhanced by the instructor's enthusiasm.
10. The instructor raised questions or problems that encouraged me to think critically.
11. The instructor explained the relevance of the subject matter.
12. The instructor established a positive learning environment.
13. The instructor was accessible outside of class (for example: held office hours, communicated via email, or offered to meet via video conferencing).
14. Overall, I was satisfied with the educational experience provided by the instructor.

RESULTS

Results are organized in parallel with the methods section. Firstly, the results of empirical survey are provided. Secondly, results of instructor's assessment are provided. Thirdly, course evaluations are shared.

Results of empirical survey

Results of the survey is explained in twofold: quantitative and qualitative feedback. In the quantitative part of the survey, demographics (self-identification, race/ethnic heritage), job status and major are given in figures 1,2, 3, and 4, respectively. Majority of the students were male and from Caucasian or Asian backgrounds. These are quite typical of engineering schools in terms of self-identification and major since majority of the graduate students are recruited from Asian countries at the Institution (see fig. 1,2). About half of the students were not actively working during the time of the assessments and a quarter of them was found to be full time employee (fig. 3). Furthermore, table 1 shows the community needs addressed with the SLPs. In this category, students could pick multiple sections as some SLPs address more than one community need. It was found that education was found to be the most frequently addressed community need with a quarter of response frequency.

<u>Quantitative assessment</u>

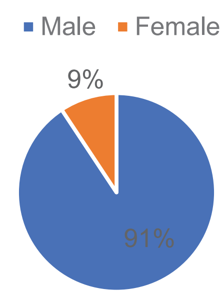


Figure 1. Self-identification (Total response: 43, Missing: 17)

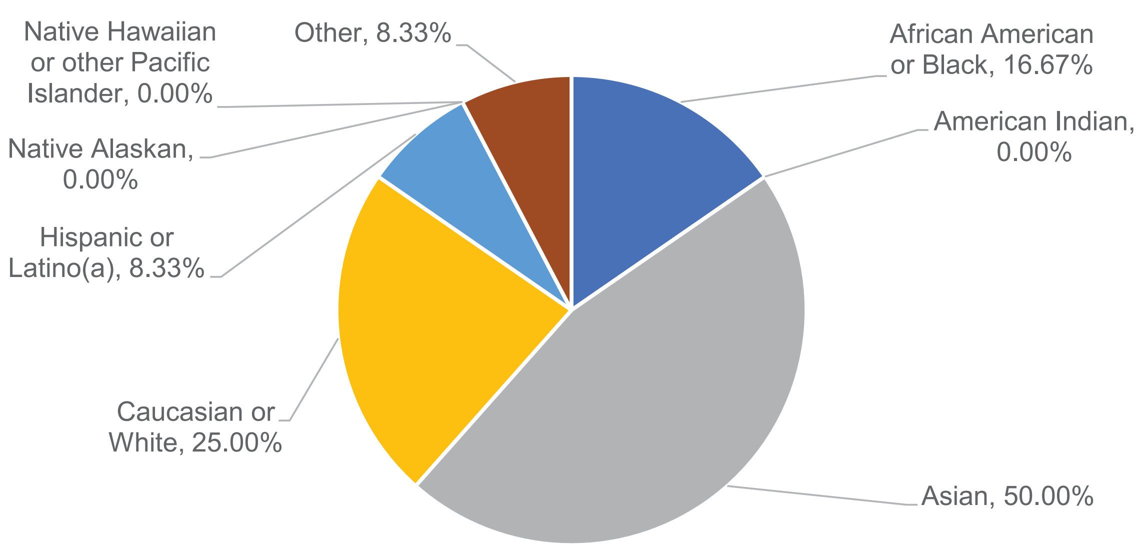


Figure 2. Race/ethnic heritage (Total response: 45, Missing: 15)

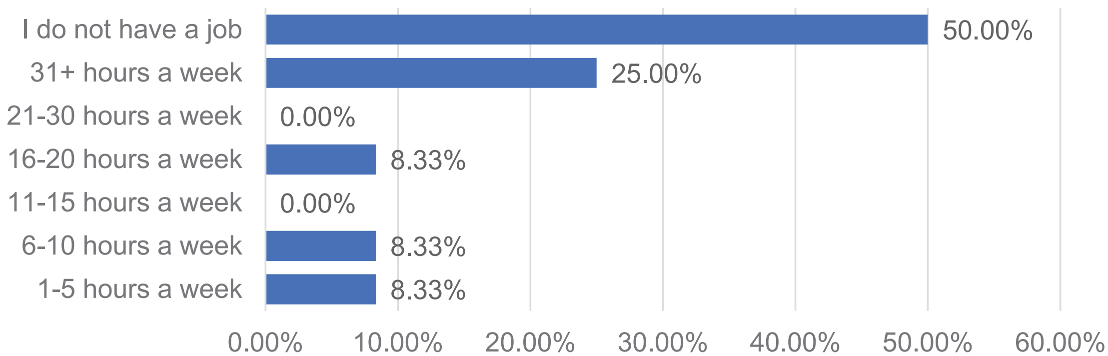


Figure 3. I have a paying job that requires me to work___ (Total response: 42, Missing: 18)

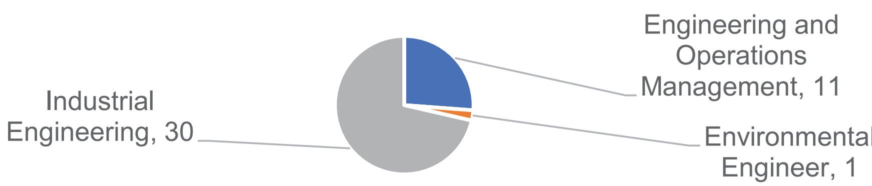


Figure 4. Major (Total response: 42, Missing: 18)

Table 1. Community Needs Addressed with the SL Experience

Community Needs Category	Responses	
Education	23.2%	23
Other	12.1%	12
Employment	9.1%	9
Youth issues	9.1%	9
Communication/Educational Outreach	8.1%	8
Financial Literacy	6.1%	6
Housing	5.1%	5
Infrastructure	5.1%	5
Environment	4.0%	4
Community Well-being	3.0%	3
Income & Economic Security	3.0%	3
Immigration	2.0%	2
Info Systems	2.0%	2
Victim Rights/ Services	2.0%	2
Community Health	2.0%	2
Disability Issues	1.0%	1
Elderly Care	1.0%	1
Mental Health	1.0%	1
Self-Image	1.0%	1
Total	100%	99

The latter section of the survey was focused on students' perception of the SL experience on a variety of categories including learning and educational growth (Table 2), community awareness and civic engagement (Table 3), personal and professional growth (Table 4). In each of the tables (2,3, and 4), the share of responses on Likert scale categories (5: Strongly agree (SA), 4: Agree (A), 3: Neutral (N), 2: Disagree (D), 1: Strongly Disagree (SD)) as well as the mean response over 5 (μ). It was found that the learning and educational growth was positively impacted by the SLPs as all of the category means were found to be over 4 in Table 2. And, the highest mean score was obtained from the question, entitled "Service-learning should be implemented into more classes at the Institution", which indicates that students were in favor of SL integration in other courses after the experience.

Table 2. Impact of SL on Learning and Educational Growth

Question	SA	A	N	D	SD	μ
The community work helped me see how course material can be used in everyday life.	39.5%	53.5%	4.7%	2.3%	0.0%	4.3
The community work helped me to better understand the material from my lectures and readings.	48.8%	37.2%	7.0%	7.0%	0.0%	4.3
My community experience was directly linked to what I learned in the classroom.	37.2%	46.5%	9.3%	7.0%	0.0%	4.1
Participation in service-learning made me take more responsibility for my own learning.	53.5%	34.9%	4.7%	4.7%	2.3%	4.3
Service-learning should be implemented into more classes at the Institution	52.4%	33.3%	11.9%	4.8%	0.0%	4.4
I would like to enroll in additional service-learning courses offered through my degree program.	35.7%	40.5%	16.7%	2.4%	4.8%	4.0

Results of community awareness and civic engagement questions were provided in Table 3. Except the initial question about students' previous service-learning experience, all categories results indicate a positive impact due to the mean scores that are greater than 4. The students strongly indicate that they do have a responsibility to serve the community with a mean score of 4.5, which is a remarkable finding since about half of them did not indicate a service-learning experience previously in question 1. In general, SL experience positively impacted and contributed to the civic engagement development of the students and their awareness about community issues, which are quite essential growths expected from especially international graduate students in terms of their adaptation and sustained success at the Institution.

Table 3. Impact of SL on Community Awareness and Civic Engagement

Question	SA	A	N	D	SD	μ
I was already volunteering before this service-learning experience.	23.3%	32.6%	20.9%	16.3%	7.0%	3.5
The community work showed me how I can become more involved in my community.	41.5%	46.3%	4.9%	4.9%	2.4%	4.2
The community work I did benefit the community.	30.2%	37.2%	32.6%	0.0%	0.0%	4.0
Service-learning helped me to become more aware of community needs.	46.3%	39.0%	9.8%	2.4%	2.4%	4.2
I will continue to be involved in the community as a result of this experience.	38.1%	45.2%	11.9%	4.8%	0.0%	4.2
I have a responsibility to serve the community.	58.5%	31.7%	7.3%	2.4%	0.0%	4.5

Table 4 depicts the survey results related to personal and professional growth. Except two categories ("The community work helped me clarify and/or confirm which major I will pursue" and "The community work helped me in defining my career choice/specialization"), all categories' results positively favor the SL experience on personal and professional growth. The highest response mean was obtained from enhancing leadership skills, which is a very critical skill that graduate students are expected to develop and continuously add on during the graduate study.

Table 4. Impact of SL on Personal and Professional Growth

	SA	A	N	D	SD	μ
The community work helped me to define my personal strengths and weaknesses.	47.6%	38.1%	7.1%	2.4%	4.8%	4.2
The community work helped me clarify and/or confirm which major I will pursue.	21.4%	28.6%	40.5%	4.8%	4.8%	3.6
The community work helped me to develop my problem-solving skills.	34.1%	46.3%	12.2%	4.9%	2.4%	4.0
The community work helped me to enhance my leadership skills.	47.6%	40.5%	7.1%	2.4%	2.4%	4.3
The community work helped me in defining my career choice/specialization.	26.2%	40.5%	28.6%	2.4%	4.8%	3.9
The community work helped me to enhance my ability to communicate my ideas in a real-world context.	42.9%	42.9%	7.1%	4.8%	2.4%	4.2
The community work I accomplished in this course has made me more marketable in my chosen profession.	35.7%	45.2%	14.3%	2.4%	4.8%	4.1

Qualitative assessment

At the end of the empirical survey, students were asked to share any other comments that they have about their service-learning experience. Total of 9 comments were received. Six of the comments were positive and indicating that the student was appreciative of SL experience. One student indicated a motivation to carry out similar project in her/his home country. Another student indicated that SLP having a social aspect made it a great experience as engineer. Two of the students indicated that the SLP experience make them more aware about environmental sustainability and sustainable development. The three students who provided negative feedback about the SLP experience. One stressed that the topic was not directly related to engineering field even though the experience was rewarding. Another student stressed that having 10+ years work experience in statistics, the SLP did not enhance his/her knowledge in statistics. Another student stated the he/she did not get to use the knowledge acquired in course towards the SLP.

Instructor Assessment

The instructor assessment was focused on SLP report, SLP presentation, and the weighted total grade. The assessment results are provided in figure 5. Figure 5a indicates the SLP report for the 3 terms the SL was integrated. The first year that SLP was integrated, the course was also a special topics course (Data Analytics, in Spring 2016) and it was offered for the first time. As shown in the SLP description, the report structure asked from students was quite demanding and the project topics were not given with sufficient level of detail, since the instructor and SL partner purposely pushed students to define the problem, collect data, identify an appropriate analytics method and answer a few pre-determined research questions. This somewhat affected the students produce not very good quality reports and projects even though the experience was found to be rewarding by majority as shown in the previous section. The second year, all student teams were given the same problem with more detailed directions and raw data in agreement with the SL partner's interest, which yielded better reports and better-quality projects. However, this had the drawback of lack of diversity in project focus areas. The third year, a well-being survey data was provided from the SL partner and students worked on the same data with specific focus areas, which provided them room for creativity as well as medium-level of directions and guidelines. This yielded highest project report grades. In terms of SLP presentations, the grades were somewhat in similar range as shown in figure 5b. In terms of weighed total grades, similar trend was observed compared to the SLP reports. Over the 3 years of implementation of SLP, student grades tend to increase in terms of both mean and the range from 2016 to 2018.

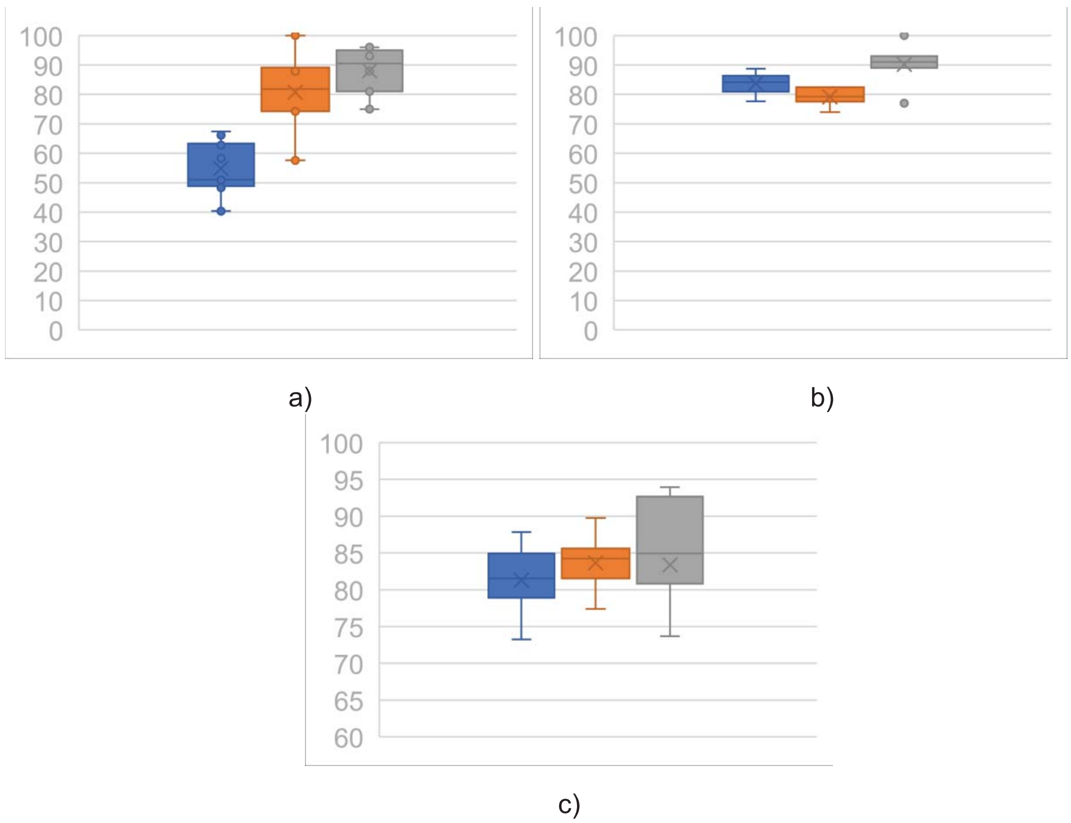


Figure 5. a) SLP Report, b) SLP Presentation (Peer-evaluation), c) Weighted Total – Course Grade (**Blue: Spring 2016**, **Orange: Spring 2017**, **Gray: Spring 2018**).

Course evaluations

In terms of the course evaluations (See figure 6), the least that could be said that the SLP did not negatively impact the course evaluations. This is important because integrating SLP into a course automatically brings challenges and complexities from both the instructor and the students' perspectives. Both stakeholders need to go out of their comfort zones to make this experience real and rewarding for each other. In all of the evaluation categories, the course was found to be positively impacting the students' learning as the mean scores were greater than 4 out of 5.

	Spring 2016		Spring 2017		Spring 2018	
Question	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
The objectives of the course were clear.	4.54	0.78	4.2	0.8	4.3	0.7
The course materials (as listed on the syllabus) contributed to my learning.	4.71	0.55	4.4	0.7	4.2	1.2
Assignments and other graded activities gave me an opportunity to demonstrate what I learned.	4.75	0.53	4.4	0.7	4	0.9
The grading system for the course was clear.	4.63	0.65	4.4	0.7	4.3	0.7
The instructor was prepared for each class.	4.67	0.56	4.3	0.7	4.4	0.5
The instructor's presentations were understandable.	4.62	0.72	4.3	0.7	4.3	0.5
The instructor provided helpful feedback.	4.67	0.56	4.4	0.7	4.4	0.7
The instructor used class time effectively.	4.79	0.51	4.3	0.8	4.6	0.7
My interest in the subject matter was enhanced by the instructor's enthusiasm.	4.46	0.72	4.3	0.8	4.2	0.6
The instructor raised questions or problems that encouraged me to think critically.	4.5	0.59	4.2	0.9	4.2	0.9
The instructor explained the relevance of the subject matter.	4.5	0.78	4.3	0.8	4.1	1
The instructor established a positive learning environment.	4.71	0.46	4.5	0.7	4.5	0.5
The instructor was accessible outside of class.	4.79	0.41	4.5	0.7	4.2	1
Overall, I was satisfied with the educational experience provided by the instructor.	4.71	0.69	4.4	0.7	4.2	0.8
Overall	4.63	0.62	4.4	0.7	4.3	0.8

Figure 6. Course Evaluations

DISCUSSION

Integrating service learning into a course could be very enjoyable or very frustrating experience from both the instructor's and the students' perspectives. The main reason is that there is always a degree of uncertainty since both stakeholders get to work with a 3rd stakeholder (the SL partner organization), which expects them to produce a project that will somewhat be beneficial for their operational or strategic decision-making processes. The experience is primarily dependent on the instructor's continuous enthusiasm to make the SLPs carried out throughout the semester. It generally demands a continuous and rigorous Project Management effort to organize all of the stakeholders to work for a common set of goals. This also becomes more challenging if student teams get to work on diverse set of projects.
One of the most impactful benefits of the SL experience is that students are exposed to community issues. This is specifically important for international students because such experience improves and fastens their adaptation to the country through being involved in the local community. This was also substaintially supported by the student survey responses. Their civic engagement and awareness about local community problems are heightened at the first minute of meeting with the SL partner organization. And, in SL course, this could go beyond the expectations of the SLP description in the course depending on the SL partner's mission and vision. Some student teams presentations went beyond the expectations and it was very rewarding to see them advocating and promoting solutions for community problems. This was also seconded by the 2nd SL partner representative as she was quite impressed with the presentations.

The lessons learned from this experience indicate that it is very crucial to find a SL partner who will support the students and instructor towards to SLP goals throughout the semester. If the SL partner did not take some of the heavy lifting in the Project Management, the instructor ends up being the solo go-to person, which cause the students' experience a limited exposure to the SL partner organization, through the instructor. This could be alleviated by producing a well-structured and detailed project charter with the SL partner prior to the semester. Furthermore, not all students will be willing to go out of their comfort zone and they would like to continue just doing the coursework and attending the lectures. In this regard, it is very important responsibility of instructor to enthusiastically promote the SLPs and invite the SL partner representative to the first lecture of the semester to have an effective kick off. In person meetings with the SL partners significantly improves the students' engagement with SLPs and with the course too.

CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

In this paper, SLP integration into Statistics and Data Analytics courses at the graduate level is introduced. Three distinct SL organizations were partnered with in Spring 2016, Spring 2017, and Spring 2018 semesters. The study sample included 60 students and 19 SLPs were completed in three terms. The SL integration process was explained as a roadmap with the DSI community. The impacts of SLPs on students learning, educational growth, civic engagement, community awareness, personal and professional growth were surveyed with a response rate of approximately 75%. Majority of the survey categories were found to be positively impacted by the SLPs, which was also supported with instructor assessment and course evaluations. This study has the following limitations. First of all, the newly proposed Data Analytics course was elective even though the Statistics course was a mandatory course for the students. Elective courses tend to create selection bias (Anderson & Sungur, 1999). The impact of selection bias could be minimized by relying on a sample who took mandatory courses, in this case Statistics course. The author(s) plan to continue integrating SL into Fall 2019 and Fall 2020 Statistics courses, which will yield an increased sample size with no selection bias, and this is left as future work. Moreover, a comparative analysis was left as future work with the other sections of the statistics course since a recent work did not find a statistically significant difference between SLPs and Case-study based approaches in terms of "mastery of course content" and " appreciation of the relevance of statistics" (Hiedemann & Jones, 2010). However, this study did not address the main benefits of SL, which are civic engagement and community awareness which are as much essential as the course learning outcomes. Therefore, the author(s) will collect more data in Fall 2019 and 2020 to conduct a comparative analysis, which is also left as a future work.

REFERENCES

Anderson, J. E., & Sungur, E. A. (1999). Community service statistics projects. *American Statistician*, *53*(2), 132–136. https://doi.org/10.1080/00031305.1999.10474446

Bennett, D., Sunderland, N., Bartleet, B. L., & Power, A. (2016). Implementing and sustaining higher education service-learning initiatives: Revisiting Young et al.'s organizational tactics. *Journal of Experiential Education*, *39*(2), 145–163. https://doi.org/10.1177/1053825916629987

Bringle, R. G., & Hatcher, J. A. (1996). Implementing Service Learning in Higher Education. *The Journal of Higher Education*, *67*(2), 221. https://doi.org/10.2307/2943981

Fletcher, C., & Cambre, C. (2016). Digital Storytelling and Implicated Scholarship in the Classroom. *Journal of Canadian Studies*, *43*(1), 109–130. https://doi.org/10.3138/jcs.43.1.109

Hiedemann, B., & Jones, S. M. (2010). Learning statistics at the farmers market? A comparison of academic service learning and case studies in an introductory statistics course. *Journal of Statistics Education*, *18*(3), 1–21. https://doi.org/10.1080/10691898.2010.11889584

Lalayants, M. (2012). Overcoming Graduate Students' Negative Perceptions of Statistics. *Journal of Teaching in Social Work*, *32*(4), 356–375. https://doi.org/10.1080/08841233.2012.705259

Nordmoe, E. D. (2007). Service-learning in introductory statistics at Kalamazoo College. *Journal of Statistics Education*, *15*(2). https://doi.org/10.1080/10691898.2007.11889466

Root, R., & Thorme, T. (2002). Community-Based Projects in Applied Statistics. *The American Statistician*, *55*(4), 326–331. https://doi.org/10.1198/000313001753272295

Schoenherr, T. (2015). Service-Learning in Supply Chain Management: Benefits, Challenges and Best Practices. *Decision Sciences Journal of Innovative Education*, *13*(1), 45–70. https://doi.org/10.1111/dsji.12052

Tsang, E., Haneghan, J., Johnson, B., Newman, E. J., & Eck, S. Van. (2001). A report on service-learning and engineering design: Service- learning's effect on student's learning engineering design in 'Introduction to Mechanical Engineering'*. *International Journal of Engineering Education*, *17*(1), 30–39. Retrieved from https://www.ijee.ie/articles/Vol17-1/Ijee1176.pdf

DECISION SCIENCES INSTITUTE

A Decision Analysis towards Mobile Financial Services Adoption and Sustainability in Togo: Structural Equation Modeling and TOPSIS Methodology

Komlan Gbongli
Institute of Finance and Accounting, Faculty of Economics, University of Miskolc, Hungary;
E-mail: pzkgbong@uni-miskolc.hu or samxp12@yahoo.fr

Levente Kovács
Institute of Finance and Accounting, Faculty of Economics, University of Miskolc, Hungary;
Secretary-General Hungarian Banking Association, Budapest, Hungary;
E-mail: kovacs.levente@uni-miskolc.hu or kovacs.levente@bankszovetseg.hu

ABSTRACT

Due to the low rate of mobile financial services (MFS) adoption, this research accesses the trust and risk factors at multidimensional level regarding MFS, using the integrated structural equation modeling-TOPSIS based on the survey data collected in Togo to offer new insights.
The results show particularly strong support for the dispositional trust and perceived privacy risk. Trust has a negative relationship with perceived risk, which finding was limited partly. For the MFS perspective, mobile money transfer (MMT) remains the core application used, followed by mobile payment (MP) and mobile banking (MB). The implications, limitations, and future research are provided.

KEYWORDS: Mobile financial services (MFS); Trust; Perceived risk; Structural equation modeling (SEM); Technique for order preference by similarity to ideal solution (TOPSIS)

1. INTRODUCTION

Innovations and technological expansion have emerged with significant advantages to the recent commercial market. Over the last few years, businesses have been redirecting their goals to making information system technology an essential part of their processes (Oliveira, Faria, Thomas, & Popovič, 2014). Therefore, more and more literature is diverted to the IS-related field (Agarwal & Prasad, 1999). The investigation of some existing studies which recommend integrating various theoretical models to understand the IT adoption has stressed that a comprehensive analysis in the context is required (Afshan, Sharif, Waseem, & Frooghi, 2018; Oliveira et al., 2014). From these perspectives, an increasing number of researchers are focused on the mobile financial services (MFS) considered as the development of IS domain (Abrahão, Moriguchi, & Andrade, 2016; Al-Jabri & Sohail, 2012; Gbongli, Dumor, & Kissi Mireku, 2016; Muñoz-Leiva, Climent-Climent, & Liébana-Cabanillas, 2017).

MFS refers to any financial transaction remotely conducted by the application of a mobile phone (e.g., smartphone or tablet) and mobile software (e.g., apps programs) either through banking service or network provider service (Gbongli, 2017; Yen & Wu, 2016). MFS providers allow their consumers the flexibility to access their financial services (access information inquiry, bill payment, and money transfers) anywhere and anytime via a mobile phone, to support and improve service relationships by investing lots of resources using wireless internet technology (Gartner, 2004; H. F. Lin, 2011; Scornavacca & Hoehle, 2007).

The studies of MFS that focused on e-money transfer fall into three major mobile technologies-related fields of study, primarily mobile banking (MB), mobile payment (MP), and mobile money transfer (MMT) (Gbongli, 2017). MB is among the latest in a sequence of new mobile technological wonders (Mohammadi, 2015). Therefore, it is expected to have a significant impact on the market (Safeena,Rahmath; Date,Hema;Kammani, Abdullaj; Hundewale, 2012). Payment today has now progressed to mobile devices (m-devices) identified as mobile financial services, particularly mobile payment (Ooi & Tan, 2016). M-payment denotes any payment in which a mobile device is used to initiate, authorize, and confirm a commercial transaction (Srivastava, Chandra, & Theng, 2010). Furthermore, mobile money has appeared as a significant innovation with a potential expansion to financial inclusion in developing countries in various ways (Gbongli, Xu, & Amedjonekou, 2019). It is, therefore, growing access to financial services for a large number of people, who are completely ignored by banks because of longer travel distances or lacking funds to meet the minimum deposit recommended to open a bank account (Jack, Ray, & Suri, 2013; Kikulwe, Fischer, & Qaim, 2014), low income population in developing countries (J. Anderson, 2010), insofar, as it has several advantages (Assadi，D. & Cudi，A, 2011; Chaix，Laetitia & Torre，Dominique, 2015; Gbongli et al., 2019). In addition to the advantages granted to certain persons and companies, there are also advantages at the national economy level, primarily in emerging economies such as Hungary. The use of increasingly more accommodating tools may incentivize the suppressed use of cash, parallel to which, the countability of economic performance with statistical instruments continues to improve; meanwhile tax payment discipline also improves and the total social cost of payments decreases, etc., that is, overall the economy begins to whiten, leading to improved competitiveness (MNB, 2019).

While tremendous benefits are associated with adopting MFS as opposed to traditional payment methods, such as physical exchanges notes, cheques, coins (J. Anderson, 2010), the adoption rate is far from full utilization in many developing countries. This is characteristically the situation of West African Countries and particularly Togo. Rendering to the Statista Portal (2016), the number of smartphone users worldwide is predicted to be over five billion marks in 2019. Approximately 67% of the Togolese population subscribed to the handphone in 2015, and the number of mobile internet users doubled between 2014 and 2015. However, the banking rate in Togo is less than 15% (Couchoro, 2016) and, the rate of consumer’s acceptance of mobile banking remains lower than expected, i.e., 1% (Financial Afrik, 2015). It is therefore coherent to merely guess that mobile money should offer a significant input to grow the rate of MFS usages. In reality, this is far from being the case. The experiences of more developed countries also suggest the same, not technological limitations were the primary obstacle of the extension of the innovative payment solutions (Divéki, Keszy-Harmath, & Helmeczi, 2010). Therefore, the motives for the success or failure together with the causes and reasons for mobile money acceptance, are still not understood sufficiently, which infers that the technology has not been extensively adopted. These trends recommend that significant growth opportunities remain, leading to predictions of potentially massive rises in the number of m-banking users. Furthermore, it enables a partial knowledge regarding the motivators and inhibitors that impact the acceptance of this mobile service, which supports the limited studies that have been conducted in this field (Hassan Hosseini, Fatemifar, & Rahimzadeh, 2015).

Understanding why it is worth it to select to use MFS can help in strategy development and allow businesses to effectively communicate benefits to their customers (Johnson, Kiser, Washington, & Torres, 2018; Liébana-Cabanillas, Sánchez-Fernández, & Muñoz-Leiva, 2014). Mobile financial service operators could increase their competitiveness if they were able to enhance their performance in meeting consumer demand. Therefore, they need to understand the requirements of MFS users and the relative weight of each factor that affects consumers' needs. One possible motive for the existence of a gap between these could be the perception of risk that limits consumers’ ability to make informed decisions to partake the advantage of MFS technology in Togo (Gbongli, Csordas, & Kissi Mireku, 2017). This is particularly true for emerging nations, mainly in an unstable country where the consideration of risk and loss of privacy in the security system played a crucial part in adopting IT (Gbongli, Peng, & Ackah, 2016). Users are concerned about whether service providers correctly gather, store, and use their financial and private information. Moreover, the studies in the past revealed that once there are risk issue concerns, the demand for trust becomes a necessity, since trust and risk are interrelated facets (Gbongli et al., 2017; Mayer, Davis, & Schoorman, 1995). Not only the developing countries facing the issue of e-business but also the reflection of the online risk has called for a considerable attention amongst developed countries like Hungary, particularly in 2014 when the case of fraud risk in electronic payment transactions ascended in Hungary (the case was discussed in the work of Kovács & David (2016) in detail).

Driven by studies towards the multiple dimensionalities for risk and trust and the central research on trust in contrast to risk in new IT perspective (David Gefen, Benbasat, & Pavlou, 2008), we suppose that initiating research into novel IT artifacts like this study could enlighten how trust and perceived risk might affect the ultimate adoption of innovative technologies in developing countries.

The goal of this study is to disclose mechanisms related to behavior associated with MFS adoption and sustainable development when decision making involves multiple criteria issues. One main research question is to understand how multi-dimensional trust and multi-faceted perceived risk perceptions affect a new emerging information technology such as MFS adoption at the individual level in an unstable country. Our approach differs from most prior studies that assess trust and risk perception of individual behavior. Indeed, most of the research that investigated the adoption and use of communicative technology has been done within highly-developed, stable, capitalist societies. Moreover, the majority of research assumes that individuals have freedom of speech, basic protection, and safety of their lives and business offered by the government. However, little has been known regarding the adoption of IT in emerging, and dynamic societies (Goodman, 2011; Marett, Pearson, Pearson, & Bergiel, 2015; Wells, Campbell, Valacich, & Featherman, 2010). The Fund for Peace identifies 31 countries around the globe that fall into a high-warning context based on different pressures they face that impact their fragilities, such as security apparatus, group grievance, human rights and the rule of law, deteriorating government, poverty, and social indicators. Therefore, we explore the fundamental trust and risk allied with MFS technology usage in high poverty.

The majority of prior research typically tests trust as a single construct (Alsaad, Mohamad, & Ismail, 2017; Gao & Waechter, 2017; K. C. Lee & Chung, 2009) or investigates trust constructs and risk dimensions disjointedly (Johnson et al., 2018; Lowry, 2008). In other words, how to effectively assess trust and risk concerns concurrently remains a black box. Drawing on research in Information Technology (Luo, Li, Zhang, & Shim, 2010; Park & Tussyadiah, 2017), we stress that multi-dimensional trust and perceived risk concepts may jointly play an integral part in individual behavior with regard to adopting a novel MFS, and it is of paramount importance for this to be investigated, particularly in developing countries such as Togo.

Furthermore, a plethora of research has been done in order to fully understand the factors that affect MFS adoption and their significance. However, most earlier research in this area has emphasized the general factors regarding the adoption of MFS, using explanatory statistical analysis as the research method (Aktepe, Ersöz, & Toklu, 2015; Alzahrani, Al-Karaghouli, & Weerakkody, 2018; Leida Chen, Meservy, & Gillenson, 2012; Chong, Chan, & Ooi, 2012). Although the beta coefficients gained in multiple regression analysis can be considered as the relative weights of the factors, their values are obtained indirectly through testing. Additionally, a negative value of beta can be found, making it complex to judge the importance of the resultant value (Shieh, Chang, Fu, Lin, & Chen, 2014). Making decisions has continually been an essential activity in day to day life. Therefore, using services such as MFS necessitates a careful decision from an individual so that he/she would not regret his/her decision, ever since decision-making has emerged as a mathematical science today (Figueira, Greco, & Ehrgott, 2005). From there, multiple criteria decision-making (MCDM) techniques constitute a critical framework through which companies focus on which strategy to implement to meet the needs of consumers, to acquire the appropriate income and to prosper in the competitive milieu (Valaskova, Kramarova, & Bartosova, 2015). A growing number of recent studies link MCDM to financial decision making (Gupta, Bhaskar, & Singh, 2017). In order to advance current IS researches, Esearch & Koppius (2011) stressed that there is a necessity to integrate decision modeling methods in IS research to generate data estimates as well as methods for assessing the analytical power of the result. Therefore, adopting a multi-analytic approach revealed how merging two diverse data analysis techniques in either methodology or analysis can enhance the validity and confidence in the output (Gbongli et al., 2019; Scott & Walczak, 2009).

Moreover, most managers make strategic decisions grounded in a single goal or dimension, but strategic planning is impacted by many different factors and regarded from several perspectives (Hung, Chou, & Tzeng, 2009). As the traditional notion of strategic planning lacks a multidimensional prominence, this paper integrates the structural equation modeling and technique for order preference by similarity to ideal solution (SEM-TOPSIS) method to construct the relationships between decision factors for MFS adoption strategic planning, while classifying the alternative of MFS. It is a unique decision support technique grounded in structural modeling.

The primary objectives of this research are: To explore the influential antecedent of trust and risk perception at the multidimensional level regarding MFS adoption. To propose and validate model MFS acceptance using SEM technique by employing data collected through experts of MFS and MFS experienced users. To develop an SEM-TOPSIS-based model for multi-criteria decision making by selecting the appropriate MFS type for MFS, based on experts' viewpoint and by ranking the effective trust-risk factors while uncovering the hidden relationship amongst factors that influence customers in the MFS.

As the main contribution, this study incorporates a complex multi-criteria decision-making problem by assessing types of multidimensional trust and risk in MFS that have rarely been investigated and touched in past studies. Furthermore, a literature review is conducted, and then SEM analysis is used to construct a hierarchical structure for trust and risk factors, which includes a total of ten sub-factors. According to the identified criteria and sub-criteria and by considering relationships among them, TOPSIS is adopted for selecting the appropriate types of MFS, based on the critical factors that influence customers' trust and risk. Therefore, the study proposes a solution that could effectively enhance trust and mitigation-perceived risk measures through a multi-level approach to planning strategy by using MFS.

2. LITERATURE AND THEORY REVIEW

2.1 Understanding Mobile Financial Services (MFS)

The rapid adoption of mobile devices in developing countries (Abu-Shanab & Abu-Baker, 2014), together with widespread mobile financial services, have recently drawn practitioners and academics' attention (Laukkanen, 2017). Since consumers are spending gradually more and more time online and are "going mobile", financial digitalization is now driving banks and network companies' providers to undertake the most extensive shift in their history. Mobile financial services (MFS) refers to financial services and transactions delivered through mobile devices (Donovan, 2012). Hence MFS has been described as a generic term, which entails the application of the devices in the conduct of various financial services, including mobile money, mobile credits, mobile insurance, mobile savings, mobile payments, and mobile banking (Adjei & Odei-Appiah, 2018).

MFS characterizes an area of innovation and strategic importance for global initiatives to counter poverty and mobile telecommunication providers (Nesse, Risnes, & Hallingby, 2018). It has been said to have carried about a positive shift in customers' perceptions in many countries. Mobile operators grasp MFS as an opportunity to engender revenue via an adjacent business (both basic payment and services) and recovery of cost and investments through enlarged data usage by consumers (Dennehy & Sammon, 2015). The goals of MFS are accompanied by various advantages for banks, such as the decreased use of cash, while cost-effectively serving the unbanked population, protecting current accounts and products. The major benefit of MFS regarding trade are higher Point-of-Sale (PoS) throughput, less cost for cash handling, and real-time messaging to users. Accessing transaction data and ownership of the user interface are further viewed as an important perceived valued of MFS. For the customer, MFS makes payments possible anytime, anywhere, and with the alleviated risk of theft (cash—especially in underdeveloped countries) (Nesse et al., 2018).

These advantages could be equally valid for Togo. Not much attention has been given to the empirical research on the adoption of MFS in Togo. Furthermore, in less affluent nations stricken with socio-political instability and vulnerability, MFS technologies may have different implications toward usage and are likely to impact the initial decisions to adopt (J. K. Lee & Rao, 2007; Li, Sarathy, & Xu, 2011). The country of Togo sometimes encounters a kind of socio-political crisis. Given a negative socio-political and external influence such as the physical atmosphere of development and growth, policies, regulations, and social environment unsupportive of adoption are suggested to hinder innovation adoption (Wisdom, Chor, Hoagwood, & Horwitz, 2014). MFS unavoidability might confront such challenges due to consumers' lack of trust in the novel wireless technology, and their risk perceptions associated with the open airwaves of a wireless medium. We thus stress that consumers' trust and risk perception may impact their espousal of MFS services.

2.2 Theory and past research

As an emerging service, mobile financial services (MFS) has not been widely adopted by users. Therefore, researchers have paid attention to assess the factors impacting its user adoption. Furthermore, technology adoption is one main area of focus for information systems (IS) researchers. A diversity of theoretical perspectives has been developed to study MFS adoption. More assertively towards another direction, the current literature on consumer behavior related to acceptance of IT, such as MFS, tends to elaborate on a theoretical model of technology adoption theories (Safeena, Date, Hundewale, & Kammani, 2013). They often employ the traditional information system models like theory of reasoned action (TRA), theory of planned behavior (TPB), technology acceptance model (TAM), diffusion of innovation (DOI), innovation diffusion theory (IDT) and unified theory of acceptance and use of technology (UTAUT) as the foundation of their research. They examine whether the models' theoretical constructs are likely to affect the consumer acceptance of an MFS (Gbongli et al., 2019; E. L. Slade, Williams, & Dwivedi, 2013; Tam & Oliveira, 2017; Yan & Yang, 2015) or assess whether consumers are ready to adopt m-payments grounded in the supposed factors (E. Slade, Williams, Dwivedi, & Piercy, 2015).

The TRA model stipulates that a particular behavior is directed by the individual's intention to conduct that action, which itself hinges on the attitude to behavior and subjective norms (Fishbein & Ajzen, 1975). For TPB model, the perceived behavior was added to the attitude towards behavior and subjective norms that affect both the intentions of people's perceived behavior and actual behavior (Ajzen, 1985). Past studies elucidated behavioral perception control as the degree to which one has control over launching a particular behavior as well as facing the circumstances, while the full volitional control over the behavior of interest is found limited (Barua, 2013; Yzer, 2012; Zolait, 2014). Although their finding pinpointed the internal and external factor of perceived control, as an example self-efficacy and facilitating condition, technology, and government sustenance, the utmost impact on the behavior is somehow associated with the type of the innovation. The TAM model, as the extension to the TRA and TPB models, bears a significance of perceived usefulness and perceived ease of use factor to affect actual behavior geared toward innovation (F. D. Davis, Bagozzi, & Warshaw, 1989). Based on the review of TAM literature, Marangunić & Granić (2015) revealed seven past TAM-related studies. However, the goal of these works and the various analysis techniques adopted differ. For instance, Legris, Ingham, & Collerette (2003) examine the question of whether the TAM explains actual use while Mortenson & Vidgen (2016) conducted the review of TAM studies employing the computational literature review (CLR). Moreover, TAM (F. Davis, 1989) and its extended version has been used in various online milieu to assess the adoption of consumer's online-system (Gbongli et al., 2019; David Gefen, Straub, Mack, & Distinguished, 2000; Paul A. Pavlou, 2001; Waite & Harrison, 2015).

The TRA model, however, has some drawbacks, comprising a major threat misleading between attitudes and norms because attitudes can commonly be viewed as norms and conversely. Similarly, further explanatory variables are required for TRA (Thompson, Higgins, & Howell, 1991; Webster & Martocchio, 1992). As such, TAM has then been successfully combined with TRA and TPB in parsimonious capability (Suh & Han, 2003). The theory of adoption, such as DOI theory (Rogers, 1962), is a handy systemic background to define either adoption or non-adoption of new technology. The theory put forward is that people will be more likely to accept innovation grounded in the innovation facets and appearance of comparative benefit, compatibility, intricacy, trialability, and observability (Plouffe, Vandenbosch, & Hulland, 2001). Regardless of the enlightened strength of this model, the weaknesses go a long way in decreasing its power. For instance, the relationship between attitude and espousal or rejection of innovation was restricted (Lei da Chen, Gillenson, & Sherrell, 2002; Karahanna, Straub, Chervany, & Karahanna, 1999); the innovation-decision process and the features of innovation remain unclear as well. The theory posits technology to pass via a linear stage; however, an intricate technology (Lyytinen & Damsgaard, 2001) has been perceived not on linear stages. Rendering to the critical review and meta-analysis of TAM (Legris et al., 2003), it was suggested as a useful model; although, it suffers from the trade-off of dropping information richness resulted from the investigation (Napaporn, 2007).

Despite the various advantages that might be incorporated into every theory or model, their competency in predicting and elucidating is due to the degree to which the predictor could get a sound proportion of variance explained in intention and usage behavior (Singleton, Straits, & Straits, 1993; Taylor & Todd, 1995). Even though the prevailing models, are indicative of e-service or MFS acceptance behavior, many researchers believe that they are not sufficiently robust with regard to assessing all the aspects clients intend obviously throughout the various phases of their decision-making process and thus require further integration (El-Kasheir, Ashour, & Yacout, 2009). George's findings (2004), after reviewing previous information acceptance models, revealed that trust consideration could be a major laudatory and backup for an online vendor.

It is important to recall that trust and risk are interrelated facets (Mayer et al., 1995), where the degree of importance of the situation depends on the impending outcome of risk. Given that the adoption of MFS becomes an important decision that consumers are required to make for a long-term impact, the function of risk is more likely to be vital. The extensive review of the literature revealed diverse antecedents to the adoption of mobile banking (Aboelmaged & Gebba, 2013; Alalwan, Dwivedi, & Rana, 2017; Johnson et al., 2018; Narteh, Mahmoud, & Amoh, 2017; Shaikh & Karjaluoto, 2014). Studies were carried out in both developing and developed countries; however, a limited number have been conducted in Togo (Afawubo, Agbagla, Couchoro, & Gbandi, 2017; Gbongli, Dumor, et al., 2016). These outcomes are therefore insufficient to offer meaningful insights into predicting which multi-dimensional trust and risk influence customers' use of MFS in Togo while providing a strategy decision analysis framework for understanding the multiple factors that entail to the decision of the acceptance. Moreover, many of these theories and models were used in developed countries, and their direct application in developing countries such as Togo might not be sufficiently robust for the economic situation of the country. Given that MFS belongs to information technology to which some adoption model might exist, it requires a distinctive conceptualization that might better pronounce the fact in emerging's countries situation.

Regarding these ends, this study uses components from both trust and risk dimensionality literature and proposes conceptual research to envisage consumer appraisal of MFS (mobile banking, mobile payment, and mobile money transfer) adoption in Togo while ranking their perspective.

3. THEORETICAL FRAMEWORK AND HYPOTHESES

3.1 Antecedent of Trust

The concept of trust remains an intricate, multi-dimensional, and context-dependent paradigm (David Gefen & Straub, 2003). Past researchers emphasize the diverse aspects of trust, a fact which frequently leads to discrepancies between numerous studies outcomes. After the appeal from Gefen et al. (2008) for additional new IT-related research on trust, there is a need to collectively assess the most crucial trust's dimension such as a disposition to trust, technology trust, and vendor trust that seems to impact MFS.

Some scholars have proposed trust dispositional, trust belief, structural assurance (Mcknight & Chervany, 2001). From others' point of view, the interpersonal trust, the dispositional trust, and institutional trust are also essential constituents of the trust dimension (F. B. Tan & Sutherland, 2004). Others found the dimension of trust to be trusting behavior, dispositional to trust, and institution-based trust (Vidotto, Massidda, Noventa, & Vicentini, 2012). Disposition to trust denotes the general susceptibility for a person to trust others (Nor & Pearson, 2008). It is grounded in the personality, which explains the reason why some of us have a tendency to either trust or mistrust and doubt others (Hallikainen & Laukkanen, 2018; Schoorman, Mayer, & Davis, 2007). Disposition to trust is, therefore, crucial for the establishment of initial trust and subsequently accommodating to less importance in the presence of pre-existed trust belief (David Gefen, Karahanna, & Straub, 2003).

Technology Trust is considered as an antecedent of trust and connotes the readiness of an individual, or individual's technological dependency, to achieve a designated task by the positive feature incorporated in the technology (Mcknight, Carter, & Thatcher, 2011) and the benefit arises from the particular technology (Muir & Moray, 1996). With this view, Technology Trust refers to the role of technology in building a trusting relationship with the user (Misiolek, Zakaria, & Zhang, 2002). From the above perspective, when an MFS user considers the technologies that are being applied to be reliable and consistent, then the probability to assess the aggregate service seems more promising, and trust will increase. Although admitting that the threefold technology aspect affects the environment of MFS (i.e., website, network, and mobile technology), the present study intends to treat them as a whole without separating them. As such, the user or potential user is called upon the strong level of comprehensive understanding purposively for MFS optimum usage. Past research has revealed much importance and many benefits of the technological trust to the behavioral field of application (Lankton, Harrison McKnight, Wright, & Thatcher, 2016; Mcknight et al., 2011; Meng, Min, & Li, 2008; Min, Meng, & Zhong, 2008).

Vendor Trust denotes the extent to which the consumer sees and believes that the vendor will accomplish the designated transactional requirements in risky or ambiguous conditions (Bailey, Gurak, & Konstan, 2002). Many situations can raise consumer's trust toward the vendor. An online consumer who perceives the vendor in presenting an opportunistic behavior can create a kind of reluctance within that particular consumer. Earlier studies have revealed a negative relationship between the online vendor's opportunism and online consumer's trust (Paul A Pavlou, Liang, & Xue, 2007). Trust, and in specific the confidence in the mobile vendor, plays an exceptionally important role in the digital environment (Liu, Min, & Ji, 2009; Nilashi, Ibrahim, Reza Mirabi, Ebrahimi, & Zare, 2015; S. Yang, 2016; S. Yang, Chen, & Wei, 2015). For Roger C. Mayer et al. (1995), vendor ability, integrity, and benevolence are crucial vendor trust features, although ability can also be regarded as vendor competence (Bhattacherjee, 2002). By relating that logic to the MFS environment, vendors with a good reputation/integrity will be less expected to bear unscrupulous behaviors and threaten their status. As a result, we posit the succeeding three assumptions to inspect the causal effect relationships between trust's antecedents and trust in the MFS perspective.

H1. Dispositional trust would significantly influence users' trust in using MFS.
H2. Technological trust would significantly influence users' trust in using MFS
H3. Vendor trust would significantly influence users' trust in using MFS.

3.2 Antecedent of Perceived Risk

Perceived risk can denote a combination of uncertainty added to the severity of consequence involved (Bauer, 1960). It is similarly taught as a kind of uncertainty and outcomes (S. M. Cunningham, 1967). In the psychological field, perceived risk is the emotional sensitivity and subjective thoughts of various objective risks. Although it is the derivative of the objectives risk, nevertheless they are different from each. From the perceptive of trust-risk relationship, prior researchers understood that the readiness to take risks is a general characteristic of all trust circumstances (Costigan, Ilter, & Berman, 1998; Gbongli et al., 2017; Johnson-George & Swap, 1982). From this point, consumer trust could be noticed and subjected to the degree of the intricate risk presented in the situations (Koller, 1988). Awkwardly perhaps, due to the complex nature of trust and risk variable, countless scholars have disregarded the function of risk perceptions (D. Gefen, Srinivasan Rao, & Tractinsky, 2003). E-commerce trust investigators have shown that, when trust increases, the trustee's perception of risk reduces and impacts their attitudes to the trustor, which successively, influences the readiness to procurement (Jarvenpaa, Tractinsky, & Vitale, 2000). In the view of risk management field, the risk is the construct associated with the cost of outcomes, empowering trust and risk as mirror images while both incorporate differing relationships (Grandison & Sloman, 2000). The study focuses on the rapport amongst trust and risk (D. Gefen et al., 2003), and the trust-related works and empirical confirmation predominantly emphasis on industrial relationships, nonetheless theoretical and empirical support encountered in MFS is limited. When people trust others, they believe that those they trust will act as anticipated, which diminishes the intricacy of the interaction. Understanding the high convolution of the relationship between trust and risk concept, and considering likewise the absence of scholarly unanimity that lack on how to account their relationship via model (Johnston, Allen C. Warkentin, 2004), this study takes the view of a mediating relationship (D. Gefen et al., 2003) instead. On the mediating standpoint, if trust exists, then the risk perceived is reduced, which successively will impact the degree of decision-making to use MFS. Consequently, trust mitigates the perceived risk effect (Lewis & Weigert, 1985; Luhmann, 1979; Mayer et al., 1995).

These ideas of risk and others will endure a detrimental dominance on the acceptance of MFS. For instance, Swaminathan et al. (1999) revealed consumers' opposition to providing their credit card information through the Internet. With MFS, the consumers are required to entrust not only their credit card information but a whole account of information in most cases. Wide-ranging, trust ameliorates the consumer's conception towards online service and the related component, diminishing the level of the risk perception allied with the transaction process.

From the attribute of risk opinion, a plethora of researchers brought that studies on consumer's risk perception are a kind of a multi-facet concept (Featherman & Pavlou, 2003; Gbongli et al., 2017; Luo et al., 2010), which becomes the root of the aggregate perceived risk.

To date, perceived risk has been employed to elucidate both offline and online risk shopping behavior. The finding derived from the work of Featherman and Pavlou (2003) on the consumer's adoption of e-services has been widely accepted which classified perceived risk dimension as an economic risk, social risk, time risk, functional risk, psychological risk, and privacy risk. Bellman et al. (1999) informed regarding the prominence of time concerns and argued that it is a substantial predictor of online buying behavior. According to the finding, consumers in a hurry who have less time are more plausible to buy on the internet. The perception of time risk can refer to the integration of time lost and determination expended in acquiring any item and service (Murray & Schlacter, 1990). Grounded in this similar logic, the current study proposes that consumers are time-oriented, time-conscious, and therefore value the potential time they might spend in implementing, searching and learning the application process of the new MFS.

Security/privacy risk is categorized as an inherent loss undeviatingly to fraud or hacktivist haggling the security of the user of an online service (M. C. Lee, 2008). Security or privacy issues mostly arise when a customer is transferring money from his/her account or dealing with his/her secluded economic information, and where others view this information without his/her consent (Littler & Melanthiou, 2006). The perception of costs applied to the MFS application reveals fear amongst consumers. Empirical evidence stressed that mobile banking adoption is highly sustained by economic aspects such as beneficial transaction service fees (A. S. Yang, 2009). Alternatively, it is impeded by economic considerations (issues centered on basic fees for assessing mobile banking), like cost burden (Cruz, Neto, Muñoz-Gallego, & Laukkanen, 2010) or high payment incorporated in using mobile banking (Yao & Zhong, 2011). Therefore, the perception of cost risk tends to negate the adoption of mobile banking (Luarn & Lin, 2005).

Centered on the work of Featherman and Pavlou (2003) predominantly, and throughout the previous studies towards risk components so far; the present study deduces four important dimensions of risk perceived, which are expected to influence the consumer's overall risk concerning the MFS adoption. They are the perceived privacy risk, time, security, and financial risk in the form cost perceived. Hence, we can posit the following assumption based on the discussion being done under this section.

H4. Consumer's trust would negatively associate with the perceived risk in MFS.
H5. Perceived privacy risk would significantly influence users' perceived risk of using MFS
H6. Perceived time risk would significantly influence users' perceived risk of using MFS
H7. Perceived security risk would significantly influence users' perceived risk of using MFS
H8. The cost perceived would significantly influence users' perceived risk of using MFS

3. 3 Antecedents of MFS Adoption

Under this section, three antecedents (dispositional trust, trust, perceived risk) of MFS adoption will be taken into consideration. Being part of a personality trait, a disposition to trust can denote an individual's predilection to show reliance on humanity and to support a trusting standpoint concerning others (McKnight, Choudhury, & Kacmar, 2002; McKnight, Cummings, & Chervany, 1998). Many researchers hypothesize the disposition to trust as partaking a positive impact on trust towards online shopping websites (McKnight et al., 2002). This relationship was also supported in various IS research, particularly in e-commerce (David Gefen, 2000; David Gefen & Straub, 2003; K. K. Kim & Prabhakar, 2004) and moreover in mobile banking (Guangming & Yuzhong, 2011). Accordingly, David Gefen et al. (2003) pointed out that disposition to trust is crucial, particularly for the development of early trust and befits less significant for established trust or pre-existing relationships trust beliefs. Once encountering people with trifling or no experience using the wireless internet as a platform for financial transactions, a disposition to trust is predictable to affect their trusting perception on the internet. People partaking high disposition to trust are more favorable to feel relaxed or secured when using wireless internet for financial transactions (Luo et al., 2010). Inferring from this lucidity to the MFS, we expect that consumers having a higher disposition to trust are more probable to espouse MFS than those with a lower disposition to trust.

Taking the antecedent of MFS from a different angle, the importance of trust has been revealed to be an extensive subject matter. Trust, combined with the previous definitions so far, denotes the readiness of one party to be exposed to the actions of another party deal with the hope that the other will accomplish the designated task needed to the trustor (Mayer et al., 1995). The empirical findings of Jarvenpaa and Tractinsky (1999) revealed the trust element to influence the decision to purchase in various manifold cultures. The prominence of trust is so decisive that it may be extended to be viewed as the “wild wild west” of the 21st century (McKnight et al., 2002). The more MFS users or potential users believe and trust the services, the more they can develop an affirmative goal for its usage. User trust, which has been revealed to be an important adoption facilitator in many IS environments, lacks adequate inspection in the context of MFS as a whole.

The next antecedent of MFS adoption resides in risk perception. Since its application amongst consumer behavior literature (Bauer, 1960), the conception of perceived risk has been reviewed from a multiplicity of viewpoints. The classical decision concept considers risk perception as a function of the distribution of probable outcomes of conduct, its likelihoods, and subjective values (Pratt, 1964). Accordingly, risk encompasses two dimensions: uncertainty and outcome where there is the possibility of experiencing a loss as a consequence of a behavior and the significance accredited to the loss (Cox, 1967; Kogan & Wallach, 1964). While various researchers have criticized this approach due to its strictness to apprehend a perceived risk variable equally to be ambiguous and indistinct (Sjoberg, 1980), some others were heightened to this concept definition as expected utility theory (Bonoma & Johnston, 1979; Currim & Sarin, 1983). Explicitly risk, therefore, carries on the subjectively driven expectancy of loss by the customer when denoting the perceived risk (L. F. Cunningham, Gerlach, Harper, & Young, 2005). Internet banking and MFS, predominantly mobile banking, rely on a similar type of risk (M. S. Y. Lee, McGoldrick, Keeling, & Doherty, 2003), only, the information media channels differ. Prior IS studies showed that the imperative attitudinal of perceived risks impact adoption behavior where much is based on the privacy risk and transaction security risk (Abrahão et al., 2016; K. K. Kim & Prabhakar, 2000; Laforet & Li, 2005; E. Lee, Kwon, & Schumann, 2005; M. Tan & Teo, 2000). Preceding studies have equally supported the negative effect of the perceived risk of online usage and purchasing behavior (D. J. Kim, Ferrin, & Rao, 2008; Liang & Huang, 1998; Liao & Cheung, 2001; P.A. Pavlou, 2003). Likewise, earlier researchers agreed that the more risk is perceived by someone in purchasing context, the less probable he/she will be resolved to buy (Dowling & Staelin, 1994).

Furthermore, the level of personal participation in the decision-making process exposes the degree of risk perceived combined with the significance attributed to the choice of the object while allowing for the desires, interest and personal values of the individuals (Assael, 1998; Coulter, Price, & Feick, 2003). Based on the perception of risk assigned in past works as the main inhibitor elements of various IS arena; similarly, it is expected to affect the acceptance of MFS negatively. In line with the literature allied with the antecedent of adoption of MFS in this study, we can, therefore, posit as follows:

H9. Disposition to trust would have a positive effect on individual’ espousal of MFS.

H10. User trust will positively influence an individual’s acceptance to use MFS.

H 11. User risk perception in MFS would have a negative impact on the adoption of MFS.

3. 4 Conceptual Framework

To assess how trust and risk perceptions at the multidimensional level affect mobile financial services (MFS) acceptance in Togo, we propose a research model, as shown in Figure 1. The proposed model is used to identify several attributes as predictors of MFS. Based on the above discussion related to the suggested hypotheses, we considered three antecedents for the general trust (dispositional trust, technology trust, and vendor trust), four antecedents for the aggregate perceived risk (privacy risk, time risk, security risk, and cost).

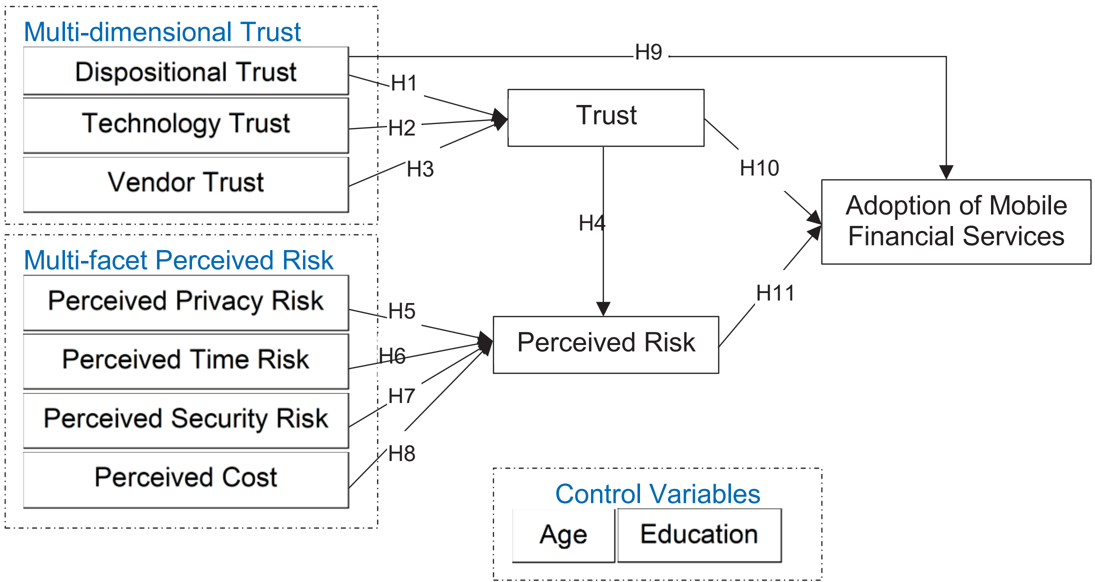


Figure 1. Proposed research model.

4 RESEARCH METHODOLOGY

4.1 Design and Data Collection

Various schools of thought questioned how data collected would be executed as well as the content of the studies. Amongst them, Cooper & Schindler (2003) have suggested two approaches of scrutinizing issues: One technique called observational approach is to gather data on people, event, situations, and behavior; while the next one, the so-called communication approach, has considered the attitudes, expectations intentions, and motivator aspect.

This research, as a result, used data collection via communication approach taking a form of the survey since the motive of the study turns to capture the influential factor of MFS adoption once testing the research model. A survey instrument was then established for indicators and criteria development, which primarily got ratified after revising the suitability of the constructs by the chosen experts of MFS. The preliminary draft of the questionnaire was prepared in English then translated into French (the official language of Togo) for its assessment as well. Both questionnaires in English and French have been retained as to avoid any confusion related to the scope, purpose, and content; so far, allowing the comparison of the versions for discrepancies concerns, steadfastness to be easily acknowledged and established. Following the advice and the opinion from the experts, redundant and confusing items were either improved or removed. As a result, new items were included in the questionnaire lastly, permitting the validity of the survey instrument employed. The research model embodies ten factors; each factor remains evaluated with multiple items. Also, all item were accommodated from existent literature to increase content validity (Straub, DetmarBoudreau & Gefen, 2004). There were two types of questionnaire. The first type (SEM questionnaire) was divided into two parts. The first part was distributed with bio-data of the sample, and the second part answered the MFS questions using the five-point Likert scale bounded from strongly disagree (1) to strongly agree (5).

For the second type of questionnaire (TOPSIS questionnaire), we arbitrarily contacted users and potential users and questioned them whether they had mobile MFS usage experience to ensure their familiarity to some extent as recommended (Gbongli, 2016, 2017). Thus, those with two or more MFS experience years were further invited to fill the TOPSIS questionnaire format.

The empirical study took almost three months of the span for data collection due to the delay in obtaining some participants' responses and an awkward time-period indicated by some of them. Data were collected at some of the busiest and most crowded places of the capital town Lomé (i.e., Assivito, Dekon, Be, and Université de Lomé) where potential users and currents users of mobile financial services (MFS) can relatively be found and inspected better than in other sectors. Literate people filled in their survey questionnaires themselves, whereas for illiterates, help was given. The questionnaire took almost 10–15 minutes to complete by a given participant.

Once the data collection procedure was completed, we examined all questionnaires and discarded cases with too many missing and or rushed responses.

As such, 538 questionnaires were both ready and yielded usable samples. Amongst them, 294(54.6%) were male and 244(45.4%) female. Seventy-five (13.9 %) respondents were aged below 18 years, 145(27%) aged between 19-24 years, 199 (37%) aged between 25-30 years, and 119(22.1%) aged above 31 years. Regarding educational qualifications, the majority of respondents (two hundred and sixty-seven) had a high school certificate or below, i.e., Baccalaureate (49.6%), 203(37.7%) had a graduate degree, while 57(10.6%) had a master degree. The remaining (2%) had a doctorate. Concerning MFS years of experiences, 187(34.8%) of respondents claimed to have no experience with MFS, 194(36.1%) used it for less than one year, 125(23.2%) MFS usage ranged from the 1-2 years, 26(4.8%) were found between 3-4 years of MFS experience. Only 6(1.1%) had MFS experience for more than five years. Hence, very few respondents had MFS experience above three years from the deduction. Moreover, they are those respondents engaged in MFS application at the early stage of its implementation (Most MFS companies in Togo started launching their activities in the year 2013) and dwell on it.

4.2 Proposed Technique of Data Analysis: SEM-TOPSIS Methods

The SEM-TOPSIS technique was employed to construct the MFS evaluation decision support system. Therefore, SEM was utilized to generate critical criteria and weights, whereas TOPSIS was used to engender the rank and score of alternatives as well permitted the fullness of the data, improved the data accuracy via group decision making.

SEM is suitable to estimate and test casual relationships by employing a combination of statistical data and qualitative assumptions (Gbongli et al., 2019; Kumar Mittal & Singh Sangwan, 2014). It remains a second-generation multivariate technique that tolerates the simultaneous assessment of multiple equations, embraces multiple regression analysis, factor analysis, and path model analysis (Hair, Black, Babin, Anderson, & Tatham, 2010). SEM incorporates the whole analysis of construct concurrently rather than separately (Chin, 1998), with this application being emergent in the social sciences (J. C. Anderson & Gerbing, 1988). Accordingly, it is the handiest method adapted for checking causative relations between predictors and adoption behavior (Hair, Black, Babin, Anderson, & Tatham, 2006; Schumacher & Lomax, 1996). It offers greater flexibility in matching a theoretical model with a data sample when compared with techniques like PCA and factor analysis (Aloini, Martini, & Pellegrini, 2011).

TOPSIS: Technique for Order Preference by Similarity to Ideal Solution. The various process of TOPSIS will concurrently be explained along with the finding under the analysis section.

5. DATA ANALYSIS

5.1 Measurement and Hypotheses Testing with SEM Analysis

We performed exploratory factor analysis (EFA) employing maximum likelihood estimation with Promax because of the large sample of data set (n=538) and its intricacy related to the outcomes elucidation which is trivial in resolving the correlated. The EFA reveals the output of KMO as 0.809 and Bartlet's test of sphericity to be significant at α=0.000 with a Chi-square of 11598.920, indicating the relevance for performing exploratory factor analysis (Kaiser, 1974). Besides, the communalities for each variable were sufficiently high (lowest was 0.343, the majority were beyond 0.597, and the greatest was 0.975) showing the evidence that these variables were effectively correlated for factor analysis. The ten-factor model obtained a total variance explained with more than 60% along with all extracted factors partaking eigenvalue beyond 1.0.

To continue assessing our quantitative model, we settled the subsequent analysis in two phases (J. C. Anderson & Gerbing, 1988): first, via confirmatory factor analysis (CFA), we appraised both reliability and discriminant validity of the ten constructs (Campbell & Fiske, 1959). The outcomes will achieve validity unless the researchers employ constructs that diverge from another construct in a similar model (Campbell & Fiske, 1959). From the second step, we valued the structural model then SEM for hypotheses testing. These last two steps are adopted from previous studies (Gbongli et al., 2017; Zhou, 2012). Hence, we estimated the reliability of each construct based on three indices, such as composite reliability (CR), average variance extracted (AVE), and Cronbach's alpha (CA). The suggested values for good measures were at least 0.70, 0.50, and 0.70, respectively (Fornell & Larcker, 1981), (see Table 1). In patronage of convergent validity, the AVE found to be higher than 0.5 for all constructs, and all item factor loadings remain beyond the minimum threshold of 0.4 (Hair, Anderson, Tatham, & Black, 2010).

Moreover, all loadings of items arose in the corresponding construct, and no item loaded with the high value in another construct. This technique was espoused in past research (Gbongli et al., 2019; Y. Hwang, 2014; Zhou, Lu, & Wang, 2010). As such, we established that our ten constructs displayed convergent validity (see Table 1 below).

Table 1. Reliability and validity in CFA

	CR	**AVE**	**MSV**	**MaxR (H)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
(1)	0.846	0.647	0.227	0.848	**0.804**									
(2)	0.933	0.779	0.133	0.965	0.108	**0.883**								
(3)	0.904	0.704	0.087	0.975	0.216	0.067	**0.839**							
(4)	0.860	0.609	0.057	0.979	0.168	0.157	0.155	**0.780**						
(5)	0.855	0.664	0.227	0.981	0.476	0.020	0.230	0.144	**0.815**					
(6)	0.843	0.577	0.056	0.984	0.236	0.061	0.114	0.011	0.235	**0.760**				
(7)	0.856	0.600	0.133	0.985	0.041	0.365	0.044	0.238	-0.022	-0.072	**0.775**			
(8)	0.811	0.594	0.065	0.987	0.127	0.155	0.113	0.232	0.091	0.035	0.255	**0.771**		
(9)	0.798	0.571	0.013	0.987	0.102	0.098	0.065	-0.004	-0.035	0.086	0.075	0.115	**0.756**	
(10)	0.820	0.610	0.087	0.988	0.228	0.051	0.295	0.064	0.198	0.216	-0.042	0.019	0.104	**0.781**

Note: (1): DTrust-Dispositional Trust; (2): TTrust-Technological Trust; (3): Vtrust-Vendor Trust; (4): PPrivR-Perceived Privacy Risk; (5): PTimeR-Perceived Time Risk; (6): PSecurR-Perceived Security Risk; (7): PCost-Perceived Cost, (8): PRisk-Perceived Risk; (9): AdMFS-Adoption of MFS; (10) Trust.

We designed Table 2 to portray the goodness of fit of CFA and SEM. Apart from the goodness-of-fit index (GFI) for CFA slightly below the recommended, as this index is sensible to sample size, and in this study, we use large sample size (n = 538); for all indexes, our measurement model and structural model indicated sufficient goodness of fit.

Table 2. The goodness of fit (CFA and SEM)

Indices	Abbreviation	CFA Value	SEM Value	Thresholds	References
Chi square	x^2	1068.904	30.445	Pval>0.05	(Wheaton, Muthen, Alwin, & Summers, 1977)
Normed chi square	$x^2_{/DF}$	2.104	1.903	1< x2/df< 3	(Bm M. Byrne, 2010)
Root mean square residual	RMS or RMR	0.066	0.015	<0.08	(Browne & Cudeck, 1993)
Goodness-of-fit index	GFI	0.889	0.991	>0.90	(Jöreskog & Sörbom, 1984)
Adjusted GFI	AGFI	0.862	0.955	>0.80	(David Gefen, Straub, & Boudreau, 2000)
Normed fit index	NFI	0.900	0.941	>0.90	(Bollen, 1989)
Comparative fit index	CFI	0.944	0.968	>0.93	(B. M Byrne, 1994)
Tucker-Lewis index	TLI	0.935	0.869	0<TLI<1 TLI > 0.9	(Bentler & Bonett, 1980)
Root mean square error of approximation	RMSEA	0.045	0.041	<0.05 excellent fit <0.08 good fit	(Browne & Cudeck, 1993)

Before the structural model, we conducted a common method bias. Since the data for the variable was led through a single method (survey), we performed a test to check if a common factor might have been impacted our outcomes. Hence, the test adopted was unmeasured latent factor suggested by Podsakoff, MacKenzie, Lee, & Podsakoff (2003) and Siemsen, Roth, & Oliveira (2009) towards studies that do not obviously measure a common factor, mentioned as a common latent factor (CLF) method. The most prevailing and best method in checking the CMB is the zero-constrained test where the CLF is involved along with Marker if accessible (Podsakoff et al., 2003). This approach checks whether the shared variance across all variable differs significantly from zero. In a case it is, then there are bias issues. To proceed, we computed the chi-square difference test among the unconstrained model and the model per all paths regarding the CLF constrained to be zero. Since the result is markedly different from zero, we can conclude that method bias does occur in our measures. Thus, moving to the causal model based on the result, CLF was retained for our structural model (by imputing composites in AMOS in the presence of CLF) which provided CMB-adjusted values.

We also check for invariance (configurable and metric) due to the presence of two groups, such as gender included in our data to see whether the factor and loading are adequately equivalent across groups. Davidov (2008) has claimed that the assessment of path coefficients could only be useful if the invariance test has been done beforehand. The result signpost that the model fit of the unconstrained measurement models (per groups loaded distinctly) presented a sufficient fit ($\chi 2$ /DF = 1.623, TLI=0.928, CFI=0.938, RMSEA=0.034) when assessing a freely estimated model across genders. Grounded on the result, the model is configurally invariant. Once the model was constrained to be equal, the result of the chi-square difference test reveals p-value (0.226) to be nonsignificant. So, the measurement model satisfies the benchmarks criteria for metric invariance across gender as well. Then and there, we move on making the composite from this measurement model, to build SEM for verification of hypotheses testing. The results of the structured model, together with parameters, was obtained while controlling for age and education. The standardized path coefficients, path significances, and explained variance R^2 of the structural model (See fig 2).

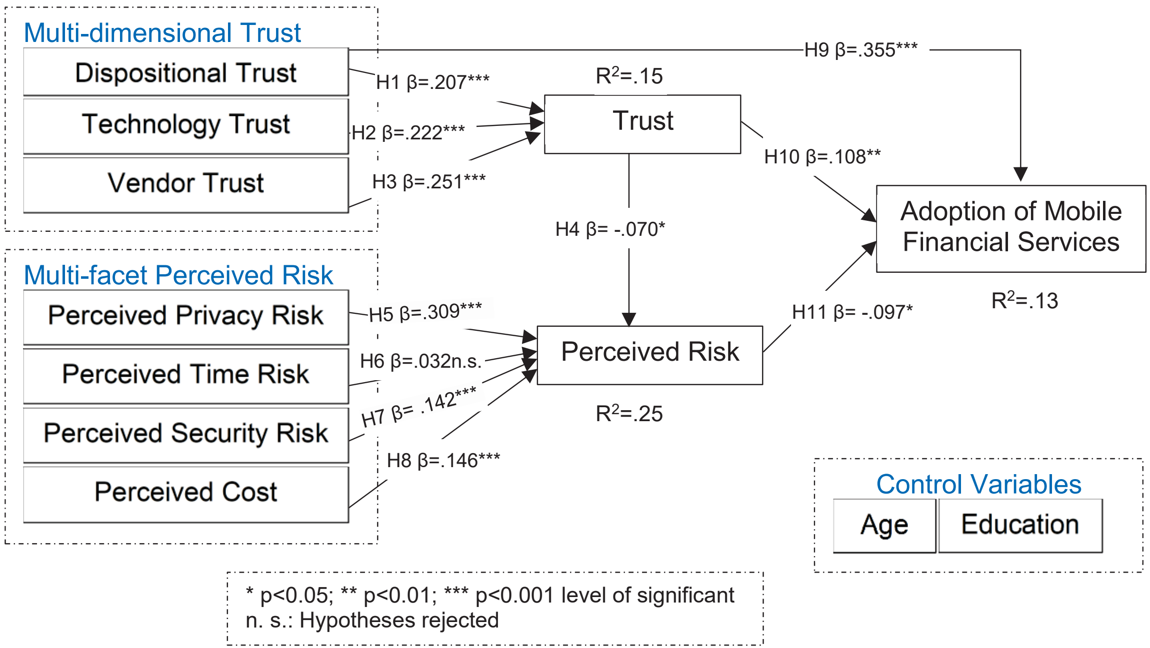


Figure 2. Final model after validation

5.2 TOPSIS Analysis

The Technique for order preference by similarity to ideal solution (TOPSIS) is a multiple criteria decision-making (MCDM) technique developed by Hwang & Yoon (1981). It is grounded in the criteria that the alternative should have the shortest distance from the positive ideal solution and the farthest from the negative ideal solution (Patil & Kant, 2014). It has been extensively employed by researchers for the ranking of alternatives centered on different criteria (Dhull & Narwal, 2018; Gbongli, 2016; Gbongli, Dumor, et al., 2016; Kumar Mittal & Singh Sangwan, 2014; Mahdevari, Shahriar, & Esfahanipour, 2014). When compared to other MCDM methods, TOPSIS:

• Necessitates limited subjective inputs from decision-makers (Vinodh, Prasanna, & Hari Prakash, 2014), it is a deterministic technique; • provides solution on both positive and negative way, which is beneficial for applications where there are considerations such as cost and benefits; and • is a rational method and works agreeably across various application areas (Behzadian, Khanmohammadi Otaghsara, Yazdani, & Ignatius, 2012). Recall that the process of the SEM-TOPSIS can be characterized as follows. Primarily, SEM was applied to compute the hierarchical criteria and their relatives to ensure their significance. This is the reason why having the relative weightage obtained from SEM is reflected more valid than via any other method. The antecedent of trust and perceived risk given by the SEM model were deliberated for the relative weightage of the sub-criteria. The relative weightage is computed from the standardized total effect, normalized (Gbongli, 2017; Punniyamoorty, Mathiyalagan, & Lakshmi, 2012) and presented in Table 3. The weightings showed the importance of each sub-criteria for the MFS companies.

Table 3. Relative weightage of sub-criteria

DTrust	TTrust	VTrust	PPrivR	PTimeR	PSecurR	PCost
0.265	0.128	0.177	0.200	0.021	0.109	0.101

To compute the relative weightage of MFS alternatives to each sub-criteria towards the criteria (Trust and Risk), the decision matrix of alternative performance evaluation (Eq.1 step 1) was created. Participants (74 MFS experienced users and experts) were asked to provide a set of values within the range of one to nine for the sub-criteria, as shown in Table 4.

Table 4. Transformation of linguistic scale into quantitative values		
Linguistic Scale	Quantitative Values	
	Benefit-Max	Cost-Min
Very High	9	1
High	7	3
Average	5	5
Low	3	7
Very Low	1	9
Intermediate values between the two-adjacent judgment: (2,4,6,8)		

The computation of TOPSIS methods grounded on Hwang & Yoon (1981), Lin & Tsai, (2010), and predominantly the one required for grouping decision Shih et al. (2007) were adopted and presented as followed:

Step 1: construction of decision matrix $\mathrm{D}^{\mathrm{k}}, k = 1, \ldots, K$ for each DM. The matrix structure can be viewed as follows:

$$\mathrm{D}^{\mathrm{k}} = \begin{array}{c} \\ \\ A_1 \\ A_2 \\ \vdots \\ A_{\mathrm{i}} \\ \vdots \\ A_{\mathrm{m}} \end{array} \begin{array}{c} n\ Criteria \\ \begin{array}{cccccc} X_1 & X_2 & \cdots & X_{\mathrm{J}} & \cdots & X_{\mathrm{n}} \end{array} \\ \begin{bmatrix} \mathrm{x}_{11}^{\mathrm{k}} & \mathrm{x}_{12}^{\mathrm{k}} & \ldots & \mathrm{x}_{1\mathrm{j}}^{\mathrm{k}} & \ldots & \mathrm{x}_{1\mathrm{n}}^{\mathrm{k}} \\ \mathrm{x}_{21}^{\mathrm{k}} & \mathrm{x}_{22}^{\mathrm{k}} & \ldots & \mathrm{x}_{2\mathrm{j}}^{\mathrm{k}} & \ldots & \mathrm{x}_{2\mathrm{n}}^{\mathrm{k}} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ \mathrm{x}_{\mathrm{i}1}^{\mathrm{k}} & \mathrm{x}_{\mathrm{i}2}^{\mathrm{k}} & \ldots & \mathrm{x}_{\mathrm{ij}}^{\mathrm{k}} & \ldots & \mathrm{x}_{\mathrm{in}}^{\mathrm{k}} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ \mathrm{x}_{\mathrm{m}1}^{\mathrm{k}} & \mathrm{x}_{\mathrm{m}2}^{\mathrm{k}} & \ldots & \mathrm{x}_{\mathrm{mj}}^{\mathrm{k}} & \ldots & \mathrm{x}_{\mathrm{mn}}^{\mathrm{k}} \end{bmatrix} \end{array} \qquad (1)$$

Where A_i refers to the likely alternatives of the decision process i with $i = 1, \ldots, m$; X_j denoting the attribute or criterion $j, j = 1, \ldots, n$; with both quantitative and qualitative data. The value $\mathrm{x}_{\mathrm{ij}}^{\mathrm{k}}$ remains, therefore, the performance score of alternative A_i in relation to attribute X_j by decision-maker $k, k = 1, \ldots, K$, while x_{ij}^k is the element of D^{k}. It is important to mention that there should be K decision-maker matrices designed for K participants of the group. Moreover, the output of the qualitative attribute from each alternative can also be set as discrete value or linguistic values (referring to Table 4) intentionally that the quantitative values could be set in the decision matrix above (C.-L. Hwang & Yoon, 1981).

Step 2: the normalized decision matrix $\mathrm{R}^k, k = 1, \ldots, K$ is generated for each DM. Vis-à-vis to any $DM\ k$, the normalized value r_{ij}^k from the decision matrix R^{k} can take any linear-scale transformation to preserve $0 \leq r_{ij}^k \leq 1$ inequality. Since we consider the vector normalization operation, then r_{ij}^k can be computed as follow:

$$r_{ij}^k = \frac{x_{ij}^k}{\sqrt{\sum_{\mathrm{j}=1}^{\mathrm{n}} \left(\mathrm{x}_{\mathrm{ij}}^{\mathrm{k}}\right)^2}}, \qquad (2)$$

Where $i = 1,2, \ldots, m$; $j = 1,2, \ldots, n$; and $k = 1,2, \ldots, K$. It is also necessary to clue that vector normalization method makes provision as to which one represents a cost criterion for additional management. Moreover, there is no need to directly assess the weighted normalized as per the case of the original TOPSIS (Shipley, de Korvin, & Obid, 1991).

Step 3. Computation of the positive ideal solution V^{k+} (PIS), and negative ideal solution V^{k-} (NIS) for each DM $k = 1, \dots, K$.
For any given $DM\ k$, his/her PIS and NIS can be characterized in the form of

$$\mathrm{PIS} = \mathrm{V}^{\mathrm{k+}} = \{\mathrm{r}_1^{\mathrm{k+}}, \dots\dots, \mathrm{r}_\mathrm{n}^{\mathrm{k+}}\} = \left\{\left(\max_i \begin{matrix} k \\ r \\ ij \end{matrix}\middle| \mathrm{j} \in \mathrm{J}\right), \left(\min_i \begin{matrix} k \\ r \\ ij \end{matrix}\middle| \mathrm{j} \in \mathrm{J}'\right)\right\} \tag{3}$$

$$\mathrm{NIS} = \mathrm{V}^{\mathrm{k-}} = \{\mathrm{r}_1^{\mathrm{k-}}, \dots\dots, \mathrm{r}_\mathrm{n}^{\mathrm{k-}}\} = \left\{\left(\min_i \begin{matrix} k \\ r \\ ij \end{matrix}\middle| \mathrm{j} \in \mathrm{J}\right), \left(\max_i \begin{matrix} k \\ r \\ ij \end{matrix}\middle| \mathrm{j} \in \mathrm{J}'\right)\right\} \tag{4}$$

Where J is related to the benefit criteria and J' allied with the cost criteria, $i = 1, \dots, m$; $j = 1, \dots, n$; and $k = 1, \dots, K$

Step 4. A weigh vector W is allocated to the attribute set for the group. Each DM will produce weights for attributes as w_j^k where $j = 1, \dots, n$ and $\sum_{\mathrm{j=1}}^{\mathrm{n}} w_j^k = 1$; and for each DM $k = 1, \dots, K$. Each element of the weigh vector W will result from the operation of the corresponding components of the attributes' weights for every DM.

Step 5. Evaluate the separation measure through the positive ideal and the negative ideal solutions, $\overline{S_I^+}$ and $\overline{S_I^-}$, relatively to the group. Due to the group decision with respect to this research, this step requires two sub-steps, where the initial one considers the distance measure for individuals while the next one aggregates the measure for the group.

- **Step 5a**. Assessment of the measure from PIS and NIS individually.
  Assuming a given DM k, his or her separation measures from PIS and NIS via Euclidean distance are

$$\mathrm{S}_\mathrm{i}^{\mathrm{k+}} = \sqrt{\sum_{\mathrm{j=1}}^{\mathrm{n}} w_j^k \left(v_{ij}^k - v_j^{k+}\right)^2} \text{, for alternative } i, i = 1, \dots, m \tag{5}$$

$$\mathrm{S}_\mathrm{i}^{\mathrm{k-}} = \sqrt{\sum_{\mathrm{j=1}}^{\mathrm{n}} w_j^k \left(v_{ij}^k - v_j^{k-}\right)^2} \text{, for alternative } i, i = 1, \dots, m. \tag{6}$$

- **Step 5b.** Assessment of the measure from PIS and NIS for the group. In this part, the individual group measure of each alternative is to be integrated via an operation ⊗ for all DMs, $k = 1, \dots, K$. As such, the twofold measure of the PIS and NIS are presented below

$$\overline{S_i^+} = \overline{S_i^{1+}} \otimes \cdots \otimes \overline{S_i^{K+}} \text{, for alternative } i, \tag{7}$$

$$\overline{S_i^-} = \overline{S_i^{1-}} \otimes \cdots \otimes \overline{S_i^{K-}} \text{, for alternative } i. \tag{8}$$

Though this operation can provide various choices like geometric mean, arithmetic means with their related extended; this study pondered only on the geometric one for the group computation. Its calculation's formulae are below shown for PIS and NIS (Eq. 9); (Eq. 10)

$$\overline{S_i^+} = \left(\prod_{k=1}^{K} S_i^{k+}\right)^{\frac{1}{K}} \text{, for alternative } i, \tag{9}$$

$$\overline{S_i^-} = \left(\prod_{k=1}^{K} S_i^{k-}\right)^{\frac{1}{K}} \text{, for alternative } i. \tag{10}$$

Where $i = 1, \dots, m; k = 1, \dots, K$

Step 6: Measure the relative closeness $\overline{C_i^*}$ to the ideal solution for the group while ranking the alternative in descending order. The following formulae (Eq.11) reflect the relative closeness of the i^{th} alternative A_i regarding PIS.

$$\overline{C_I^*} = \frac{\overline{S_i^-}}{S_i^+ + S_i^-}, i = 1, \dots, m \tag{11}$$

With $0 \leq \overline{C_I^*} \leq 1$. The larger the index values, the higher the rank order and so, the better the alternative' performance.
Step 7: Ranking the preference order which remains to define the rank a set of alternatives by comparing their $\overline{C_I^*}$ values in descending order are starting from values closest to 1.

Following the procedure of the TOPSIS method, through a TOPSIS algorithm built-in MATLAB technical computing tool, the relative weightage of MFS allied with each sub-criterion is calculated and shown in Table 5. After aggregating the individual PIS and NIS via geometric mean from the Step 5b eq. (7) and eq. (8), then the final score $\overline{C_I^*}$ is computed using Eq. (11) of Step 6, followed by the ranking of MFS perspective from step 7 as being portrayed in Table 6 and Fig. 3.

Table 5. Summary of the relative weightage of MFS to each sub-criterion

Sub-criteria weightage	0.265	0.128	0.177	0.200	0.021	0.109	0.101
Sub-criteria	DTrust	TTrust	VTrust	PPrivR	PTimeR	PSecurR	PCost
MB	5.40	4.78	6.23	-3.56	-7.53	-4.20	-6.50
MP	8.50	5.40	4.43	-2.34	-8.20	-5.00	-7.00
MMT	7.30	4.70	5.11	-1.42	-8.00	-4.48	-7.20

Table 6. MFS perspectives ranking

MFS	$\overline{C_I^*}$	Rank	% distribution of coefficient
MMT	0.7454	1	46.68%
MP	0.6106	2	38.24%
MB	0.2407	3	15.07%

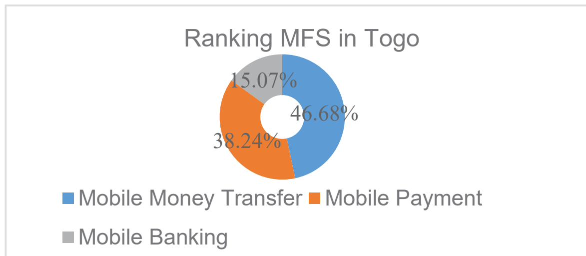


Figure 3. MFS perspectives in % level of classification

6. DISCUSSION

New technology adoptions are impacted mainly by many factors, which may vary from technology concerns to the trust dimension, the perception of risk facets, and the behavior of users, to mention a few. The intricacy and significance related to the effort in elucidating the motives or reasons for users' adoption or rejection of new IT have led to the development of various concepts. Furthermore, there are a plethora of studies on the influence of trust and perceived risk with their determinant towards the adoption decision in an online environment.

Conversely to prior research works, this study scrutinizes the influence of critical variables such as multi-dimensional trust and perceived risk facets on the consumers' adoption behavior of MFS and incorporates each of them into the MFS alternatives decision-making scenario. Some postulations were made towards the possible relationship amongst the factors. The findings are yet to be probed purposively to draw an important conclusion and implication. The result of the MFS structural model analysis regarded as a final model after validation is summarized and portrayed in Figure 2. To be specific, the discussion section is scheduled to be under two sections. The first section will be made with SEM methodology grounded on hypotheses results, which are comprised of three sets. First set: hypotheses associated with trust; second set: hypotheses associated with perceived risk; and third set: hypotheses associated with MFS adoption constructs. The last section of the discussion is booked for a succinct analysis of TOPSIS output obtained via SEM-TOPSIS hybrid technique.

It should be mentioned that all hypotheses were tested when controlling for age and education. The reason for controlling variables is to support mitigating the unrelated effect.

Moreover, its use contributes to improving the robustness and validity of the outcome. In terms of relationships, the study account for the p-value column allied with each variable where the related p-value of less than 0.05 indicates a significant relation associated. The results of the entire tested eleven hypotheses were statistically significant except for the relationship between perceived time risk (PTimeR) and perceived risk (PRisk, i.e., H6 as displayed from Figure 2.

The first set is hypotheses related to trust, which was scrutinized by H1-H3. Empirical evidence is found to accept hypothesis H1 ($\beta = 0.207$, $p < 0.001$), which refers to the positive effect the disposition to trust has on trust in MFS**.** Payne and Clark (2003) showed that the general disposition to trust exerts a substantial control on the trust amongst senior managers in an industrial context. Moreover, Consumers' disposition to trust has been revealed to maintain a strong influence on their trust in an e-vendor (McCord & Ratnasingam, 2004). Although most of the previous studies did not plainly define the direction of the impact, the present study ratifies that disposition to trust and trust are positively related in MFS. Such information comes to back the knowledge that consumers who unveil a greater disposition to trust will more willingly trust the e-vendor (David Gefen, 2000) compared to those who will require more info (Salam, Iyer, Palvia, & Singh, 2005). However, our results are contradicted by earlier e-services (Liu et al., 2009), particularly in mobile banking. The reason might be that when consumers are to encounter a choice within MFS perspectives (mobile banking, mobile payment, mobile money), their trust disposition significantly affects the general trust more or less that of the single type of MFS. As a result, companies dealing with MFS should be aware of this critical effect and prepare for any competitive advantage strategies in the marketplace.

H2 ($\beta = 0.222$, $p < 0.001$) tested the effect of trust in technology on trust in general, and the findings stressed that technology trust has a strong positive impact on trust. Given technological trust as a sole antecedent of trust whereby the object upon which the trust remained imparted when referring to the inert technology (Lippert & Forman, 2006), then, our empirical results are in line with previous findings in the context of mobile banking (Liu et al., 2009). Furthermore, previous works (P. Pavlou & Ratnasingam, 2001) implicitly incorporated the concept of trust technology to trust with its importance being emphasized as a facilitator of e-commerce adoption.

Trust in the vendor was also found to have a positive influence on general trust, which supports H3 ($\beta = 0.251$, $p < 0.001$). The results of this research are reliable with the previous finding in which vendor trust has been defined as multi-dimensional and influential levers that the vendors could employ to build consumer trust (Harrison McKnight, Choudhury, & Kacmar, 2002; Nilashi et al., 2015). Vendor trust remains so vital to promoting trust in changing a potential consumer from curious viewer to one that will be ready to perform MFS. Thoughtful discerning of the essence and antecedents of consumer trust in MFS can support e-vendors with a set of controllable, strategic levers to develop such trust, which will encourage greater MFS acceptance and usage.

As a result, lack of consumer trust (Trust Disposition, Technology Trust and Vendor Trust) in the overall online environment has been, and persists in being, a hamper to IS adoption (Aldridge, White, & Forcht, 1997; Hoffman, Novak, & Peralta, 1999) and thus to MFS. All these could serve as a clue to the concept that the consumers' espousal of MFS may be shaped accordingly.

The second set is hypotheses associated with perceived risk. From this part, perceived risk has five antecedents such as H4 and H5-H8. The investigation of the relationship between trust and perceived risk has been one of the main issues in the development of IS (A. P. Pavlou, 2003). Our result shows that trust has a negative influence on perceived risk, which supports H4 ($\beta = -0.070$, $p < 0.05$). The literature offers supportive studies on the import of this relationship (Muñoz-Leiva et al., 2017; A. P. Pavlou, 2003). Various researchers have also contributed to the belief that trust mitigates consumers' perceived risk (Cheung & Lee, 2000; Fukuyama, 1995; Kesharwani & Bisht, 2012) as well as affecting perceived benefit in e-commerce (Ratnasingham & Kumar, 2000). Lots of incentives that increase trust are similar incentives that reduce perceived risk. This result clarifies to some extent, the doubt related to the direction of the causality between trust and risk, which were found deficient from the past literature (D. Gefen et al., 2003; Mayer et al., 1995; Rousseau, Sitkin, Burt, & Camerer, 1998). From H5-H8, the empirical study found to patronize all the hypotheses at a different level of p-value mention that each dimension of perceived risk has a positive influence on the overall perceived risk except H6 (see Figure 2). At that point, the moderate to weak positive relationships between the perceived risk (aggregate) and the risk component offers further reinforcement that risk can be researched as a multidimensional phenomenon (Zikmund & Scott, 1974). These results are also consistent with the work of Featherman and Pavlou (2003), which validated a majority of these antecedents as a risk dimension; therefore, being the influential element of the aggregated risk. Again, the outcomes reveal the multidimensional nature of perceived risk in information technology, mainly MFS. Boksberger et al. (2007) are supporters of these findings in the area of air travel. Again, the results show that perceived privacy risk H5 ($\beta = 0.309$, $p < 0.001$) is indeed the predominant perceived risk dimension for the partakers of MFS, shadowed by the perceived cost H8 ($\beta = 0.146$, $p < 0.001$) and perceived security risk H7 ($\beta = 0.142$, $p < 0.001$). Moreover, this study confirms the positive effect of the perceived cost on the consumers' perceived risk, such as that the lower the cost, the more minor the perception of risk and the more the likelihood of MFS adoption. As such, the involvement aspect of the risk (Choffee & McLeod, 1973) is importantly observed when the price or cost is high, and the consumers risk losing money.

This research reveals no statistical evidence to support the hypothesis H6 ($\beta = 0.032$, $p < 0.342$) that perceived time risk has a positive influence on the aggregate perceived. Although H6 is rejected; the expected direction of the relationship is kept just so that the p-value is not statistically significant at 0.05. However, our findings are controverted by prior online payment research that has indicated a positive relationship between time risk and perceived risk (Zhang, Tan, Xu, & Tan, 2012). It has stressed that consumers lack patience in waiting a long time because they always delight in pursuing new things (Zhang et al., 2012). Then, a longer waiting time for service delivery would deter the desire, impact their buying disposition or decision to adopt as well. In the view of this current study, the perceived risk dimension, such as perceived time risk, does not appear to impact specific information technology acceptance, at least for the Togolese MFS investigated in this research. The reason may be related to the participants (user and potential user) MFS experience. Since quite many of them lack experience in MFS, they might not be conscious regarding the real time needed for a service done. This implies that the effect of time risk perceived is worthy of further development in future researches and MFS companies are encouraged to continue easing the transaction process of MFS in terms of time spent.

The third set is hypotheses associated with the adoption of MFS. Among them, the hypothesis associated with the positive relationship that the dispositional trust has with MFS adoption was supported by the test result, hence, H 9 ($\beta = 0.355$, $p < 0.001$) is accepted (Figure 2). This infers that when increasing the level of trust disposition, individuals tend to adopt MFS technologies without necessarily cogitating on the general trust. The finding is consistent with e-commerce adoption for SMEs (Chakuthip, Brunetto, Rod, & Sheryl, 2007). Moreover, the scholars reported that indicators for the dispositional trust should be incorporated into empirical studies either as a moderating variable or as a precursor of trusting beliefs, intentions, and behaviors (Grabner-Kräuter & Kaluscha, 2003). Being an antecedent of trust, a disposition to trust remains one of the most operative elements required during the launch phases of a relationship when parties are generally unacquainted with each other (Rotter, 1971). Given that MFS is still in the early stage of adoption in Togo, services providers are recommended to promote the variable that could increase the consumer's dispositional trust.

The study entails and accepts the hypothesis H10 ($\beta = 0.108$, $p < 0.008$) in which general trust has a positive influence in MFS adoption. Generally, trust remains a vital factor in various economic and social relations involving uncertainty and reliance (Hosmer, 1995; Rousseau et al., 1998), particularly those regarding important decisions (Luhmann, 1979) and new technology

(Fukuyama, 1995) as an MFS perspective. Accordingly, our findings are sustained via the idea that trust in business rests on the pertinent and the crucial stimulus of behavior in general (Konovsky & Pugh, 1994; Rossiter & Barnett, 1975; Schurr & Ozanne, 1985), and the facilitator factors for MFS adoption and usage in particular.

Last but not least, from Figure 3, perceived risk significantly negates the adoption and usage of MFS, rendering the support of H11 (β= -0.097, $p < 0.022$). It is so crucial to signpost the feasibility of this outcome to be enlightened by the theory of consumer behavior (Bauer, 1960) allied with risk perception. The importance of perceived risk in the study also confirms previous studies that demonstrate that consumers' perceived risk is more efficacious at clarifying purchasing or adoption behavior inasmuch as consumers are more recurrently driven to avert mistakes than to capitalize on utility in purchasing (Mitchell, 1999). This output is also coherent with a recent report on mobile payment adoption, which underlines rapid technology innovation while stressing the importance of perceived risk in the form of security (De Fouchier & Larduinat, 2016; Manchiraju, Vudayagiri, & Garg, 2016).

Under this set, it can then be deduced that both improving trust and decreasing risk continue to raise the likelihood level of consumers' engagement in MFS transaction. Companies are required to take the necessary precaution to balance the trade-off.

SEM-TOPSIS: It is noteworthy to recall that the second section of discussion concern the output of MFS alternatives computation. The overall result from the TOPSIS technique shows the preference of each alternative regarding the various sub-criteria. The relative closeness $\overline{C_I^*}$ results obtained satisfies the sine qua non-condition, i.e. $0 \leq \overline{C_I^*} \leq 1$. Furthermore, TOPSIS technique is grounded on the principle that the higher the value of $\overline{C_I^*}$, the high the rank order, and consequently, the more the chosen alternatives are favored over others. The final result reveals that mobile money transfer (MMT) is the most preferable MFS to adopt and use with $\overline{C_I^*}$ tantamount to 0.7454 signifying 46.68% compared to the last two remaining. Mobile payment (MP) with 0.6106 (38.24%) was found to be the second MFS alternative used, whereas mobile banking (MB) adoption with 0.2407 (15.07%) is considered minor. This finding is relatively supported from the prior study on mobile banking and mobile payment, where 82% of participants under 35 years old have made mobile payment as compared to 79% who used mobile banking (Fox, Causey, & Cencula, 2016). The similar past study has further shown that mobile payment usage amongst USA millennials was higher than that of mobile banking generally. The likely motive of the MFS preference acknowledged in this study can be explained based on the significant issues of concern towards perceived privacy risk. Using mobile money transfer or mobile payment service do not necessarily involve consumers' personal information or an account that needs to be connected to a bank account. By that, lots of end-users would rather opt for mobile money transfer and mobile payment than for mobile banking accordingly.

7. CONCLUSIONS

The objective of this study was to examine the influence of both multidimensional trust and perceived risk facets at the individual level concurrently on the acceptance of mobile financial services (MFS) when prioritizing MFS perspective. This paper's goal is to illuminate, to some extent, the MFS accessibility in Togo allied with the potential facilitators or inhibitor factors. Also, to evaluate them based on the consumer's experience and experts through a benchmark robust SEM-TOPSIS methodology. A qualitative study in the context of the Togolese was performed together with a literature review to derive the most probable factors that might influence end-users' perception of MFS since there was a scarcity of research investigating trust and perceived risk antecedents. A quantitative study was then propelled to test the hypotheses formulated through the collected information obtained.

Our research model efficaciously integrates these dimensions, such as trust (dispositional trust, technological trust, and vendor trust), perceived risk (privacy, time, security, and cost) viewed as complex multidimensional factors. The data support the underlying assumption of the study except for H6 (see Figure 2). Mainly, our study is partially similar to the recent study done in Ghana (neighboring country of Togo) in which the perceived risk found to be related to the customer's trust in service providers regarding the adoption of mobile money (Abdul-Hamid, Shaikh, Boateng, & Hinson, 2019). In this line, our study provides more information to the various role-players of MFS about the necessity to emphasize on the trust and risk at the multidimensional level while making strategic and multicriteria decision-making.

Amongst the MFS alternative, the ranking result revealed mobile money to be the preferable MFS type used, followed by mobile payment and mobile banking with a minor percentage.

7.1 Implication for Practice

The outcomes of this study expose and validate the factors that impact consumers' adoption of MFS. Firstly, the relative level of the path coefficients in our analysis model recommended that disposition to trust (an antecedent of trust) be the most salient factor that facilitates either directly or indirectly the adoption of MFS. The perceived privacy risk (an antecedent of perceived risk) as the next influential factor, however, hinders MFS adoption. Given this trusting disposition is developed throughout a lifetime (Rotter, 1971) and reveals social impact over broad periods (Fukuyama, 1995), it implies that there might be a presence of a cross-cultural difference in trust. If so, MFS companies' providers must expect various levels of trust, and thus, different proportions of MFS adoption as well. As a deduction, companies are recommended to be acquainted in building trust-based tools and for instance, increasing awareness and firms' reputations by keeping their promises while treating the customer as individuals, mainly in societies which acknowledge exhibiting a lower level of trust. MFS service providers could meritoriously upsurge adoption behavior by publicizing the advantages of MFS to potential consumers, seeing that the findings supported trust with all of its antecedents.

Moreover, by modeling perceived risk with various facets, this study's finding imparts numerous risk effect concern. From this perspective, when companies propagandize their MFS services to ease the adoption issues, they should realistically underline a neutralizer or counter step for those risks' perceptions. The prominence of privacy risk and financial risk in the form of perceived cost as confirmed by this study and others prior research (José Liébana-Cabanillas, Sánchez-Fernández, & Muñoz-Leiva, 2014) signposts that customers still have doubts about the security of virtual transactions. For instance, these companies may stimulate a privacy risk protection strategy and grant technological support and anti-fraud to guarantee potential end-users minimal security risk. It is typical in the practice of emerging and developed countries (and it should be considered in developing countries as well) for payment service providers to try and promote trust in mobile financial services, in payments in general, as well as in other banking services by improving the general financial literacy of the population and small and medium enterprises. This is important, as those individuals who are familiar with financial processes and concepts demonstrate more trust towards financial services and can assess their risks better (Kovács & Terták, 2016). An increasingly popular practice is for certain governments to aid this process through an appropriate strategy and programs that serve the execution of that strategy. Because perceived time risk did not hold statistical significance in Togo, this phenomenon pinpoints that using MFS has little to do with time spent. As such, service providers should preserve those features that ease MFS application in time-frame.

Lastly, the outcome of TOPSIS through an SEM-TOPSIS integrated study specifies that mobile money transfer (MMT) is indeed the predominant mobile financial service (MFS) alternative used in Togo followed by mobile payment (MP), while mobile banking (MB) is reflected as trifling. In general, MFS companies should concede that consumer trust and risk with their antecedent create a tremendous barrier to MFS transactions. This study still demonstrates that amongst MFS companies, mobile money transfer companies are not powerless. It provides a practical guideline towards mobile financial service companies compared to the prevailing competitors within the related field such as online banking and ATM, for constructing more trust-based strategies to manipulate favorable consumer attitudes certainly, actions and eventual transaction behavior whereas mitigating the perceived risk factors. Regarding MFS, companies offering mobile money transfer are suggested to sustain the adoption growth, while those performing mobile payment, mobile banking predominantly, are to bear their target consumers at the core of the business model by diversifying market strategy.

With regard to the above, we cannot ignore the network nature of the payments market, an essential characteristic of which is that the market's dynamics (all the services provided and their prices) depend on the cooperation between many actors. Therefore the optimizing and maximizing effect of the traditional, individual competition on efficiency does not necessarily prevail by itself (Divéki et al., 2010). This may account for why collaboration between actors plays a positive and decisive role in improvements in the fields of trust and encountered risks.

7.2 Implication for Methodology and Theory

This research remains the first to assess the multi-dimensional trust and perceived risk facet concurrently towards consumers' adoption decision in mobile financial services while ranking their perspective.

The result obtained will open doors for scholars to explore further trust and perceived risk antecedents. It will support the theory of trust and risk literature in general, and IT in particular, since many prior studies lacked conclusive outcomes about the directivity of the causative relationship between trust and perceived risk (D. Gefen et al., 2003; Mayer et al., 1995; Rousseau et al., 1998). Our finding acknowledges the trust to be the potential predictor of risk in technology adoption. The scale items employed were greatly adopted from the prevailing studies in developed countries that are allied with technology acceptance adoption behavior, trust, and perceived risk. This section provides a crucial methodological implication for the marketing scholar, who might require a hint to cross-cultural appraisal concerning the application of scales, like those established in the USA and their relevance or relatedness in Togo. Our study outcomes not only enhance the clarification of mobile financial services adoption via the effect of trust and perceived risk but also hold some strategic implications for the global expansion of managerial implementation decision tools. This study provides a benchmark integrated methodology based on an SEM-MCDM application which found lacking in the adoption decision in general. The theoreticians and practitioners should comprehend that the prominence of the integrated SEM-TOPSIS is rooted in its robustness to test multifarious postulations made, combined with the high level in ranking the countless alternatives when multiple criteria issues arise in decision making.

7.3 Limitation and Future Research

Notwithstanding some contributions to the literature, practical, theoretical and methodological applications, all research unavoidably entails drawbacks that should be addressed. Our study outcomes are unique to Togo, although they are similar to IT in general and mobile financial transactions studies, predominantly. Preferably, a longitudinal study on our framework might need to gain a better understanding of how the variables relay over time. We expect future research will address these concerns. This research displays that time risk concerns are not significant antecedents of perceived risk. We hope that future research will further elucidate the relationship between time risk issues and adoption behavior in other populaces and circumstances.

Emphasizing multi-dimensional trust and perceived risk influences, this research projected to offer a wide-ranging still parsimonious decision-making model for MFS acceptance. However, the present model expounds only 13.1% of the variance in behavior to adopt. Future studies can incorporate additional variables, such as usefulness, perceived ease of use, and familiarity in an attempt to enhance the explanatory power. Based on the respondent's educational background, our distributed questionnaire appears to be limited to the more educated and technically competent elements of society, who would be more inclined to accept MFS applications. Therefore, researchers interested in MFS for adoption and sustainability should focus more on the underbanked population where illiterate people might be found in the majority.

REFERENCES

Abdul-Hamid, I. K., Shaikh, A. A., Boateng, H., & Hinson, R. E. (2019). Customers' Perceived Risk and Trust in Using Mobile Money Services—an Empirical Study of Ghana. *International Journal of E-Business Research*, *15*, 1–19.

Aboelmaged, M. G., & Gebba, R. T. (2013). Mobile Banking Adoption : An Examination of Technology Acceptance Model and Theory of Planned Behavior. *International Journal of Business Research and Development*, *2*, 35–50.

Abrahão, R., Moriguchi, S. N., & Andrade, D.-F. (2016). Intention of Adoption of Mobile Payment: An analysis in the light of the Unified Theory of Acceptance and Use of Technology (UTAUT). *RAI Revista de Administração e Inovação*, *13*, 221–230.

Abu-Shanab, E. A., & Abu-Baker, A. N. (2014). Using and buying mobile phones in Jordan: Implications for future research and the Development of New Methodology. *Technology in Society*, *38*, 103–110.

Adjei, J. K., & Odei-Appiah, S. (2018). Mobile financial services in emerging countries: Technology, adoption, and regulatory issues. In *Mobile Technologies and Socio-Economic Development in Emerging Nations* (pp. 109–128).

Afawubo, K., Agbagla, M., Couchoro, K. M., & Gbandi, T. (2017). *Socioeconomic determinants of the mobile money adoption process: the case of Togo*.

Afshan, S., Sharif, A., Waseem, N., & Frooghi, R. (2018). Internet banking in Pakistan: an extended technology acceptance perspective. *International Journal of Business Information Systems*, *27*, 383.

Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences*, *30*, 361–391.

Ajzen, I. (1985). From intentions to action: a theory of planned behavior. In *Action Control* (pp. 11–39).

Aktepe, A., Ersöz, S., & Toklu, B. (2015). Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling. *Computers & Industrial Engineering*, *86*, 95–106.

Al-Jabri, brahim M., & Sohail, M. S. (2012). Mobile banking adoption: Application of diffusion of innovation theory. *Journal of Electronic Commerce Research*, *13*, 379–391.

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management*, *37*, 99–110.

Aldridge, A., White, M., & Forcht, K. (1997). Security considerations of doing business via the Internet: cautions to be considered. *Internet Research*, *7*, 9–15.

Aloini, D., Martini, A., & Pellegrini, L. (2011). A structural equation model for continuous improvement: a test for capabilities, tools and performance. *Production Planning & Control: The Management of Operations*, *22*, 628–648.

Alsaad, A., Mohamad, R., & Ismail, N. A. (2017). The moderating role of trust in business to business electronic commerce (B2B EC) adoption. *Computers in Human Behavior*, *68*, 157–169.

Alzahrani, L., Al-Karaghouli, W., & Weerakkody, V. (2018). Investigating the impact of citizens' trust toward the successful adoption of e-government: A multigroup analysis of gender, age, and internet experience. *Information Systems Management*, *35*, 124–146.

Anderson, J. (2010). M-banking in developing markets: competitive and regulatory implications. *Info*, *12*, 18–25.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin*, *103*, 411–423.

Assadi，D., & Cudi，A. (2011). Le potentiel d'inclusion financiere du Mobile Banking. Une etude exploratoire. *Management et Avenir*, *6*, 227–243.

Assael, H. (1998). *Consumer behavior and marketing action* (6th ed.). Cincinnati, OH: South-Western College Publishing.

Bailey, B. P., Gurak, L. J., & Konstan, J. . (2002). *Trust in Cyberspace. In Human Factors and Web Development* (2nd ed.). J. Ratner. New Jersey: Lawrence Erlbaum Associates.

Barua, P. (2013). The Moderating Role of Perceived Behavioral Control : The Literature Criticism and Methodological Considerations. *International Journal of Business and Social Science*, *4*, 57–59.

Bauer, R. A. (1960). Consumer behavior as risk taking. In *Risk Taking and Information Handling in Consumer Behavior* (pp. 389–398).

Behzadian, M., Khanmohammadi Otaghsara, S., Yazdani, M., & Ignatius, J. (2012). A state-of the-art survey of TOPSIS applications. *Expert Systems with Applications*, *39*, 13051–13069.

Bellman, S., Lohse, G. L., & Johnson, E. J. (1999). Predictors of online buying behavior. *Communications of the ACM*, *42*, 32–38.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Bhattacherjee, A. (2002). Individual Trust in Online Firms: Scale Development and Initial test. *Journal of Management Information Systems*, *19*, 211–242.

Boksberger, P. E., Bieger, T., & Laesser, C. (2007). Multidimensional analysis of perceived risk in commercial air travel. *Journal of Air Transport Management*, *13*, 90–96.

Bollen, K. A. (1989). A New Incremental Fit Index for General Structural Equation Models. *Sociological Methods & Research*, *17*, 303–316.

Bonoma, T. V., & Johnston, W. J. (1979). Decision making under uncertainty: A direct measurement approach. *Journal of Consumer Research*, *6*, 177.

Browne, M. W., & Cudeck, R. (1993). *Alternative ways of assessing model fit* (In K. A. B). Newsbury Park, CA: Sage.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/windows*. Thousand Oaks, CA: Sage Publications.

Byrne, Bm M. (2010). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. *Routledge* (Vol. 22). doi:10.4324/9781410600219

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Chaix，Laetitia, & Torre，Dominique. (2015). Le double role du paiement mobile dans les pays en d´eveloppement. *Revue Economique*, *66*, 703–727.

Chakuthip, A., Brunetto, Y., Rod, F.-W., & Sheryl, R. (2007). Trust, Social Network and Electronic Commerce Adoption. *In Handbook of Research on Electronic Collaboration and Organizational Synergy*. Information Science Reference, Hershey. New York.

Chen, Lei da, Gillenson, M. L., & Sherrell, D. L. (2002). Enticing online consumers: An extended technology acceptance perspective. *Information and Management*, *39*, 705–719.

Chen, Leida, Meservy, T. O., & Gillenson, M. (2012). Understanding Information Systems Continuance for Information-Oriented Mobile Applications. *Communications of the Association for Information Systems*, *30*, 127–146.

Cheung, C., & Lee, M. K. O. (2000). Trust in Internet Shopping : A Proposed Model and Measurement Instrument. In *Americas Conference on Information Systems* (pp. 680–689).

Chin, W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*, 295–336.

Choffee, S. H., & McLeod, J. M. (1973). *Consumer decisions and information use, in Ward, S. and Robertson, T.S. (Eds), consumer behavior: theoretical sources*. Prentice-Hall Inc., Englewood Cliffs, NJ.

Chong, A. Y. L., Chan, F. T. S., & Ooi, K. B. (2012). Predicting consumer decisions to adopt mobile commerce: Cross country empirical examination between China and Malaysia. *Decision Support Systems*.

Cooper, D. R., & Schindler, P. S. (2003). *Business research methods*. *Business*. Retrieved from http://130.209.236.149/headocs/31businessresearch.pdf

Costigan, R. D., Ilter, S. S., & Berman, J. J. (1998). A Multi-Dimensional Study of Trust in Organizations. *Journal of Managerial Issues*, *10*, 303–316.

Couchoro, M. K. (2016). Challenges faced by MFIs in adopting Management information system during their growth phase: the case of Togo. *Enterprise Development and Microfinance*, *17*.

Coulter, R. a, Price, L. L., & Feick, L. (2003). Rethinking the Origins of invoivement and Brand Commitment : Insights from Postsociaiist Centrai Europe. *Journal of Consumer Research*, *30*, 151–169.

Cox, D. (1967). *Risk taking and information handling in consumer behavior*. *Risk Taking and Information Handling in Consumer Behavior* (Vol. 32). doi:10.2307/1249770

Cruz, P., Neto, L. B. F., Muñoz-Gallego, P., & Laukkanen, T. (2010). Mobile banking rollout in emerging markets: evidence from Brazil. *The International Journal of Bank Marketing*, *28*, 342–371.

Cunningham, L. F., Gerlach, J. H., Harper, M. D., & Young, C. E. (2005). Perceived risk and the consumer buying process: internet airline reservations. *International Journal of Service Industry Management*, *16*, 357–372.

Cunningham, S. M. (1967). *The major dimensions of perceived risk, in risk taking and information handling in consumer behavior*. *Consumer Behavior* (Harvard Un, Vol. 82–108).

Currim, I. S., & Sarin, R. K. (1983). A Procedure for Measuring and Estimating Consumer Preferences Under Uncertainty. *Journal of Marketing Research*, *20*, 249–256.

Davidov, E. (2008). A cross-country and cross-time comparison of the human values measurements with the second round of the European social survey. *Survey Research Methods*, *2*, 33–46.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*, 319–340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, *35*, 982–1003.

De Fouchier, R., & Larduinat, X. (2016). *Connected living: How technology could impact daily lives by 2025 – An International Study of Consumers Expectations for the Future of Mobile Technology*. Retrieved from http://www.gemalto.com/brochures-site/download-site/Documents/documentgating/tel-wp-connected-living-2025.pdf

Dennehy, D., & Sammon, D. (2015). Trends in mobile payments research: A literature review. *Journal of Innovation Management*, *3*, 49–61.

Dhull, S., & Narwal, M. S. (2018). Prioritizing the Drivers of Green Supply Chain Management in

Indian Manufacturing Industries Using Fuzzy TOPSIS Method: Government, Industry, Environment, and Public Perspectives. *Process Integration and Optimization for Sustainability*, *2*, 47–60.

Divéki, É., Keszy-Harmath, Z., & Helmeczi, I. (2010). *Innovative payment solutions, MNB surveys No. 85*. Budapest: Magyar Nemzeti Bank. Retrieved from https://www.mnb.hu/kiadvanyok/elemzesek-tanulmanyok-statisztikak/mnb-tanulmanyok/mnb-tanulmanyok/mt-85-diveki-eva-keszy-harmath-zoltanne-helmeczi-istvan-innovativ-fizetesi-megoldasok

Donovan, K. (2012). Mobile Money for Financial Inclusion. *Information and Communications for Development*. doi:http://dx.doi.org/10.3362/1755-1986.2011.016

Dowling, G. R., & Staelin, R. (1994). A Model of Perceived Risk and Intended Risk-Handling Activity. *Journal of Consumer Research*, *21*, 119.

El-Kasheir, D., Ashour, a, & Yacout, O. (2009). Factors affecting continued usage of internet banking among Egyptian customers. *Communications of the IBIMA*, *9*, 252–263.

Esearch, S. Y. R., & Koppius, O. R. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, *35*, 553–572.

Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: a perceived risk facets perspective. *International Journal of Human Computer Studies*, *59*, 451–474.

Figueira, J., Greco, S., & Ehrgott, M. (2005). Multiple criteria decision analysis: state of the art surveys. *Multiple Criteria Decision Analysis State of the Art Surveys*, *78*, 859–890.

Financial Afrik. (2015). Faible taux du «mobile banking» au Togo. Retrieved March 14, 2019, from http://news.alome.com/cc/24339.html

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: an introduction to theory and research*. *Reading, MA: Addison-Wesley* (Vol. 1–578). doi:10.1017/CBO9781107415324.004

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, *18*, 382–388.

Fox, C., Causey, E., & Cencula, D. (2016). *Study of Mobile Banking & Payments Third Edition August 2016*. Retrieved from http://www.firstannapolis.com/wp-content/uploads/2016/08/First-Annapolis-Study-of-Mobile-Banking-Payments-3rd-Ed_Aug-2016.pdf

Fukuyama, F. (1995). Trust: The Social Virtues and the Creation of Prosperity. *Trust the Social Virtues and the Creation of Prosperity*.

Gao, L., & Waechter, K. A. (2017). Examining the role of initial trust in user adoption of mobile payment services: an empirical investigation. *Information Systems Frontiers*, *19*, 525–548.

Gartner. (2004). Predicts 2005: banks to target underserved Customer segments. *Gartner Research*.

Gbongli, K. (2016). Integrating AHP-TOPSIS approach on prioritizing self-service technology (SST) decision making in financial institution (Togo). *British Journal of Mathematics & Computer Science*, *16*, 1–22.

Gbongli, K. (2017). A two-staged SEM-AHP technique for understanding and prioritizing mobile financial services perspectives adoption. *European Journal of Business and Management*, *9*, 107–120.

Gbongli, K., Csordas, T., & Kissi Mireku, K. (2017). Impact of consumer multidimensional online trust-risk in adopting Togolese mobile money transfer services: structural equation modelling approach. *Journal of Economics, Management and Trade*, *19*, 1–17.

Gbongli, K., Dumor, K., & Kissi Mireku, K. (2016). MCDM technique to evaluating mobile banking adoption in the Togolese banking industry based on the perceived value: perceived benefit and perceived sacrifice factors. *International Journal of Data Mining &*

Knowledge Management Process (IJDKP), *6*, 37–56.

Gbongli, K., Peng, Y., & Ackah, O. (2016). Selection and ranking of perceived risk associated with mobile banking in West Africa: An AHP Approach from customers' perspective. *International Journal of Scientific & Engineering Research*, *7*, 80–86.

Gbongli, K., Xu, Y., & Amedjonekou, K. M. (2019). Extended Technology Acceptance Model to Predict Mobile-Based Money Acceptance and Sustainability: A Multi-Analytical Structural Equation Modeling and Neural Network Approach. *Sustainability*, *11*, 1–33.

Gefen, D., Srinivasan Rao, V., & Tractinsky, N. (2003). The conceptualization of trust, risk and their electronic commerce: The need for clarifications. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, HICSS 2003, 6-9 January, Hawai, USA*. Big Island, USA. doi:10.1109/HICSS.2003.1174442

Gefen, David. (2000). E-commerce: the role of familiarity and trust. *Omega*, *28*, 725–737.

Gefen, David, Benbasat, I., & Pavlou, P. A. (2008). A Research Agenda for Trust in Online Environments. *Journal of Management Information Systems*, *24*, 275–286.

Gefen, David, Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated mode. *MIS Quarterly*, *27*, 51–90.

Gefen, David, & Straub, D. (2003). Managing User Trust in B2C e-Services. *E-Service Journal*, *2*, 7–24.

Gefen, David, Straub, D., & Boudreau, M. (2000). Structural equation modeling and regression: guidelines for research practice. *Communications of the Association for Information Systems*, *4*, 7.

Gefen, David, Straub, D., Mack, J., & Distinguished, R. (2000). The relative importance of perceived ease of use in IS adoption: a study of e-commerce adoption. *Journal of the Association for Information Systems*, *1*, 1–30.

George, J. F. (2004). The theory of planned behavior and Internet purchasing. *Internet Research*, *14*, 198–212.

Goodman, M. (2011). Corporate communication in times of stress. *Corp Commun An Int J*, *16*.

Grabner-Kräuter, S., & Kaluscha, E. A. (2003). Empirical research in on-line trust: a review and critical assessment. *International Journal of Human-Computer Studies*, *58*, 783–812.

Grandison, T., & Sloman, M. (2000). A survey of trust in internet applications. *IEEE Communications Surveys & Tutorials*, *3*, 2–16.

Guangming, Y., & Yuzhong, M. (2011). A research on the model of factors influencing consumer trust in mobile business. *E -Business and E -Government (ICEE), 2011 International Conference On*.

Gupta, K. P., Bhaskar, P., & Singh, S. (2017). Prioritization of factors influencing employee adoption of e-government using the analytic hierarchy process. *Journal of Systems and Information Technology*, *19*, 116–137.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2010). *Multivariate data analysis (7th Edition)*. *Uppersaddle River, New Jersey: Pearson Education International.* doi:10.1016/j.ijpharm.2011.02.019

Hair, J. F., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate data analysis*. *Pearson Prentice Hall.* (6th Editio, Vol. 6). Pearson Prentice Hall, Pearson Education, Inc., Upper Saddle River, New Jersey. doi:10.1080/19447013008687143

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis*. *Prentice Hall*.

Hallikainen, H., & Laukkanen, T. (2018). National culture and consumer trust in e-commerce. *International Journal of Information Management*. doi:10.1016/j.ijinfomgt.2017.07.002

Harrison McKnight, D., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *Journal of Strategic Information Systems*, *11*, 297–323.

Hassan Hosseini, M., Fatemifar, A., & Rahimzadeh, M. (2015). Effective factors of the adoption of mobile banking services by customers. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *6*, 1–13.

Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM*, *42*, 80–85.

Hosmer, L. T. (1995). Trust : the connecting link between organizational theory and philosophical ethics. *Academy of Management Review*, *20*, 379–403.

Hung, Y. H., Chou, S. C. T., & Tzeng, G. H. (2009). Using MCDM methods to adopt and assess knowledge management. In *Communications in Computer and Information Science* (Vol. 35, pp. 840–847).

Hwang, C.-L., & Yoon, K. (1981). *Multiple attribute decision making : methods and applications : a state-of-the-art survey* (Vol. 259). Berlin ; New York : Springer-Verlag, 1981.

Hwang, Y. (2014). User experience and personal innovativeness: An empirical study on the Enterprise Resource Planning systems. *Computers in Human Behavior*, *34*, 227–234.

Jack, W., Ray, A., & Suri, T. (2013). Transaction networks: Evidence from mobile money in Kenya. In *American Economic Review* (Vol. 103, pp. 356–361).

Jarvenpaa, S. L., & Tractinsky, N. (1999). Consumer trust in an Internet store: a cross cultural validation. *Journal of Computer Mediated Communications*, *5*, 1–5.

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer Trust in an Internet Store. *Information Technology and Management*, *1*, 45–71.

Johnson-George, C., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality & Social Psychology*, *43*, 1306–1317.

Johnson, V. L., Kiser, A., Washington, R., & Torres, R. (2018). Limitations to the rapid adoption of M-payment services: Understanding the impact of privacy risk on M-Payment services. *Computers in Human Behavior*, *79*, 111–122.

Johnston, Allen C. Warkentin, M. (2004). The Online Consumer Trust Construct: A Web Merchant Practitioner Perspective. In *SAIS 2004 Proceedings. 36* (pp. 220–226).

Jöreskog, K. G., & Sörbom, D. (1984). *Lisrel VI. analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Mooresville, Indiana: Scientific Software.

José Liébana-Cabanillas, F., Sánchez-Fernández, J., & Muñoz-Leiva, F. (2014). Role of gender on acceptance of mobile payment. *Industrial Management & Data Systems*, *114*, 220–240.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36.

Karahanna, E., Straub, D. W., Chervany, N. L., & Karahanna, E. (1999). Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly*, *23*, 183–213.

Kesharwani, A., & Bisht, S. S. (2012). The impact of trust and perceived risk on internet banking adoption in India: An extension of technology acceptance model. *International Journal of Bank Marketing*, *30*, 303–322.

Kikulwe, E. M., Fischer, E., & Qaim, M. (2014). Mobile money, smallholder farmers, and household welfare in Kenya. *PLoS ONE*, *9*. doi:10.1371/journal.pone.0109804

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*, 544–564.

Kim, K. K., & Prabhakar, B. (2000). Initial trust, perceived risk, and the adoption of internet banking. *Proceedings of the Twenty First International Conference on Information Systems*, 537–543.

Kim, K. K., & Prabhakar, B. (2004). Initial Trust and the Adoption of B2C e-Commerce: The Case of Internet Banking. *ACM SIGMIS Database*, *35*, 50–64.

Kogan, N., & Wallach, M. A. (1964). *Risk-taking: A study in cognition and personality* (New York:). Rhinehart & Winston.
Koller, M. (1988). Risk as a Determinant of Trust. *Basic and Applied Social Psychology*, *9*, 265–276.
Konovsky, M. A., & Pugh, S. D. (1994). Citizenship behavior and social exchange. *Academy of Management Journal*, *37*, 656–669.
Kovács, L., & David, S. (2016). Fraud risk in electronic payment transactions. *Journal of Money Laundering Control*. doi:10.1108/JMLC-09-2015-0039
Kovács, L., & Terták, E. (2016). *Financial Literacy, Panacea or Placebo - A Central European Perspective*. Retrieved from http://www.bankszovetseg.hu/Content/Publikaciok/Financial-literacy-20161122-Kovács.pdf
Kumar Mittal, V., & Singh Sangwan, K. (2014). Development of a structural model of environmentally conscious manufacturing drivers. *Journal of Manufacturing Technology Management*, *25*, 1195–1208.
Laforet, S., & Li, X. (2005). Consumers' attitudes towards online and mobile banking in China. *International Journal of Bank Marketing*, *23*, 362–380.
Lankton, N. K., Harrison McKnight, D., Wright, R. T., & Thatcher, J. B. (2016). Using expectation disconfirmation theory and polynomial modeling to understand trust in technology. *Information Systems Research*, *27*, 197–213.
Laukkanen, T. (2017). Mobile banking. *International Journal of Bank Marketing*, *35*, 1042–1043.
Lee, E., Kwon, K., & Schumann, D. W. (2005). Segmenting the non-adopter category in the diffusion of internet banking. *International Journal of Bank Marketing*, *23*, 414–437.
Lee, J. K., & Rao, H. R. (2007). Perceived risks, counter-beliefs, and intentions to use anti-/counter-terrorism websites: An exploratory study of government-citizens online interactions in a turbulent environment. *Decision Support Systems*, *43*, 1431–1449.
Lee, K. C., & Chung, N. (2009). Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified DeLone and McLean's model perspective. *Interacting with Computers*, *21*, 385–392.
Lee, M. C. (2008). Predicting Behavioural Intention to Use Online Banking. In *Proceedings of the 19th International Conference on Information Management*. Taiwan.
Lee, M. S. Y., McGoldrick, P. J., Keeling, K. A., & Doherty, J. (2003). Using ZMET to explore barriers to the adoption of 3G mobile banking services. *International Journal of Retail & Distribution Management*, *31*, 340–348.
Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? a critical review of the technology acceptance model. *Information and Management*, *40*, 191–204.
Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces*, *63*, 967–985.
Li, H., Sarathy, R., & Xu, H. (2011). The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. *Decision Support Systems*, *51*, 434–445.
Liang, T.-P., & Huang, J.-S. (1998). An empirical study on consumer acceptance of products in electronic markets : a transaction cost model. *Decision Support Systems*, *24*, 29–43.
Liao, Z., & Cheung, M. T. (2001). Internet-based e-shopping and consumer attitudes: An empirical study. *Information & Management*, *38*, 299–306.
Liébana-Cabanillas, F., Sánchez-Fernández, J., & Muñoz-Leiva, F. (2014). Antecedents of the adoption of the new mobile payment systems: The moderating effect of age. *Computers in Human Behavior*, *35*, 464–478.
Lin, C. T., & Tsai, M. C. (2010). Location choice for direct foreign investment in new hospitals in China by using ANP and TOPSIS. *Quality and Quantity*, *44*, 375–390.
Lin, H. F. (2011). An empirical investigation of mobile banking adoption: the effect of innovation

attributes and knowledge-based trust. *International Journal of Information Management*, *31*, 252–260.

Lippert, S. K., & Forman, H. (2006). A supply chain study of technology trust and antecedents to technology internalization consequences. *International Journal of Physical Distribution & Logistics Management*, *36*, 271–288.

Littler, D., & Melanthiou, D. (2006). Consumer perceptions of risk and uncertainty and the implications for behaviour towards innovative retail services: The case of Internet Banking. *Journal of Retailing and Consumer Services*, *13*, 431–443.

Liu, Z., Min, Q., & Ji, S. (2009). An empirical study on mobile banking adoption: The role of trust. In *2009 Second International Symposium on Electronic Commerce and Security* (Vol. 2, pp. 7–13).

Lowry, P. B. (2008). Explaining and Predicting the Impact of. *Journal of Management Information Systems*, 199–224.

Luarn, P., & Lin, H.-H. (2005). Toward an understanding of the behavioral intention to use mobile banking. *Computers in Human Behavior*, *21*, 873–891.

Luhmann, N. (1979). Trust and Power. *Cooperation without Trust*, *8*, 208.

Luo, X., Li, H., Zhang, J., & Shim, J. P. (2010). Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services. *Decision Support Systems*, *49*, 222–234.

Lyytinen, K., & Damsgaard, J. (2001). What ' s Wrong with the diffusion of innovation theory ? The case of a complex and networked technology. *Reports Aalborg University Department of Computer Science R985010*, *187*, 1–20.

Mahdevari, S., Shahriar, K., & Esfahanipour, A. (2014). Human health and safety risks management in underground coal mines using fuzzy TOPSIS. *Science of the Total Environment*, *488–489*, 85–99.

Manchiraju, S., Vudayagiri, G., & Garg, G. (2016). Capgemini Top 10 trends in payment in 2016. *Capgeminin Consulting Technology Outsourcing*, 1–26.

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, *14*, 81–95.

Marett, K., Pearson, A. W., Pearson, R. A., & Bergiel, E. (2015). Using mobile devices in a high risk context: The role of risk and trust in an exploratory study in Afghanistan. *Technology in Society*, *41*, 54–64.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*, 709–734.

McCord, M., & Ratnasingam, P. (2004). The Impact of Trust on the Technology Acceptance Model in Business to Consumer E-Commerce. *Idea Group Publishing*, 921–925.

Mcknight, D. H., Carter, M., & Thatcher, J. B. (2011). Trust in a Specific Technology: An Investigation of Its Components and Measures Trust in a specific technology: An investigation of its components and measures. *ACM Trans. Manag. Inform. Syst*, *2*. doi:10.1145/1985347.1985353

Mcknight, D. H., & Chervany, N. L. (2001). What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology. *International Journal of Electronic Commerce*, *6*, 35–59.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: an integrative typology. *Information Systems Research*, *13*, 334–359.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review*, *23*, 473–490.

Meng, D., Min, Q., & Li, Y. (2008). Study on trust in mobile commerce adoption - A conceptual model. In *Proceedings of the International Symposium on Electronic Commerce and Security, ISECS 2008* (pp. 246–249).

Min, Q., Meng, D., & Zhong, Q. (2008). An empirical study on trust in mobile commerce adoption. In *Proceedings of 2008 IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE/SOLI 2008* (Vol. 1, pp. 659–664).

Misiolek, N., Zakaria, N., & Zhang, P. (2002). Trust in organizational acceptance of information technology: A conceptual model and preliminary evidence. In *Proceedings of the Decision Sciences Institute 33rd Annual Meeting, November 23-26* (pp. 1–7). San Diego, California.

Mitchell, V. (1999). Consumer perceived risk: conceptualisations and models. *European Journal of Marketing*, *33*, 163–195.

MNB. (2019). *National Bank of Hungary's Competitiveness programme in 330 points*. Budapest: Magyar Nemzeti Bank. Retrieved from https://www.mnb.hu/letoltes/competitiveness-programme-in-330-points.pdf

Mohammadi, H. (2015). A study of mobile banking loyalty in Iran. *Computers in Human Behavior*, *44*, 35–47.

Mortenson, M. J., & Vidgen, R. (2016). A computational literature review of the technology acceptance model. *International Journal of Information Management*, *36*, 1248–1259.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, *39*, 429–460.

Muñoz-Leiva, F., Climent-Climent, S., & Liébana-Cabanillas, F. (2017). Determinants of Intention to Use The Mobile Banking Apps: An Extension of The Classic TAM Model. *Spanish Journal of Marketing - ESIC*, *21*, 25–38.

Murray, K. B., & Schlacter, J. L. (1990). The impact of services versus goods on consumers' assessment of perceived risk and variability. *Journal of the Academy of Marketing Science*, *18*, 51–65.

Napaporn, K. (2007). *Examining a Technology Acceptance Model of Internet Usage by Academics Within Thai Business Schools*. Business education.

Narteh, B., Mahmoud, M. A., & Amoh, S. (2017). Customer behavioural intentions towards mobile money services adoption in Ghana. *Service Industries Journal*, *37*, 426–447.

Nesse, P. J., Risnes, O., & Hallingby, H. S. (2018). Management of Mobile Financial Services—Review and Way Forward. In *In: Tan LM., Lau Poh Hock E., Tang C. (eds) Finance & Economics Readings* (Springer). Singapore.

Nilashi, M., Ibrahim, O., Reza Mirabi, V., Ebrahimi, L., & Zare, M. (2015). The role of Security, Design and Content factors on customer trust in mobile commerce. *Journal of Retailing and Consumer Services*, *26*, 57–69.

Nor, K. M., & Pearson, J. M. (2008). An Exploratory Study Into The Adoption of Internet Banking in a Developing Country: Malaysia. *Journal of Internet Commerce*, *7*, 29–73.

Oliveira, T., Faria, M., Thomas, M. A., & Popovič, A. (2014). Extending the understanding of mobile banking adoption: when UTAUT meets TTF and ITM. *International Journal of Information Management*, *34*, 689–703.

Ooi, K. B., & Tan, G. W. H. (2016). Mobile technology acceptance model: An investigation using mobile users to explore smartphone credit card. In *Expert Systems with Applications* (Vol. 59, pp. 33–46).

Park, S., & Tussyadiah, I. P. (2017). Multidimensional Facets of Perceived Risk in Mobile Travel Booking. *Journal of Travel Research*, *56*, 854–867.

Patil, S. K., & Kant, R. (2014). A fuzzy AHP-TOPSIS framework for ranking the solutions of Knowledge Management adoption in Supply Chain to overcome its barriers. *Expert Systems with Applications*, *41*, 679–693.

Pavlou, A. P. (2003). Consumer acceptance of electronic commerce : integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*, 69–103.

Pavlou, P., & Ratnasingam, P. (2001). Technology trust in B2B electronic commerce:

conceptual foundations. In *Business Strategies for Information Technology Management* (pp. 200–215). in Kangas, K. (Ed.), Business Strategies for Information Technology Management, Idea Group Publishing, Hershey, PA.
Pavlou, P.A. (2003). Consumer intentions to adopt electronic commerce – Incorporating trust and risk in the technology acceptance model. *International Journal of Electronic Commerce*, *7*, 101–134.
Pavlou, Paul A. (2001). Integrating trust in electronic commerce with the technology acceptance model : model development and validation. In *Proceedings of the Seventh Americas Conference in Information Systems AMCIS. New York, USA* (pp. 816–822).
Pavlou, Paul A, Liang, H., & Xue, Y. (2007). Understanding and Mitigating Uncertainty in Online Exchnge Relationships: A Principle - Agent Perspective. *MIS Quarterly*, *31*, 105–136.
Payne, R., & Clark, M. (2003). Dispositional and situational determinants of trust in two types of managers. *The International Journal of Human Resource Management*, *14*, 128–138.
Plouffe, C. R., Vandenbosch, M., & Hulland, J. (2001). Intermediating technologies and multi-group adoption: A comparison of consumer and merchant adoption intentions toward a new electronic payment system. *Journal of Product Innovation Management*, *18*, 65–81.
Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, *88*, 879–903.
Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, *32*, 122–136.
Punniyamoorty, M., Mathiyalagan, P., & Lakshmi, G. (2012). A combined application of structural equation modeling (SEM) and analytic hierarchy process (AHP) in supplier selection. *Benchmarking: An International Journal*, *19*, 70–92.
Ratnasingham, P., & Kumar, K. (2000). Trading partner trust in electronic commerce participation. In *ICIS* (pp. 544–552).
Rogers, E. M. (1962). *Diffusion of innovations*. *Newyork Free Press*. doi:citeulike-article-id:126680
Rossiter, C. M., & Barnett, P. W. (1975). *Communicating Personally: A Theory of Interpersonal Communication and Human Relationships*. Indianapolis : The Bobbs-Merrill Company, Inc., 1975.
Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist*, *26*, 443–452.
Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: a cross-discipline view of trust. *Academy of Management Review*, *23*, 393–404.
Safeena,Rahmath; Date,Hema;Kammani, Abdullaj; Hundewale, N. (2012). Technology Adoption and Indian Consumers: Study on Mobile Banking. *International Journal of Computer Theory and Engineering* , *4*, 1020–1024.
Safeena, R., Date, H., Hundewale, N., & Kammani, A. (2013). Combination of TAM and TPB in Internet Banking Adoption. *International Journal of Computer Theory and Engineering*, *5*, 146–150.
Salam, A. F., Iyer, L., Palvia, P., & Singh, R. (2005). Trust in e-commerce. *Communications of the ACM*, *48*, 72–77.
Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*, 344–354.
Schumacher, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. *A Beginner's Guide to Structural Equation Modeling*.
Schurr, P. H., & Ozanne, J. L. (1985). Influences on Exchange Processes: Buyers' Preconceptions of a Seller's Trustworthiness and Bargaining Toughness. *Journal of Consumer Research*.
Scornavacca, E., & Hoehle, H. (2007). Mobile banking in Germany: a strategic perspective.

International Journal of Electronic Finance, *1*, 304.
Scott, J. E., & Walczak, S. (2009). Cognitive engagement with a multimedia ERP training tool: Assessing computer self-efficacy and technology acceptance. *Information and Management*, *46*, 221–232.
Shaikh, A. A., & Karjaluoto, H. (2014). Mobile banking adoption: A literature review. *Telematics and Informatics*.
Shieh, L. F., Chang, T. H., Fu, H. P., Lin, S. W., & Chen, Y. Y. (2014). Analyzing the factors that affect the adoption of mobile services in Taiwan. *Technological Forecasting and Social Change*, *87*, 80–88.
Shih, H. S., Shyur, H. J., & Lee, E. S. (2007). An extension of TOPSIS for group decision making. *Mathematical and Computer Modelling*, *45*, 801–813.
Shipley, M. F., de Korvin, A., & Obid, R. (1991). A decision making model for multi-attribute problems incorporating uncertainty and bias measures. *Computers and Operations Research*, *18*, 335–342.
Siemsen, E., Roth, A., & Oliveira, P. (2009). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods*, *13*, 456–476.
Singleton, R. A., Straits, B. C., & Straits, M. M. (1993). *Approaches to social research* (2nd ed.). New Yor: New York: Oxford University Press.
Sjoberg, L. (1980). The risks of risk analysis. *Acta Psychologica*, *45*, 301–321.
Slade, E. L., Williams, M. D., & Dwivedi, Y. K. (2013). Mobile payment adoption: Classification and review of the extant literature. *The Marketing Review*, *13*, 167–190.
Slade, E., Williams, M., Dwivedi, Y., & Piercy, N. (2015). Exploring consumer adoption of proximity mobile payments. *Journal of Strategic Marketing*, *23*, 209–223.
Srivastava, S. C., Chandra, S., & Theng, Y.-L. (2010). Evaluating the Role of Trust in Consumer Adoption of Mobile Payment Systems: An Empirical Analysis. *Communications of the Association for Information Systems*, *27*, 561–588.
Straub, DetmarBoudreau, M., & Gefen, D. (2004). Validation Guidelines for IS Positivist Validation Guidelines for IS Positivist. *Communications of the Association for Information Systems*, *13*, 380–427.
Suh, B., & Han, I. (2003). The Impact of Customer Trust and Perception ofSecurity Control on the Acceptance ofElectronic Commerce. *International Journal of Electronic Commerce*, *7*, 135–161.
Swaminathan, V., Lepkowska-White, E., & Rao, B. P. (1999). Browsers or buyers in cyberspace? An investigation of factors influencing electronic exchange. *Journal of Computer-Mediated Communication*, *5*.
Tam, C., & Oliveira, T. (2017). Understanding mobile banking individual performance. *Internet Research*, *27*, 538–562.
Tan, F. B., & Sutherland, P. (2004). Online Consumer Trust: A Multi-Dimensional Model. *Journal of Electronic Commerce in Organizations*, *2*, 40–58.
Tan, M., & Teo, T. S. H. (2000). Factors Influencing the Adoption of Internet Banking. *Journal of the AIS*, *1*, 1–42.
Taylor, S., & Todd, P. A. (1995). Understanding information technology usage: a test of competing models. *Information Systems Research*, *6*, 144–176.
Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal Computing - toward a Conceptual-Model of Utilization. *Mis Quarterly*, *15*, 125–143.
Valaskova, K., Kramarova, K., & Bartosova, V. (2015). Multi Criteria Models Used in Slovak Consumer Market for Business Decision Making. *Procedia Economics and Finance*, *26*, 174–182.
Vidotto, G., Massidda, D., Noventa, S., & Vicentini, M. (2012). Trusting beliefs: A functional measurement study. *Psicologica*, *33*, 575–590.

Vinodh, S., Prasanna, M., & Hari Prakash, N. (2014). Integrated Fuzzy AHP-TOPSIS for selecting the best plastic recycling method: A case study. *Applied Mathematical Modelling*, *38*, 4662–4672.

Waite, K., & Harrison, T. (2015). Online banking adoption: We should know better 20 years on. *Journal of Financial Services Marketing*, *20*, 258–272.

Webster, J., & Martocchio, J. J. (1992). Microcomputer Playfulness: Development of a Measure with Workplace Implications. *MIS Quarterly*, *16*, 201.

Wells, J. D., Campbell, D. E., Valacich, J. S., & Featherman, M. (2010). The Effect of Perceived Novelty on the Adoption of Information Technology Innovations: A Risk/Reward Perspective. *Decision Sciences*, *41*, 813–843.

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84–136.

Wisdom, J. P., Chor, K. H. B., Hoagwood, K. E., & Horwitz, S. M. (2014). Innovation adoption: A review of theories and constructs. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*, 480–502.

Yan, H., & Yang, Z. (2015). Examining mobile payment user adoption from the perspective of trust. *International Journal of U-and e-Service, Science and Technology*, *8*, 117–130.

Yang, A. S. (2009). Exploring Adoption difficulties in mobile banking services. *Canadian Journal of Administrative Sciences*, *26*, 136–149.

Yang, S. (2016). Role of transfer-based and performance-based cues on initial trust in mobile shopping services: a cross-environment perspective. *Information Systems and E-Business Management*, *14*, 47–70.

Yang, S., Chen, Y., & Wei, J. (2015). Understanding consumers' web-mobile shopping extension behavior: A trust transfer perspective. *Journal of Computer Information Systems*, *55*, 78–87.

Yao, H., & Zhong, C. (2011). The analysis of influencing factors and promotion strategy for the use of mobile banking/L'analyse d'influencer des facteurs et la stratégie de promotion pour l'usage des opérations bancaires mobiles. *Canadian Social Science*, *7*, 60–63.

Yen, Y. S., & Wu, F. S. (2016). Predicting the adoption of mobile financial services: The impacts of perceived mobility and personal habit. *Computers in Human Behavior*, *65*, 31–42.

Yzer, M. (2012). Perceived Behavioral Control in Reasoned Action Theory: A Dual-Aspect Interpretation. *The ANNALS of the American Academy of Political and Social Science*, *640*, 101–117.

Zhang, L., Tan, W., Xu, Y., & Tan, G. (2012). Dimensions of perceived risk and their influence on consumers' purchasing behavior in the overall process of B2C. *In: Zhang L., Zhang C. (Eds) Engineering Education and Management. Lecture Notes in Electrical Engineering, Springer, Berlin, Heidelberg*, *111*, 1–11.

Zhou, T. (2012). Examining mobile banking user adoption from the perspectives of trust and flow experience. *Information Technology and Management*, *13*, 27–37.

Zhou, T., Lu, Y., & Wang, B. (2010). Integrating TTF and UTAUT to explain mobile banking user adoption. *Computers in Human Behavior*, *26*, 760–767.

Zikmund, G. W., & Scott, E. J. (1974). A Multivariate Analysis of Perceived Risk Self-Confidence and Information Sources. *In NA - Advances in Consumer Research Volume 01*.

Zolait, A. H. S. (2014). The nature and components of perceived behavioural control as an element of theory of planned behaviour. *Behaviour & Information Technology*, *33*, 1–21.

DECISION SCIENCES INSTITUTE

Boards' Corruption Culture and Bank Loan Contracts

Robin
Yuan Ze University
Email: robin@uib.ac.id

Chia-Wei Huang
National Taiwan Normal University
Email: d94723002@ntu.edu.tw

Chih-Yung Lin
National Chiao-Tung University
Email: d95723009@ntu.edu.tw

Huai-Chun Lo
Yuan Ze University
Email: hclo@saturn.yzu.edu.tw

ABSTRACT

We investigate the impacts of board's corruption culture on the financing costs of firms. Evidence shows that lending banks attach higher loan spreads, higher total costs of borrowing, and stricter covenants to firms with a strong corruption culture in their boards. The results are robust to controlling for the endogenous matching between firms and boards using a variety of econometric techniques. We further document that the effect of the board's corruption becomes stronger when firms have weak governance mechanisms. Our paper suggests that banks recognize board's culture of corruption as a sort of agency problem when they make lending decisions.

KEYWORDS: Corruption culture, Bank loan spread, Corporate governance, Financial cost, Agency problem

INTRODUCTION

Boards have an important role in a firm's corporate governance. Conventionally, boards perform monitoring and advisory roles (Schwartz-Ziv & Weisbach, 2013). The empirical evidence usually focuses on the structure of the board's composition, such as inside and outside directors, in which the directors are assumed to maximize shareholders' wealth. In contrast, a few recent studies have found several cases of inefficiency in the board structure. For example, the powerful CEOs are more likely to influence the selection of board members and invite related directors on to the board, which leads to weak monitoring by board directors (Fracassi & Tate, 2012; Coles et al, 2014). However, studies have not examined whether the board's culture of corruption influences the functions of board governance. Specifically, this paper investigates the effects of this corruption on the loan costs of firms.

The board of directors can be a critical function of board governance (Core et al, 1999; Fich & Shivdasani, 2006; Adams & Ferreira, 2009; Adams et al, 2010; Li & Srinivasan, 2011; Masulis et al, 2012; Baldenius et al, 2014; Coles et al, 2014; Kim et al, 2014). Boards with a strong level of corruption tend to be more tolerant of corrupt behavior, and thus the function of board governance is more likely to be invalid. This invalidity indicates that the culture of corruption will exacerbate the agency problem. The findings confirm this conjecture that boards' corruption culture is associated with firms' misbehavior (DeBacker et al, 2015; Liu, 2016).

In practical cases, some corporate scandals are related to the board's corruption. WorldCom disclosed US$3.8 billion in expenses, US$11 billion in inflated assets, and US$180 billion in losses for investors. This scandal was one of the biggest accounting frauds in the US and involved the CEO and outside directors. Accused of manipulating earnings and profits, the CEO was sentenced to 25 years in prison and outside directors were fined US$54 million that included US$18 million from their own pockets. Furthermore, the total fine was US$500 million and in 2002 WorldCom filed for Chapter 11 bankruptcy (The New York Times, 2002; Washington Post, 2005).

The Rigas family was a major stockholder of Adelphia Communications and held several top executive positions. The Rigas family hid US$2.3 billion in bank debt in accounting reports that they used for personal expenses. Once the case was uncovered, the CEO and the executives were charged 15 and 20 years in prison respectively, and the internal corruption forced Adelphia into bankruptcy in 2002 (Newton, 2010).

The banking literature finds that agency risk is the foremost consideration in the design of loan contracts (Rajan & Winton, 1995; Hart & Moore, 1998; Hart, 2001; DeMarzo & Fishman, 2007). The agency risk that lenders face results from the potential misbehavior of management, such as shirking, expropriation, entrenchment, myopia, and empire building, all of which not only reduce shareholders' wealth but also increase the probability of default (Sufi, 2007; Bharath et al, 2008; Graham et al, 2008). Accordingly, we hypothesize that lending banks charge higher loan spreads, have higher total borrowing fees, and attach stricter covenants to contracts if the borrowing firms have a board that displays a strong corruption culture.

To test our hypothesis, we measure the corruption culture of the firms' boards by using the average level of director corruption. For each director, we use his or her surname to identify the country of ancestry. Then, we match the country of ancestry with the Corruption Perception Index (CPI) to measure the level of their corruption.

We use a sample of 7,517 loans for US firms between 1996 and 2015. We find that the loan spread is positively associated with the level of corruption in the borrowing firm's board, which means that banks charge higher rates for firms with boards that have stronger corruption. The results are robust to various sets of control variables (firm characteristics, loan characteristics, macroeconomic variables, industry fixed effects, loan purposes, and loan types). Our main result is also economically important. For example, the average natural logarithm of the loan spread in our sample is 4.8143 (123.26 basis points). A one-standard-deviation increase in corruption results in an increase of 0.1176 that is about 15.38 basis points, which is around 12.48% of the average loan spread.

To address endogeneity concerns, we use four strategies. First, we construct a two-stage least square (2SLS) with an instrumental variable of political rights (Brunetti & Weder, 2003; Jong-sung & Khagram, 2005; Arezki & Bruckner, 2011). The 2SLS results show a positive relation between the corruption culture and the loan spread, which is consistent with our findings in the baseline regressions. Second, our results hold even if we use a firm fixed effect to control for the influence of unobservable omitted variables. Third, we use propensity score matching (PSM) to control for a self-selection bias (Rosenbaum & Rubin, 1983; Heckman et al, 1997; 1998). We sort firms into quartiles and take the top quartile as the treated firms and the bottom quartile as the control firms. Our evidence shows that the treated firms are charged a higher loan spread than the control firms.

Last, we use a quasi-natural experiment, the enactment of the Sarbanes-Oxley Act (SOX) of 2002, to generate a plausibly exogenous shock to board composition during our sample period (Coles et al, 2014; Balsmeier et al, 2017). Based on a difference-in-differences (hereafter, DiD) approach, we use the change ratio of the board's corruption culture to differentiate treated firms (top 50%) and control firms (bottom 50%) before SOX. We find that banks charged treated firms a higher loan spread after SOX that indicates that the boards' corruption culture is viewed as an agency concern for firms. Overall, our four identification tests confirm that the board's corruption culture has a positive causal impact on the bank's loan spread to the firm.

Furthermore, we investigate the cross-sectional variation in the effect of the board's corruption culture on loan contracting. We expect that the corruption culture is particularly crucial for firms with poor governance mechanisms. This is because the board's corruption culture should have a substitution effect with other governance mechanisms. In the literature, several studies have confirmed that the substitution effect exists among various governance mechanisms, such as the board of directors and severance pay (Almazan & Suarez, 2003), independent boards and board incentive pay (Adams et al, 2010), CEO ownership and external governance (Kim & Lu, 2011), external governance and weak internal governance (Fracassi & Tate, 2012), and independent boards and external governance mechanisms (Guo et al, 2015).

We then provide evidence from a conditional test based on strong and weak governance subsamples. The corruption culture leads to a firm's weak corporate governance that raises an agency problem between banks and firms. Based on the agency argument, the agency problems induced by the board's corruption culture should be deteriorated if the firm has weak governance mechanisms. This is because if the corruption culture is viewed as an agency problem, then the effect of the corruption culture on loan prices should be more significant when firms are associated with weak corporate governance.

We use several dummy variables to measure the corporate governance, such as institutional ownership, entrenchment index (Bebchuk et al, 2009), financial experts, and busy boards. Our evidence indicates that weak corporate governance increases the effect of the corruption culture, and hence banks punish these firms with higher loan spreads. This evidence supports our conjecture that banks consider stronger corruption as one type of agency cost when they make lending decisions.

We conduct several additional tests to corroborate our main results. First, banks also charge unfavorable non-price terms in loan contracts when the borrowing firms have boards with a strong corruption culture. Second, these firms are given low ratings. Third, we test the effect of the board's corruption culture with a combination of sin and politically sensitive industries (Hong & Kacperczyk, 2009; 2012). Because these industries are more likely to allow corrupt behaviors, the effect of the corruption culture should lessen. Fourth, we find similar results even if we define the corruption culture as coming from independent directors. Finally, we use alternative measures of the board's corruption culture (Mauro, 1998; Kaufmann et al, 2007). All results are consistent with our main finding.

Our paper contributes to the literature in several ways. First, it is related to the studies on failed governance mechanisms. For example, friendly boards monitor the CEO less and are not too independent (Adams & Ferreira, 2007). CEO's network connections with the board engage in more frequent value-destroying acquisitions and that connected independent directors provide less monitoring (Fracassi & Tate, 2012). Co-opted boards are more sympathetic to the CEO and not all independent directors are monitoring the effectiveness of the board (Coles et al, 2014). Directors with multiple directorships are less willing to relinquish their prestigious directorships (Masulis & Mobbs, 2014). Newly independent boards with newly appointed CEOs need to better assess their CEOs than CEOs from parent firms (Denis et al, 2015). Independent directors will depart before a bad event to protect their own reputation or to avoid work overload (Fahlenbrach et al, 2017). However, studies have ignored the role of the board's corruption culture in corporate governance. Our paper shows that weak governance mechanisms and an agency problem influence the firm's financing costs for private debt.

Second, we contribute to the literature on corporate culture. Previous studies have focused on cultural beliefs and societal organizations (Greif, 1994), culture and economic outcomes (Guiso et al, 2006), culture distance (Giannetti & Yafeh, 2012), culture of integrity and a firm's performance (Guiso et al, 2015), national culture and M&A (Ahern et al, 2015), domestic culture and stock price movements (Eun et al, 2015), unethical culture and corporate misbehavior (Biggerstaff et al, 2015), and corporate risk culture and investments (Pan et al, 2017). Our paper complements these studies by exploring how the board's corruption culture influences a firm's financing costs.

Third, our study contributes to the recent studies on the determinants of bank loan contracting, including shareholder rights (Chava et al, 2008), corporate misreporting (Graham et al, 2008), political connections (Houston et al, 2014), tax evasion (Hasan et al, 2014), private information (Carrizosa & Ryan, 2017), and social capital (Hasan et al, 2017). We complement this line of research by showing that the board's corruption culture has a significant impact on loan contracting.

LITERATURE REVIEW

Corruption Culture

Empirical studies have examined the impacts of corruption on corporate behaviors. Private monitoring and regulations force banks to disclose more information and this information can reveal the private gains of bank officials that improve the integrity of lending (Beck et al, 2006). Unpaid parking violations are based from diplomats from strong corruption countries and this evidence shows that social norms or cultural is strongly correlated with home country (Fisman & Miguel, 2007). Information sharing and bank competition reduces the corruption in lending (Barth et al, 2009). Beyond these findings, the structure of ownership and the legal environment play important roles in the level of corruption in bank lending. Strong corruption state is associated with higher default risks of credit and bond yield (Butler et al, 2009). However, firms with propensity to corrupt are significantly perform better than those without a propensity to corrupt firms (Mironov, 2015).

Outside of corruption, a couple of studies investigate bribes also harmful to corporate value. One of the reasons that firms spend on lobbying is to increase stimulus funds (Adelino & Dinc, 2014). The effect of spending on lobbying is to decrease market value because investors perceive lobbying as an unethical practice that increases the potential of conspiracy between firms and government officials (Borisov et al, 2016). Bribery activities reflect lower abnormal returns (Zeume, 2017). Fraudulent firms that lobby delay the detection of corporate fraud because they can allocate resources and sell their shares in the meantime (Yu & Yu, 2011). Firms typically pay bribes when they need something from government officials, for example, export-import services or public infrastructure services and innovative firms are more likely to pay bribes to public officials than non-innovative firms (Svensson, 2003; Ayyagari et al, 2014).

Corruption Culture and Agency Problem

Over the past decade, the conflicts of interest between managers and shareholders have persisted as corporate finance issues. In the banking industry, the loan officer may enjoy private benefits from corruption at the cost of borrowers and depositors because he or she approves the loans (Akins et al, 2017). Firm with a stronger culture of corruption displays an increasing opportunistic behavior and that the corrupt firm attracts a similar behavior in employees (Liu, 2016). State corruption is a part of political misbehavior that interacts with governments and financial institutions (Butler et al, 2009).

In our study, we are interested in whether the board's culture of corruption affects the firm's loan spread. We argue that this culture creates an agency problem. Moreover, the link between this culture and the loan spread provides evidence on the pricing of a bank's risk. Thus, we propose the following hypothesis:

Hypothesis 1. Firms with a stronger corruption culture in their boards are charged higher loan spreads*.*

Boards' Corruption Culture and Corporate Governance Roles

Most studies show that corporate governance bridges the gap between managers and stakeholders. Therefore, the agency problem is related to corporate governance mechanisms. Firms with a higher ownership concentration mitigate the agency conflicts (Jensen & Meckling, 1976; Mitton, 2002; Baek et al, 2004). The outside directors should increase the firm's reputation through better monitoring (Beasley, 1996; Beneish et al, 2017). A number of good corporate governance mechanisms can reduce the agency problem and increase the monitoring of firms (Lin et al, 2018).

From a different perspective, a weak corporate governance is associated with the value destruction that arises from the agency problem (Hoechle et al, 2012). A poor internal monitoring raises the loan spread (Kim et al, 2011).Thus, we propose our second hypothesis:

Hypothesis 2. The effect of the board's corruption culture on the loan spread is stronger for firms with weaker corporate governance.

METHODS

Data and Variables

We compute the corruption culture with the corruption perception index (CPI) as the main explanatory variable (DeBacker et al, 2015; Liu, 2016). Following previous studies, we measure the corruption culture of a firm with the average corruption culture of all directors (excluding the CEO) that we determine from the ancestral country of their surname. The corruption perception index (CPI) ranges from 0 (Denmark) to 9.6 (Bangladesh), where the United States index is 2.6.

We use surnames to capture the corruption culture for more than 187,000 boards of US public firms from Risk-Metrics. These data are from Origins Info Ltd. We hand-check their country of ancestry by using sources from ancestry.com if the surnames are unmatched. We are able to match 98.43% of the surname-ancestry countries and board names.

Loan data are collected from the DealScan database such as the loan spread and other non-price terms. The dependent variable is *Spread* (the natural logarithm of the amount of pay scaled by the London inter-bank offered rate). The other loan characteristics are *TCB* (the natural logarithm of loan spread plus other fees), *Maturity* (the natural logarithm of maturity in a month), *Loan Size* (the natural logarithm of the loan amount), *GenCov* (the natural logarithm of the general covenants), *FinCov* (the natural logarithm of the financial covenants), *Performance* (a dummy variable equal to one if the loan uses performance pricing and zero otherwise), and *Collateral* (dummy variable equal to one if the loan has collateral and zero otherwise).

We obtain the firm characteristics and macroeconomic control variables from Compustat, The Center for Research in Security Prices (CRSP), and Datastream. The firm characteristics are *Assets* (the natural logarithm of the total assets), *Market-to-Book* (total assets – book value of equity + price x common shares outstanding / total assets), *Leverage* (current liabilities plus long term debt scaled by the total assets), *Tangibility* (the sum of net property, plant, and equipment scaled by the total assets), *Profitability* (the earnings before interest, taxes, depreciation, and amortization scaled by the total assets), *Z-Score* (modified Altman's Z-score 1.2 x working capital + 1.4 x retained earnings + 3.3 x earnings before interest and taxes + 0.999 x sales scaled by the total assets), *Cash Flow Volatility* (standard deviation of quarterly cash flow from operation scaled by the total assets), and *Rating* (the Standard & Poor's debt rating that is converted into 0= no rating to 22=AAA (Klock et al, 2005; Jiraporn et al, 2014).

The macroeconomics variables are *Credit Spread* (the difference between the AAA corporate bond yield and the BAA corporate bond yield) and *Term Spread* (the difference between 10-year and 2-year Treasury yields).

Descriptive Statistics

Table 1 presents the descriptive statistics of the corruption culture, the loan spread, the total cost of borrowing (*TCB*), and the other control variables. The value of the corruption culture from 1996 to 2015 is positive, which is reasonable and consistent with the previous studies on the corruption culture. The mean value is 2.4702 with a standard deviation of 0.6083.

For firm characteristics, the mean value of *Assets* is US$4.416 million (8.3931 in natural logarithm). For loan characteristics, the mean value of *Spread* is 123.3 Bps or 1.233% (4.8143 in natural logarithm) and a mean of *TCB* is US$68.993 (4.2340 in natural logarithm). A loan contract in the sample has an average *Loan Size* of US$329 million (5.7957 in natural logarithm) with a mean *Maturity* of 41 months (3.6954 in natural logarithm). The average of the general covenants (*GenCov*) is 3.1918 (maximum of 11 covenants), and for the financial covenants (*FinCov*), it is 1.8784 (maximum of 7 covenants).

Table 1: Descriptive statistics

Variable	Mean	Q25	Median	Q75	Std Dev
BCC	2.470	2.013	2.454	2.864	0.608
Spread	4.814	4.318	5.011	5.416	0.813
TCB	4.234	3.531	4.202	4.884	0.929
Assets	8.393	7.288	8.272	9.456	1.541
Market-to-Book	0.678	0.614	0.693	0.768	0.127
Leverage	0.300	0.187	0.291	0.400	0.170
Profitability	0.134	0.090	0.124	0.165	0.069
Tangibility	0.317	0.126	0.254	0.479	0.232
Cash Flow Volatility	0.345	0.037	0.077	0.171	1.358
Z-Score	1.837	1.034	1.718	2.454	1.062
Loan Size	5.796	5.011	5.784	6.620	1.257
Maturity	3.695	3.584	4.094	4.094	0.699
Performance	0.474	0.000	0.000	1.000	0.499
Collateral	0.344	0.000	0.000	1.000	0.475
Gencov	3.192	2.000	4.000	6.000	0.760
Fincov	1.878	1.000	2.000	3.000	0.555
Credit Spread	-1.013	-1.110	-0.920	-0.830	0.316
Term Spread	-1.437	-2.510	-0.240	-0.020	2.091

This table reports the descriptive statistics for the main variables used in the study. The main sample consists of firm-year observations in the US from 1996 to 2015. All the data are winsorized at 1% and 99%. See the appendix for the variable definitions.

Table 2 presents the mean differences between firms with strong or weak corruption cultures. For firms with a stronger corruption culture, banks charge a higher loan spread, higher total cost of borrowing (*TCB*), require more *Collateral*, and attach a higher number of covenants (*GenCov*). Lower *Profitability* and *Tangibility* and higher *Market-to-Book* and *Cash Flow Volatility* are also observed in firms with a stronger corruption culture.

Table 2: Mean difference between strong and weak corruption culture

Variable	Board Corruption Culture			
	High	Low	Diff	*t*-Statistics
Spread	5.006	4.622	0.384***	21.06
TCB	4.426	4.122	0.304***	9.77
Assets	8.394	8.392	0.002	0.05
Market-to-Book	0.683	0.674	0.009***	2.94
Leverage	0.291	0.308	-0.017***	-4.25
Profitability	0.133	0.136	-0.003*	-1.91
Tangibility	0.295	0.339	-0.044***	-8.01
Cash Flow Volatility	0.402	0.291	0.111***	3.37
Z-Score	1.822	1.852	-0.030	-1.16
Loan Size	5.909	5.682	0.227***	7.88
Maturity	3.801	3.587	0.214***	13.29
Performance	0.456	0.492	-0.036***	-3.06
Collateral	0.357	0.331	0.026**	2.39
Gencov	1.220	1.102	0.118***	6.75
Fincov	0.638	0.623	0.015	1.22
Credit Spread	-1.032	-0.993	-0.039***	-5.45
Term Spread	-0.616	-2.257	1.641***	36.97

This table reports the mean and mean difference between firms with strong or weak corruption cultures. Superscripts *, **, and *** denote significance of the *t*-test for the difference in the means between the two subsamples at the 10%, 5%, and 1% levels, respectively.

Table 3 reports the Pearson's correlation coefficient matrix. We find a positive and significant coefficient for the correlation between *Spread* and the corruption culture at the 1% level. The coefficients for the correlation between the board's corruption culture and loan terms (*Loan Size*, *Maturity*, *Collateral*, *GenCov*, and *FinCov*) are also positive and significant at the 1% level. The coefficients for the correlations between the spread and firm characteristics (*Assets*, *Profitability*, *Leverage*, and *Tangibility*) are also negative and significant.

Table 3: Correlation matrix

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) *BCC*	1.00								
(2) *Spread*	0.25***	1.00							
(3) *TCB*	0.17***	0.87***	1.00						
(4) *Assets*	-0.04***	-0.23***	-0.24***	1.00					
(5) *Market-to-Book*	0.04***	0.15***	0.21***	-0.23***	1.00				
(6) *Leverage*	-0.05***	0.18***	0.30***	0.24***	0.24***	1.00			
(7) *Profitability*	-0.03**	-0.28***	-0.27***	-0.18***	0.05***	-0.16***	1.00		
(8) *Tangibility*	-0.11***	-0.02	0.02	0.16***	0.07***	0.23***	-0.01	1.00	
(9) *Cash Flow Volatility*	0.05***	0.01	-0.03*	-0.18***	-0.07***	-0.35***	0.08***	-0.10***	1.00
(10) *Z-Score*	0.01	-0.18***	-0.25***	-0.32***	-0.15***	-0.46***	0.49***	-0.35***	0.18***
(11) *Loan Size*	0.07***	-0.24***	-0.32***	0.59***	-0.12***	0.13***	0.02	0.06***	-0.12***
(12) *Maturity*	0.15***	0.28***	0.16***	-0.11***	0.09***	0.04***	-0.01	-0.03***	-0.01
(13) *Performance*	-0.03***	-0.04***	-0.11***	-0.09***	0.02**	-0.08***	0.06***	-0.05***	-0.01
(14) *Collateral*	0.05***	0.49***	0.58***	-0.25***	0.11***	0.15***	-0.18***	-0.07***	0.02*
(15) *Gencov*	0.09***	0.22***	0.25***	-0.27***	0.09***	-0.01	-0.02*	-0.08***	0.04***
(16) *Fincov*	0.03**	0.18***	0.23***	-0.32***	0.13***	-0.01	0.04***	-0.11***	0.04***
(17) *Credit Spread*	-0.11***	-0.21***	-0.25***	0.09***	-0.04***	0.08***	0.04***	0.01	-0.06***
(18) *Term Spread*	0.41***	0.38***	0.25***	0.05***	0.07***	-0.07***	-0.07***	-0.05***	0.01
Variable	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(10) *Z-Score*	1.00								
(11) *Loan Size*	-0.12***	1.00							
(12) *Maturity*	0.01	-0.02	1.00						
(13) *Performance*	0.07***	0.05***	0.06***	1.00					
(14) *Collateral*	-0.11***	-0.21***	0.24***	0.03***	1.00				
(15) *Gencov*	0.02	-0.02*	0.13***	0.50***	0.36***	1.00			
(16) *Fincov*	0.05***	-0.10***	0.13***	0.51***	0.31***	0.77***	1.00		
(17) *Credit Spread*	-0.01	0.06***	0.08***	-0.07***	-0.06***	-0.09***	-0.05***	1.00	
(18) *Term Spread*	-0.02*	0.13***	0.21***	0.01	0.09***	0.13***	0.06***	-0.42***	1.00

This table reports the Pearson correlation matrix between all variables. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

RESULTS AND DISCUSSION

Boards' Corruption Culture and Their Loan Spread

To examine the effect of the board's corruption culture on the loan spread, we follow the regression setting of previous studies (Hasan et al, 2014; 2017).

$$Spread_{i,t} = \alpha_0 + \alpha_1 BCC_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}' Z_{i,t} + v_k + \varepsilon_{i,t}, \quad (1)$$

Table 4: Boards' corruption culture and loan spreads

	(1)	(2)	(3)	(4)	(5)
Constant	4.067***	3.971***	5.958***	5.706***	6.700***
	(31.31)	(31.03)	(32.12)	(32.17)	(39.35)
BCC	0.224***	0.189***	0.195***	0.200***	0.118***
	(16.40)	(13.75)	(15.45)	(16.24)	(9.93)
SOX		0.253***	0.316***	0.290***	-0.646***
		(9.74)	(12.89)	(12.46)	(-16.83)
Assets			-0.177***	-0.093***	-0.096***
			(-24.77)	(-11.01)	(-12.11)
Market-to-Book			0.017	0.055	-0.048
			(0.24)	(0.81)	(-0.77)
Leverage			1.013***	0.833***	0.807***
			(16.96)	(14.14)	(14.79)
Profitability			-2.833***	-2.282***	-2.035***
			(-18.39)	(-15.45)	(-14.56)
Tangibility			-0.231***	-0.191***	-0.213***
			(-4.39)	(-3.79)	(-4.45)
Cash Flow Volatility			0.025***	0.017***	0.014**
			(4.26)	(3.16)	(2.48)
Z-Score			-0.055***	-0.031***	-0.045***
			(-4.48)	(-2.69)	(-4.17)
Loan Size				-0.071***	-0.070***
				(-8.23)	(-8.46)
Maturity				-0.054**	-0.005
				(-2.55)	(-0.22)
Performance				-0.071***	-0.064***
				(-4.44)	(-4.30)
Collateral				0.350***	0.329***
				(19.61)	(19.88)
GenCov				0.0001	-0.013
				(0.01)	(-0.90)
FinCov				0.064***	0.071***
				(3.06)	(3.69)
Credit Spread					-0.216***
					(-8.01)
Term Spread					0.207***
					(26.96)
Control for					
Industry FE	Yes	Yes	Yes	Yes	Yes

Loan Purpose	Yes	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes	Yes
Obs.	7,498	7,498	6,512	6,382	6,382
Adj. R^2	0.377	0.387	0.532	0.566	0.628

where the dependent variable $Spread_{i,t}$ is the natural logarithm of the loan spread for firm *i* in year *t*. The main independent variable is an average board's corruption culture (hereafter, *BCC*) in year *t-1*, which is calculated based on previous studies (DeBacker et al, 2015; Liu, 2016). SOX_t is a dummy variable equal to one if the year is 2002 or later and zero otherwise. $Firm_{i,t-1}$ is the characteristics of *Assets*, *Market-to-Book*, *Leverage*, *Tangibility*, *Profitability*, *Z-Score*, and *Cash Flow Volatility*. $Z_{i,t}$ is the characteristics of *Loan Size*, *Maturity*, *Performance*, *Collateral*, *GenCov*, and *FinCov* that are all measured in year *t*. The v_k and $\varepsilon_{i,t}$ represent the 2-digit SIC industry fixed effect and the error of the regression.

Table 4 shows the regression results. The coefficients are significantly positive in all models that indicate that firms with a stronger corruption culture are charged higher loan spreads (H1). The results are robust to several models when using the corruption culture as the sole explanatory variable or with various sets of control variables. The average natural logarithm of the loan spread in our sample is 4.8143 (=123.26 basis points). Thus, a one-standard-deviation increase in the board's corruption culture results in an increase of 0.1176 that is about 15.38 basis points ($e^{(4.8143+0.1176)} - e^{4.8143}$), which is around 12.48% (15.38 ÷ $e^{4.8143}$) of the average loan spread. Furthermore, 0.1176 is similar to the estimated coefficient for the dummy variable for repeat borrowers in Bharath, Dahiya, Saunders, and Srinivasan (2011) and for those for the politically connected firms in Houston et al. (2014). This result indicates that the board's corruption culture has not only statistically significant but also economically meaningful effects on banks' loan spreads.

In addition, firms with larger *Assets*, higher *Profitability*, higher *Tangibility*, and a better *Z-Score* tend to obtain a lower loan spread. However, firms with higher *Leverage* and higher *Cash Flow Volatility* are charged higher loan spreads.

Endogeneity Problems

The endogeneity problems can influence the effect of the corruption culture on the loan spread. The problems are reverse causality, measurement error, and omitted variables. First, reverse causality is the situation in which firm can improve the board's corruption culture to pursue a lower loan spread. To address the reverse causality, we estimating a two-stage least squares (2SLS) procedure (Griliches & Hausman, 1986; Berger & Hannan, 1998). Second, the coefficient becomes inconsistent if the corruption culture has a measurement error (Biddle & Hilary, 2006; Rajgopal & Shevlin, 2002; Chen et al, 2011). Third, if the board's corruption culture correlates with an omitted variable, then the estimator is inconsistent and biased and a regression with a firm fixed effect can eliminate this bias (Roberts & Whited, 2013).

<u>Two-Stage Regressions</u>

We eliminate endogeneity with the 2SLS with an instrumental variable.

$$BCC_{i,t} = \pi_0 + \pi_1 Political\ Rights_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + u_k + \varepsilon_{i,t}, \quad (2)$$

$$Spread_{i,t} = \alpha_0 + \alpha_1 \widehat{BCC} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + v_k + \varepsilon_{i,t}, \quad (3)$$

Table 5:2SLS and firm fixed effect				
	(1)	(2)		(3)
	Two-stage least square			Firm fixed effect
	Step 1	Step 2		*Spread*
	BCC	*Spread*		
Constant	1.346***	6.884***		5.095***
	(9.26)	(39.57)		(15.94)
Political Rights	1.329***			
	(38.06)			
$\widehat{BCC}$		0.047**		
		(2.02)		
BCC				0.232***
				(7.34)
SOX	-0.252***	-0.663***		-0.572***
	(-7.02)	(-17.21)		(-12.02)
Assets	-0.013*	-0.096***		0.050
	(-1.91)	(-12.28)		(1.56)
Market-to-Book	-0.182***	-0.057		-0.383**
	(-3.23)	(-0.91)		(-2.02)
Leverage	-0.028	0.804***		0.719***
	(-0.56)	(14.83)		(4.89)
Profitability	0.106	-2.026***		-1.053***
	(0.82)	(-14.56)		(-2.86)
Tangibility	-0.174***	-0.221***		-0.607***
	(-3.85)	(-4.67)		(-2.83)
Cash Flow Volatility	0.006	0.015***		0.022***
	(1.03)	(2.70)		(2.79)
Z-Score	-0.060***	-0.049***		-0.098**
	(-5.98)	(-4.54)		(-2.17)
Loan Size	0.022***	-0.068***		-0.061***
	(3.09)	(-8.35)		(-5.41)
Maturity	0.067***	-0.001		-0.032
	(4.10)	(-0.04)		(-1.14)
Performance	-0.076***	-0.070***		-0.064***
	(-5.22)	(-4.68)		(-3.06)
Collateral	-0.051***	0.326***		0.248***
	(-3.11)	(19.97)		(7.29)
GenCov	0.053***	-0.009		0.003
	(3.97)	(-0.60)		(0.12)
FinCov	-0.020	0.069***		0.017
	(-1.06)	(3.58)		(0.44)
Credit Spread	0.050**	-0.210***		-0.280***
	(2.49)	(-7.83)		(-8.39)
Term Spread	0.132***	0.211***		0.172***
	(19.67)	(26.63)		(19.10)
Control for				
Industry FE	Yes	Yes		No
Firm FE	No	No		Yes
Loan Purpose	Yes	Yes		Yes
Loan Type	Yes	Yes		Yes

Obs.	6,382	6,382		6,389
Adj. R^2	0.436	0.622		0.751

where *BCC* (Board Corruption Culture) is the average corruption values for all boards, and *Spread* is the natural logarithm of the amount of pay scaled by London inter-bank offered rate. *Political Rights* is the rights to political participation. Firm characteristics are measured in year *t-1*. *SOX* is a dummy variable equal to one if the year is 2002 or later and zero otherwise. Loan and macroeconomic characteristics are measured in year *t*.. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The first-stage regression, we use political rights (defined as the rights of political participation) as the instrumental variable by following previous studies (Brunetti & Weder, 2003; Jong-sung & Khagram, 2005; Arezki & Bruckner, 2011). They find that higher political rights are likely to have high agency problems. The data for political rights is obtained from Freedom House. The measurement of the political rights score is equal to an average of the boards' countries of ancestry. The political rights range from 1 to 7, where the lower score means more freedom. In the second-stage regression, we use the fitted value of the corruption culture from the first-stage regression to estimate the effect on the loan spread.

In column 1 of Table 5 presents the results of the first-stage regression of political rights on *BCC* and all other control variables. The coefficient for political rights is positive and significant. The Cragg-Donald Wald *F*-statistic is 1,936.187, which rejects the null hypothesis that the political rights are a weak instrumental variable. In the second-stage regression (in column 2), we replace the fitted value ($\widehat{BCC}$) from the first-stage regression. The coefficient for $\widehat{BCC}$ is still positive and significant for the loan spread. These results show that a stronger corruption culture leads a higher loan spread.

Regression with Firm Fixed Effect

We use a regression with a firm fixed effect to control for the influence of unobservable omitted variables.

$$Spread_{i,t} = \alpha_0 + \alpha_1 BCC_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + \mu_i + \varepsilon_{i,t}, \qquad (4)$$

In column 3 of Table 5, the coefficient for the corruption culture is positive and significant at the 1% level. After controlling for the firm fixed effect, the result still supports H1.

Propensity Score Matching

The hire of corrupt directors as board members could be self-selection. Even if we control for the firm fixed effect, there still is a concern for the self-selection bias. To solve the self-selection bias, we use propensity score matching (PSM) by following previous studies (Rosenbaum & Rubin, 1983).

We sort firms into quartiles and take the top quartile as the stronger corruption culture (treated firms) and the bottom quartile as the weakest corruption culture (control firms). We use a Probit regression on the top and bottom quartiles and include all control variables. For the robustness results, we use five different matching methods: (1) Nearest neighbor (n=1), (2) Mahalanobis, (3) Nearest neighbor (n=2), (4) Gaussian Kernel, and (5) Radius (1.0).

In Table 6, we show the differences between the treatment and control firms. In all models, the *t*-statistic's value is positive and significant at the 1% level, which confirms that firms with a stronger corruption culture in their boards are charged higher loan spreads. The results show that banks consider the stronger corruption culture as an agency problem and thus charge high loan spreads.

Table 6: Propensity score matching (PSM)

Matching method	Treatment	Control	Difference	*t*-Statistic
(1) Near neighbor (n=1)	5.0559	4.8390	0.2169***	2.79
(2) Mahalanobis	5.0591	4.5680	0.4911***	8.54
(3) Near neighbor (n=2)	5.0559	4.8223	0.2336***	3.29
(4) Kernel Gaussian	5.0559	4.7664	0.2895***	5.14
(5) Radius (0.1)	5.0559	4.7669	0.2890***	5.10

The dependent variable is *Spread* (the natural logarithm of the amount of pay scaled by London inter-bank offered rate). In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Difference-in-Differences Analysis

We use the Sarbanes-Oxley Act of 2002 (SOX) as an exogenous shock in a DiD analysis Coles et al, 2014; Balsmeier et al, 2017). We focus on the sample period 2001 – 2003 and calculate the change in the board's corruption culture. We use the change ratio to differentiate treated firms (top 50%) from the control firms (bottom 50%) before SOX. The sample size is 537 observations with 118 firms (55 treated firms and 58 control firms) after we use the PSM for selecting control firms. We use the same setting as in Equation (1) and run the OLS as follows:

$$Spread = D_{Tr}\times Y_{2002} + D_{Tr}\times Y_{2003} + D_{Tr} + Y_{2002} + Y_{2003} + \varepsilon, \quad (5)$$

Table 7: Difference-in-differences (DiD)

Panel A: Difference-in-differences matching analysis			
	Pre	Post	Difference (Post - Pre)
Control	4.429	4.445	0.016 (0.10)
Treated	4.106	4.572	0.466*** (3.68)
Difference			0.450**
			(2.31)
Panel B: Difference-in-differences regression			
		(1)	
Constant		4.431***	
		(47.70)	
$D^{Tr}\times Y_{2002}$		0.303	
		(1.53)	
$D^{Tr}\times Y_{2003}$		0.449**	
		(2.34)	
D^{Tr}		-0.324**	
		(-2.51)	
Y_{2002}		-0.191	
		(-1.33)	

Y_{2003}		0.015	
		(0.11)	
Obs.		537	
Adj. R^2		0.021	

where the dependent variable is *Spread*. Y_{2002} and Y_{2003} are time dummy variables. D_{Tr} is a treated group dummy that equals to one if the change ratio of the corruption culture is larger than the median and zero otherwise. $D_{Tr}{\times}Y_{2003}$ is a DiD estimator. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

In Table 7, the interaction of $D_{Tr}{\times}Y_{2003}$ is positive and significant for banks charging firms a higher loan spread if the firm's corruption culture on the board increases after SOX. Therefore, banks consider a stronger corruption culture is an agency problem.

The Moderating Effect of Corporate Governance

The empirical studies show that corporate governance mechanisms can reduce agency problems and increase monitoring roles. Firms with stronger institutional ownership are associated with lower loan spreads (Lin et al, 2018). In other ways, weak corporate governance destroys rather than creates value because of fewer financial experts and entrenchment index (Farber, 2005; Bebchuk et al, 2009). Following the previous studies, we use busier boards as a proxy for weak corporate governance (Hoechle et al, 2012). We then do a conditional test based on strong and weak governance subsamples. If the corruption culture is an agency problem, then its effect on loan prices should be more significant when firms are associated with weak corporate governance. We use several corporate governance measures such as institutional ownership (*IO_Low*), financial expertise (*Expert_Low*), entrenchment index (*E_Index_High*), and busy boards (*BB_High*). All of the corporate governance measures are defined in the appendix.

We set the dummy variables for corporate governance as follows: (1) firms with institutional ownership lower than the median; (2) firms with no financial expertise on the boards; (3) firms with an entrenchment index higher than the median; and (4) firms with board directors who serve on at least three boards. We set a dummy with these criteria equal to one and zero otherwise. We use the following regression to test the moderating effect of corporate governance:

$$Spread_{i,t} = \alpha_0 + \alpha_1 BCC_{i,t-1} + \alpha_2 BCC_{i,t-1} \times CG_{i,t-1} + \alpha_3 CG_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + v_k + \varepsilon_{i,t}, \quad (6)$$

Table 8 presents the interaction effect of corporate governance on the corruption culture. The coefficients in all models remain significant and support our expectations. In all models, the coefficients for the interaction terms between the corruption culture with institutional ownership (*IO_Low*), financial expertise (*Expert_Low*), entrenchment index (*E_Index_High*), and busy boards (*BB_High*) are positive and significant. These interaction terms confirm H2.

Table 8: The moderating effect of corporate governance

	(1)	(2)	(3)	(4)
Constant	7.029***	6.750***	7.217***	6.741***
	(35.76)	(39.51)	(41.58)	(39.32)
BCC	0.026	0.104***	-0.037**	0.097***
	(0.60)	(7.79)	(-2.30)	(6.50)
BCC × IO_Low	0.098**			

	(2.22)			
IO_Low	-0.349***			
	(-3.25)			
BCC × E_Index_High		0.048**		
		(2.04)		
E_Index_High		-0.145**		
		(-2.35)		
BCC × Expert_Low			0.085***	
			(3.45)	
Expert_Low			-0.678***	
			(-10.59)	
BCC × BB_High				0.043**
				(1.99)
BB_High				-0.109*
				(-1.92)
SOX	-0.639***	-0.649***	-0.405***	-0.643***
	(-16.58)	(-16.93)	(-10.45)	(-16.73)
Assets	-0.096***	-0.097***	-0.099***	-0.095***
	(-12.12)	(-12.22)	(-12.98)	(-11.79)
Market-to-Book	-0.038	-0.051	-0.106*	-0.047
	(-0.60)	(-0.81)	(-1.76)	(-0.75)
Leverage	0.801***	0.807***	0.845***	0.805***
	(14.67)	(14.78)	(16.02)	(14.70)
Profitability	-2.035***	-2.028***	-2.022***	-2.028***
	(-14.59)	(-14.52)	(-15.29)	(-14.47)
Tangibility	-0.222***	-0.212***	-0.215***	-0.215***
	(-4.63)	(-4.43)	(-4.66)	(-4.49)
Cash Flow Volatility	0.014**	0.014**	0.016***	0.014**
	(2.49)	(2.48)	(3.06)	(2.51)
Z-Score	-0.044***	-0.045***	-0.048***	-0.046***
	(-4.09)	(-4.17)	(-4.60)	(-4.23)
Loan Size	-0.069***	-0.070***	-0.082***	-0.070***
	(-8.42)	(-8.47)	(-10.25)	(-8.47)
Maturity	-0.003	-0.004	-0.035*	-0.005
	(-0.13)	(-0.18)	(-1.71)	(-0.22)
Performance	-0.064***	-0.064***	-0.036**	-0.065***
	(-4.28)	(-4.31)	(-2.46)	(-4.35)
Collateral	0.328***	0.328***	0.341***	0.330***
	(19.90)	(19.86)	(20.89)	(19.93)
GenCov	-0.012	-0.014	-0.019	-0.014
	(-0.84)	(-0.98)	(-1.33)	(-0.99)
FinCov	0.069***	0.072***	0.079***	0.072***
	(3.58)	(3.72)	(4.21)	(3.72)
Credit Spread	-0.206***	-0.218***	-0.264***	-0.217***
	(-7.60)	(-8.06)	(-10.26)	(-8.05)
Term Spread	0.205***	0.201***	0.108***	0.200***
	(26.97)	(27.05)	(12.99)	(26.84)
Control for				
Industry FE	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes

Loan Type	Yes	Yes	Yes	Yes
Obs.	6,382	6,382	6,382	6,382
Adj. R^2	0.630	0.628	0.658	0.628

where the dependent variable is *Spread* (the natural logarithm of the amount of pay scaled by London inter-bank offered rate). *BCC* (Board Corruption Culture) is the average corruption values for all boards. $CG_{i,t-1}$ is the corporate governance with institutional ownership (*IO_Low*) that equals one if no percentage of shares is held by institutional ownership and zero otherwise, e-index (*E-Index_High*) equals one if the Entrenchment Index is higher than the median and zero otherwise, financial experts (*Expert_Low*) equal one if no financial expertise exists on the boards and zero otherwise, and busy boards (*BB_High*) equal one if the a board director serves on at least three boards and zero otherwise. *SOX* is a dummy variable equal to one if the year is 2002 or later and zero otherwise. Firm characteristics are measured in year *t-1*. Loan and macroeconomic characteristics are measured in year *t*. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Additional Supporting Evidence

Boards' Corruption Culture, Non-Price Terms, and Firms' Rating

Following the previous studies, we use non-price terms (covenants) to test the effect of the board's corruption culture (Graham et al, 2008; Hasan et al, 2014; 2017). We use an ordinary least square (OLS) to construct the effect of *BCC* on non-price loan terms and the firm's rating as follows:

$$Y_{i,t} = \alpha_0 + \alpha_1 BCC_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + v_k + \varepsilon_{i,t}, \tag{7}$$

Table 9 shows the results. In columns 1 and 3, the coefficients for the corruption culture are positive and significant at the 1% level. These results show that firms with a stronger corruption culture are given more general and total covenants.

Table 9: Boards' corruption culture, non-price terms, and firm rating

	(1)	(2)	(3)	(4)
	GenCov	*FinCov*	*TotalCov*	*Rating*
Constant	1.246***	1.015***	1.577***	-0.614***
	(7.39)	(8.27)	(8.71)	(-5.92)
BCC	0.049***	0.009	0.056***	-0.051***
	(3.65)	(0.87)	(3.75)	(-5.59)
SOX	-0.030	-0.010	-0.031	0.048*
	(-0.76)	(-0.36)	(-0.70)	(1.90)
Assets	-0.095***	-0.080***	-0.112***	0.178***
	(-10.87)	(-12.82)	(-11.68)	(32.61)
Market-to-Book	-0.135**	0.020	-0.119*	-0.273***
	(-2.09)	(0.41)	(-1.67)	(-5.70)
Leverage	0.213***	0.197***	0.243***	0.432***
	(3.51)	(4.22)	(3.59)	(9.78)
Profitability	-0.084	-0.025	-0.126	0.438***
	(-0.58)	(-0.23)	(-0.80)	(4.33)
Tangibility	0.008	-0.085**	-0.016	-0.040
	(0.14)	(-2.11)	(-0.26)	(-1.05)

Cash Flow Volatility	0.007	0.003	0.008	-0.025***
	(1.25)	(0.62)	(1.31)	(-6.11)
Z-Score	-0.028**	-0.009	-0.026**	-0.006
	(-2.48)	(-0.98)	(-2.05)	(-0.78)
Loan Size	0.061***	0.017**	0.060***	0.006
	(6.34)	(2.55)	(5.84)	(1.16)
Maturity	-0.038*	-0.021	-0.048**	-0.008
	(-1.74)	(-1.42)	(-2.07)	(-0.73)
Performance	0.657***	0.503***	0.807***	0.008
	(42.7)	(43.63)	(47.52)	(0.66)
Collateral	0.459***	0.239***	0.474***	-0.007
	(24.26)	(16.84)	(23.24)	(-0.54)
GenCov				0.037***
				(3.77)
FinCov				-0.069***
				(-4.85)
Credit Spread	-0.019	0.032*	0.015	0.023*
	(-0.80)	(1.73)	(0.58)	(1.80)
Term Spread	0.013*	0.008	0.021**	-0.012**
	(1.72)	(1.44)	(2.44)	(-2.44)
Control for				
Industry FE	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes
Obs.	6,382	6,382	6,382	6,381
Adj. R^2	0.458	0.441	0.491	0.447

where the dependent variables of non-price terms and rating include general covenants (*GenCov*), financial covenants (*FinCov*), total covenants (*TotalCov*), and *Rating* is Standard & Poor's debt rating. *BCC* (Board Corruption Culture) is the average corruption values for all boards. *SOX* is a dummy variable equal to one if the year is 2002 or later and zero otherwise. Firm characteristics are measured in year *t-1*. Loan and macroeconomic characteristics are measured in year *t*. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

In addition, the credit rating is one good measure for firms' credit risk (Jiraporn et al, 2014). To test whether the corruption culture causes a lower credit rating, we use credit rating as the dependent variable in column 4 of Table 9. The coefficient for the corruption culture is negative and significant at the 1% level, which confirms that firms with a stronger corruption culture have a lower credit rating.

Specific Industries: Combination of Sin and Politically Sensitive Industries

We classify alcohol and gaming industries sin industries and tobacco, guns and defense, and natural resources industries as "politically sensitive industries" (Hong & Kacperczyk, 2009; 2012). We expect that the effect of the corruption culture should be weak in these industries. This weakness is because these industries are more likely to allow corrupt behaviors. Consistent with our expectation, we find insignificant and significant results for the corruption culture in specific industries and other industries (Panels A and B of Table 10).

Table 10: Boards' corruption culture in specific industries

Panel A:	(1)	(2)	(3)	(4)
Specific industries	*Spread*	*Rating*	*GenCov*	*FinCov*
Constant	6.403***	1.839***	1.742	2.207
	(4.17)	(3.50)	(0.94)	(1.53)
BCC	0.102	-0.042	0.116	-0.013
	(1.01)	(-1.13)	(1.30)	(-0.19)
Control for Control Variables	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes
Obs.	158	158	158	158
Adj. R^2	0.726	0.584	0.546	0.532
Panel B:	(5)	(6)	(7)	(8)
Other industries	*Spread*	*Rating*	*GenCov*	*FinCov*
Constant	6.719***	-0.630***	1.276***	1.017***
	(39.21)	(-5.97)	(7.51)	(8.24)
BCC	0.120***	-0.050***	0.045***	0.007
	(10.05)	(-5.33)	(3.27)	(0.68)
Control for Control Variables	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes
Obs.	6,224	6,223	6,224	6,224
Adj. R^2	0.628	0.435	0.459	0.440

This table presents the regression results for the effect of the board's corruption culture on loan spreads. *BCC* (Board Corruption Culture) is the average corruption value for all boards. Firm characteristics are measured in year *t-1*. Loan and macroeconomic characteristics are measured in year *t*. In panel A, following the study of Hong & Kostovetsky (2009, 2012), the specific industries are alcohol, gaming, tobacco, guns and defense, and natural resources. In Panel B, the other industries are all others except finance and the specific industries. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Robustness Checks

Alternative Measures of Corruption Culture

In this subsection, we use other corruption measures such as *CEO Corruption*, *Independent Corruption*, and *Insider Corruption* to examine the effect of the corruption culture. In columns 1 and 2 of Table 11, the results show that the corruption culture of board and CEO are significant for the loan spread. In columns 3, 4, 5, and 6, the corruption culture of board, independent boards, and insider boards are positive and significant for the loan after controlling for the *CEO Corruption*. Thus, our findings confirm that the firm's corruption culture comes from independent boards and insider boards but not from the CEO.

We also use two alternative measures of the corruption culture from the International Country Risk Guide (Mauro, 1998) and the Worldwide Governance Indicators (Kaufmann et al, 2007). In columns 7 and 8 of Table 11, the evidence shows that the board's corruption culture is positively and significantly related to the loan spread, which supports H1 again.

Table 11: Alternative measures of corruption culture

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	6.700***	6.898***	6.611***	6.657***	6.850***	6.614***	6.950***	6.710***
	(39.35)	(42.30)	(39.05)	(39.25)	(41.94)	(25.36)	(42.26)	(37.23)
BCC	0.118***		0.128***					
	(9.93)		(10.06)					
Independent Corruption				0.092***		0.091***		
				(7.61)		(4.89)		
Insider Corruption					0.041***	0.0390***		
					(4.61)	(2.85)		
CEO Corruption		0.014***	-0.001	0.006	-0.006	-0.012		
		(2.78)	(-0.07)	(1.25)	(-0.82)	(-1.06)		
BCC_WGI							0.006***	
							(4.39)	
BCC_ICRG								0.403***
								(3.38)
SOX	-0.646***	-0.661***	-0.640***	-0.643***	-0.658***	-0.641***	-0.673***	-0.669***
	(-16.83)	(-16.91)	(-16.43)	(-16.42)	(-16.78)	(-14.56)	(-17.44)	(-17.30)
Assets	-0.096***	-0.098***	-0.097***	-0.096***	-0.099***	-0.097***	-0.097***	-0.097***
	(-12.11)	(-12.24)	(-12.04)	(-11.98)	(-12.27)	(-7.42)	(-12.16)	(-12.21)
Market-to-Book	-0.048	-0.072	-0.062	-0.062	-0.082	-0.071	-0.057	-0.072
	(-0.77)	(-1.12)	(-0.97)	(-0.96)	(-1.27)	(-0.64)	(-0.91)	(-1.14)
Leverage	0.807***	0.841***	0.843***	0.841***	0.844***	0.842***	0.808***	0.780***
	(14.79)	(15.07)	(15.18)	(15.08)	(15.06)	(8.40)	(14.71)	(14.51)
Profitability	-2.035***	-1.994***	-2.012***	-2.019***	-1.998***	-2.024***	-2.033***	-2.013***
	(-14.56)	(-13.92)	(-14.19)	(-14.07)	(-13.98)	(-8.64)	(-14.43)	(-14.25)
Tangibility	-0.213***	-0.250***	-0.239***	-0.229***	-0.263***	-0.242***	-0.220***	-0.230***
	(-4.45)	(-5.14)	(-4.92)	(-4.71)	(-5.40)	(-2.85)	(-4.58)	(-4.80)
Cash Flow Volatility	0.014**	0.016***	0.014**	0.015***	0.0160***	0.015*	0.015***	0.016***
	(2.48)	(2.85)	(2.43)	(2.65)	(2.78)	(1.96)	(2.66)	(2.94)
Z-Score	-0.045***	-0.052***	-0.046***	-0.048***	-0.052***	-0.048***	-0.049***	-0.054***
	(-4.17)	(-4.72)	(-4.23)	(-4.33)	(-4.72)	(-2.75)	(-4.53)	(-4.97)
Loan Size	-0.070***	-0.068***	-0.070***	-0.070***	-0.068***	-0.070***	-0.068***	-0.067***
	(-8.46)	(-8.21)	(-8.43)	(-8.39)	(-8.13)	(-6.26)	(-8.24)	(-8.24)

Maturity	-0.005	0.006	-0.001	0.004	0.005	0.002	0.001	0.004
	(-0.22)	(0.29)	(-0.06)	(0.16)	(0.21)	(0.05)	(0.01)	(0.19)
Performance	-0.064***	-0.071***	-0.063***	-0.063***	-0.072***	-0.064***	-0.071***	-0.073***
	(-4.30)	(-4.64)	(-4.16)	(-4.14)	(-4.64)	(-3.14)	(-4.77)	(-4.87)
Collateral	0.329***	0.326***	0.329***	0.330***	0.325***	0.328***	0.325***	0.323***
	(19.88)	(19.61)	(19.68)	(19.75)	(19.45)	(13.62)	(19.66)	(19.58)
GenCov	-0.013	-0.007	-0.012	-0.009	-0.008	-0.010	-0.0089	-0.003
	(-0.90)	(-0.46)	(-0.82)	(-0.64)	(-0.51)	(-0.45)	(-0.61)	(-0.18)
FinCov	0.071***	0.063***	0.068***	0.067***	0.064***	0.067**	0.069***	0.065***
	(3.69)	(3.16)	(3.41)	(3.35)	(3.18)	(2.28)	(3.55)	(3.34)
Credit Spread	-0.216***	-0.206***	-0.215***	-0.215***	-0.206***	-0.216***	-0.205***	-0.218***
	(-8.01)	(-7.42)	(-7.79)	(-7.79)	(-7.44)	(-6.89)	(-7.55)	(-7.98)
Term Spread	0.201***	0.215***	0.200***	0.204***	0.212***	0.201***	0.214***	0.221***
	(26.96)	(28.72)	(26.49)	(26.66)	(28.22)	(22.15)	(29.12)	(30.02)
Control for								
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	6,382	6,206	6,176	6,179	6,167	6,140	6,382	6,382
Adj. R^2	0.628	0.626	0.631	0.629	0.627	0.631	0.623	0.622

This table presents the regression results for the robustness checks. In Models (1)–(8), the dependent variable is *Spread* (the natural logarithm of the amount of pay scaled by London inter-bank offered rate). *BCC* (Board Corruption Culture) is the average corruption values for all boards. *Independent Corruption* is the average corruption values for all independent directors. *Insider Corruption* is the average corruption values for the executives. *CEO Corruption* is the average corruption values for CEOs. *BCC_WGI* (The Worldwide Governance Indicators) follows Kaufmann et al, (2007), and *BCC_ICRG* (International Country Risk Guide) follows Mauro (1998). Firm characteristics are measured in year *t-1*. Loan and macroeconomic characteristics are measured in year *t*. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Total Cost of Borrowing (*TCB*)

Throughout the study, we use loan spread as the measure of the loan price. However, Berg et al., (2016) recently recommended that over 80% of US syndicated loans contained at least one of type fee, and contracts often specify a whole menu of spreads and types of fees. Therefore, we use their measure as another dependent variable in the robustness check: total cost of borrowing (*TCB*) as the dependent variable as follows:

$$TCB_{i,t} = \alpha_0 + \alpha_1 BCC_{i,t-1} + SOX_t + \boldsymbol{\beta}'\boldsymbol{Firm}_{i,t-1} + \boldsymbol{\theta}'\boldsymbol{Z}_{i,t} + v_k + \varepsilon_{i,t}, \quad (8)$$

Table 12: Total cost of borrowing (*TBC*)

	(1)	(2)	(3)	(4)	(5)
Constant	4.072***	4.018***	5.975***	6.375***	6.910***
	(19.20)	(18.91)	(22.12)	(26.94)	(31.63)
BCC	0.187***	0.177***	0.166***	0.159***	0.082***
	(10.57)	(9.87)	(11.03)	(12.00)	(6.51)
SOX		0.111***	0.186***	0.171***	-0.470***
		(4.07)	(7.47)	(7.85)	(-13.98)
Assets			-0.181***	-0.101***	-0.100***
			(-21.07)	(-9.99)	(-10.74)
Market-to-Book			0.170*	0.197**	0.105
			(1.93)	(2.55)	(1.44)
Leverage			1.139***	0.893***	0.846***
			(14.82)	(13.13)	(13.52)
Profitability			-2.679***	-1.809***	-1.579***
			(-14.95)	(-11.99)	(-11.07)
Tangibility			-0.109	-0.094	-0.105*
			(-1.64)	(-1.55)	(-1.88)
Cash Flow Volatility			0.027***	0.012**	0.009
			(3.79)	(1.98)	(1.36)
Z-Score			-0.084***	-0.050***	-0.068***
			(-5.53)	(-3.84)	(-5.62)
Loan Size				-0.043***	-0.040***
				(-4.20)	(-4.19)
Maturity				-0.362***	-0.292***
				(-15.36)	(-13.28)
Performance				-0.159***	-0.158***
				(-7.80)	(-8.30)
Collateral				0.489***	0.465***
				(22.43)	(23.17)
GenCov				-0.025	-0.035**
				(-1.29)	(-1.97)
FinCov				0.096***	0.106***
				(3.76)	(4.54)
Credit Spread					-0.209***
					(-7.29)
Term Spread					0.136***
					(19.73)
Control for					
Industry FE	Yes	Yes	Yes	Yes	Yes
Loan Purpose	Yes	Yes	Yes	Yes	Yes
Loan Type	Yes	Yes	Yes	Yes	Yes

Obs.	3,749	3,749	3,514	3,514	3,514
Adj. R^2	0.580	0.581	0.708	0.774	0.807

where the dependent variable is *TCB* (the natural logarithm of loan spread plus other fees). *BCC* (Board Corruption Culture) is the average corruption values for all boards. *SOX* is a dummy variable equal to one if the year is 2002 or later and zero otherwise. Firm characteristics are measured in year *t-1*. Loan and macroeconomic characteristics are measured in year *t*. In all models, the *t*-values are computed for the heteroskedasticity-robust standard errors. Coefficients: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 12 presents the regression results. The coefficients of *BBC* for *TCB* are positive and significant at the 1% level. The results thus support our hypothesis that firms with a stronger corruption culture have higher loan costs.

CONCLUSIONS

An increasing number of studies discuss the influences of corporate culture or corruption culture on corporate decisions. However, whether the board's corruption culture affects the firm's financing cost has not been investigated yet. Thus, this study examines whether this corruption culture influences the bank's loan terms by using 7,517 loans of US firms between 1996 and 2015. We use the average corruption culture of boards (excluding the CEO) by determining their ancestral country from their surname to measure the corruption culture of firms.

Our empirical results show that a stronger corruption culture increases the loan spread and total cost of borrowing. In addition, we use several methods to address endogeneity concerns. All the results support the main hypotheses. In further testing, our evidence indicates that the effect of the corruption culture on the costs of bank loans increases when firms have weak corporate governance. Such evidence indicates that (1) a stronger corruption culture is one kind of weak corporate governance for firms and (2) banks consider a stronger corruption culture as a type of agency problem when they make lending decisions.

We also conduct several tests and robustness checks to corroborate our main hypotheses. First, we find that the bank requires more covenants (especially general covenants) for firms with a stronger corruption culture. Second, our evidence shows that the stronger corruption culture of firms also reduces their ratings. Third, we find the effect of the corruption culture is less in specific industries that encourage more corrupt behaviors. Fourth, we find similar results even if we use alternative measures of the corruption culture (Mauro, 1998; Kaufmann et al, 2007).

Overall, our findings support the banks' view that the firms' with a strong corruption culture are an agency risk, which provides a reference for scholars, policy makers, and market investors for assessing the role of the corruption culture in regulating corporate governance.

REFERENCES

Adams, R. B., & Ferreira, D. (2007). A theory of friendly boards. *Journal of Finance*, *62*(1), 217-250.
Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291-309.
Adams, R. B., Hermalin, B. E., & Weisbach, M. S. (2010). The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature, 48*(1), 58-107.
Adelino, M., & Dinc, I. S. (2014). Corporate distress and lobbying: Evidence from the Stimulus Act. *Journal of Financial Economics*, *114*(2), 256-272.
Ahern, K. R., Daminelli, D., & Fracassi, C. (2015). Lost in translation? The effect of cultural values on mergers around the world. *Journal of Financial Economics*, *117*(1), 165-189.
Akins, B., Dou, Y., & Ng, J. (2017). Corruption in bank lending: The role of timely loan loss recognition. *Journal of Accounting and Economics*, *63*(2-3), 454-478.
Almazan, A., & Suarez, J. (2003). Entrenchment and severance pay in optimal governance structures. *Journal of Finance*, *58*(2), 519-547.
Arezki, R., & Brückner, M. (2011). Oil rents, corruption, and state stability: Evidence from panel data regressions. *European Economic Review*, *55*(7), 955-963.
Asquith, P., Beatty, A., & Weber, J. (2005). Performance pricing in bank debt contracts. *Journal of Accounting and Economics*, *40*(1-3), 101-128.
Ayyagari, M., Demirgüç-Kunt, A., & Maksimovic, V. (2014). Bribe payments and innovation in developing countries: Are innovating firms disproportionately affected?. *Journal of Financial and Quantitative Analysis*, *49*(1), 51-75.
Baek, J. S., Kang, J. K., & Park, K. S. (2004). Corporate governance and firm value: Evidence from the Korean financial crisis. *Journal of Financial Economics*, *71*(2), 265-313.
Baldenius, T., Melumad, N., & Meng, X. (2014). Board composition and CEO power. *Journal of Financial Economics*, *112*(1), 53-68.
Ball, R., Bushman, R. M., & Vasvari, F. P. (2008). The debt-contracting value of accounting information and loan syndicate structure. *Journal of Accounting Research*, *46*(2), 247-287.
Balsmeier, B., Fleming, L., & Manso, G. (2017). Independent boards and innovation. *Journal of Financial Economics*, *123*(3), 536-557.
Barth, J. R., Lin, C., Lin, P., & Song, F. M. (2009). Corruption in bank lending to firms: Cross-country micro evidence on the beneficial role of competition and information sharing. *Journal of Financial Economics*, *91*(3), 361-388.
Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *Accounting Review*, *71*(4), 443-465.
Bebchuk, L., Cohen, A., & Ferrell, A. (2008). What matters in corporate governance?. *Review of Financial Studies*, *22*(2), 783-827.
Beck, T., Demirgüç-Kunt, A., & Levine, R. (2006). Bank supervision and corruption in lending. *Journal of Monetary Economics*, *53*(8), 2131-2163.
Beneish, M. D., Marshall, C. D., & Yang, J. (2017). Explaining CEO retention in misreporting firms. *Journal of Financial Economics*, *123*(3), 512-535.
Berg, T., Saunders, A., & Steffen, S. (2016). The total cost of corporate borrowing in the loan market: Don't ignore the fees. *Journal of Finance*, *71*(3), 1357-1392.
Berger, A. N., & Hannan, T. H. (1998). The efficiency cost of market power in the banking industry: A test of the "quiet life" and related hypotheses. *Review of Economics and Statistics*, *80*(3), 454-465.
Bharath, S. T., Sunder, J., & Sunder, S. V. (2008). Accounting quality and debt contracting. contracting. *Accounting Review*, *83*(1), 1-28.

Bharath, S. T., Dahiya, S., Saunders, A., & Srinivasan, A. (2011). Lending relationships and loan contract terms. *Review of Financial Studies*, *24*(4), 1141-1203.
Biddle, G. C., & Hilary, G. (2006). Accounting quality and firm-level capital investment. *Accounting Review*, *81*(5), 963-982.
Biggerstaff, L., Cicero, D. C., & Puckett, A. (2015). Suspect CEOs, unethical culture, and corporate misbehavior. *Journal of Financial Economics*, *117*(1), 98-121.
Borisov, A., Goldman, E., & Gupta, N. (2016). The corporate value of (corrupt) lobbying. *Review of Financial Studies*, *29*(4), 1039-1071.
Bosch, O., & Steffen, S. (2011). On syndicate composition, corporate structure and the certification effect of credit ratings. *Journal of Banking and Finance*, *35*(2), 290-299.
Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, *87*(7-8), 1801-1824.
Butler, A. W., Fauver, L., & Mortal, S. (2009). Corruption, political connections, and municipal finance. *Review of Financial Studies*, *22*(7), 2873-2905.
Carrizosa, R., & Ryan, S. G. (2017). Borrower private information covenants and loan contract monitoring. *Journal of Accounting and Economics*, *64*(2-3), 313-339.
Chava, S., Livdan, D., & Purnanandam, A. (2008). Do shareholder rights affect the cost of bank loans?. *Review of Financial Studies*, *22*(8), 2973-3004.
Chen, H. J., Kacperczyk, M., & Ortiz-Molina, H. (2011). Labor unions, operating flexibility, and the cost of equity. *Journal of Financial and Quantitative Analysis*, *46*(1), 25-58.
Coles, J. L., Daniel, N. D., & Naveen, L. (2014). Co-opted boards. *Review of Financial Studies* *27*(6), 1751-1796.
Core, J. E., Holthausen, R. W., & Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*, *51*(3), 371-406.
DeBacker, J., Heim, B. T., & Tran, A. (2015). Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics, 117*(1), 122-138.
DeMarzo, P. M., & Fishman, M. J. (2007). Optimal long-term financial contracting. *Review of Financial Studies*, *20*(6), 2079-2128.
Denis, D. J., Denis, D. K., & Walker, M. D. (2015). CEO assessment and the structure of newly formed boards. *Review of Financial Studies*, *28*(12), 3338-3366.
Eun, C. S., Wang, L., & Xiao, S. C. (2015). Culture and R^2. *Journal of Financial Economics, 115*(2), 283-303.
Fahlenbrach, R., Low, A., & Stulz, R. M. (2017). Do independent director departures predict future bad events?. *Review of Financial Studies*, *30*(7), 2313-2358.
Farber, D. B. (2005). Restoring trust after fraud: Does corporate governance matter?. *Accounting Review*, *80*(2), 539-561.
Feldhütter, P., Hotchkiss, E., & Karakaş, O. (2016). The value of creditor control in corporate bonds. *Journal of Financial Economics*, *121*(1), 1-27.
Fich, E. M., & Shivdasani, A. (2006). Are busy boards effective monitors?. *Journal of Finance, 61*(2), 689-724.
Fisman, R., & Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, *115*(6), 1020-1048.
Fracassi, C., & Tate, G. (2012). External networking and internal firm governance. *Journal of Finance*, *67*(1), 153-194.
Giannetti, M., & Yafeh, Y. (2012). Do cultural differences between contracting parties matter? Evidence from syndicated bank loans. *Management Science*, *58*(2), 365-383.
Graham, J. R., Li, S., & Qiu, J. (2008). Corporate misreporting and bank loan contracting. *Journal of Financial Economics*, *89*(1), 44-61.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, *102*(5), 912-950.
Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, *31*(1), 93-118.
Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes?. *Journal of Economic Perspectives*, *20*(2), 23-48.
Guiso, L., Sapienza, P., & Zingales, L. (2015). The value of corporate culture. *Journal of Financial Economics*, *117*(1), 60-76.
Guo, L., Lach, P., & Mobbs, S. (2015). Tradeoffs between internal and external governance: evidence from exogenous regulatory shocks. *Financial Management*, *44*(1), 81-114.
Hart, O. (2001). Financial contracting. *Journal of Economic Literature*, *39*(4), 1079-1100.
Hart, O., & Moore, J. (1998). Default and renegotiation: A dynamic model of debt. *Quarterly Journal of Economics*, *113*(1), 1-41.
Hasan, I., Hoi, C. K. S., Wu, Q., & Zhang, H. (2014). Beauty is in the eye of the beholder: The effect of corporate tax avoidance on the cost of bank loans. *Journal of Financial Economics, 113*(1), 109-130.
Hasan, I., Hoi, C. K., Wu, Q., & Zhang, H. (2017). Social capital and debt contracting: Evidence from bank loans and public bonds. *Journal of Financial and Quantitative Analysis*, *52*(3), 1017-1047.
Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies, 64*(4), 605-654.
Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, *65*(2), 261-294.
Hoechle, D., Schmid, M., Walter, I., & Yermack, D. (2012). How much of the diversification discount can be explained by poor corporate governance?. *Journal of Financial Economics, 103*(1), 41-60.
Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15-36.
Hong, H., & Kostovetsky, L. (2012). Red and blue investing: Values and finance. *Journal of Financial Economics*, *103*(1), 1-19.
Houston, J. F., Jiang, L., Lin, C., & Ma, Y. (2014). Political connections and the cost of bank loans. *Journal of Accounting Research*, *52*(1), 193-243.
Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of Financial Economics*, *92*(2), 300-319.
Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305-360.
Jiraporn, P., Jiraporn, N., Boeprasert, A., & Chang, K. (2014). Does corporate social responsibility (CSR) improve credit ratings? Evidence from geographic identification. *Financial Management*, *43*(3), 505-531.
Jong-Sung, Y., & Khagram, S. (2005). A comparative study of inequality and corruption. *American Sociological Review*, *70*(1), 136-157.
Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007). The worldwide governance indicators project: Answering the critics. Working Paper, World Bank.
Kim, J. B., Song, B. Y., & Zhang, L. (2011). Internal control weakness and bank loan contracting: Evidence from SOX Section 404 disclosures. *Accounting Review*, *86*(4), 1157-1188.
Kim, E. H., & Lu, Y. (2011). CEO ownership, external governance, and risk-taking. *Journal of Financial Economics*, *102*(2), 272-292.

Klock, M. S., Mansi, S. A., & Maxwell, W. F. (2005). Does corporate governance matter to bondholders?. *Journal of Financial and Quantitative Analysis*, *40*(4), 693-719.
Li, F., & Srinivasan, S. (2011). Corporate governance when founders are directors. *Journal of Financial Economics*, *102*(2), 454-469.
Lin, C. Y., Tsai, W. C., & Hasan, I. (2018). Private benefits of control and bank loan contracts. *Journal of Corporate Finance*, *49*, 324-343.
Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Economics, 122*(2), 307-327.
Livingston, M., Poon, W. P., & Zhou, L. (2018). Are Chinese credit ratings relevant? A study of the Chinese bond market and credit rating industry. *Journal of Banking and Finance*, *87*, 216-232.
Masters, B. A., & Day. K. (2005). 10 ex-WorldCom directors agree to settlement. Retrieved from Washington Post, January 6th 2005.
Masulis, R. W., Wang, C., & Xie, F. (2012). Globalizing the boardroom—The effects of foreign directors on corporate governance and firm performance. *Journal of Accounting and Economics, 53*(3), 527-554.
Masulis, R. W., & Mobbs, S. (2014). Independent director incentives: Where do talented directors spend their limited time and energy?. *Journal of Financial Economics*, *111*(2), 406-429.
Mauro, P. (1998). Corruption and the composition of government expenditure. *Journal of Public Economics*, *69*(2), 263-279.
Mironov, M. (2015). Should one hire a corrupt CEO in a corrupt country?. *Journal of Financial Economics*, *117*(1), 29-42.
Mitton, T. (2002). A cross-firm analysis of the impact of corporate governance on the East Asian financial crisis. *Journal of Financial Economics*, *64*(2), 215-241.
Newton, M. (2010). *Crime and criminal*. Chelsea House Publishers.
Pan, Y., Siegel, S., & Wang, T. Y. (2017). Corporate risk culture. *Journal of Financial and Quantitative Analysis*, *52*(6), 2327-2367.
Rajan, R., & Winton, A. (1995). Covenants and collateral as incentives to monitor. *Journal of Finance*, *50*(4), 1113-1146.
Rajgopal, S., & Shevlin, T. (2002). Empirical evidence on the relation between stock option compensation and risk taking. *Journal of Accounting and Economics*, *33*(2), 145-171.
Roberts, M. R., & Whited, T. M. (2013). Chapter 7 - Endogeneity in Empirical Corporate Finance 1. In *Handbook of the Economics of Finance* Eds. Constantinides, G. M., In Harris, M. & Stulz, R. M. Elsevier.
Romero, S., & R. A. Atlas. 2002. WorldCom's collapse: The overview; WorldCom files for bankruptcy; Largest U.S. case. Retrieved from The New York Times, July 22th 2002.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
Schwartz-Ziv, M., & Weisbach, M. S. (2013). What do boards really do? Evidence from minutes of board meetings. *Journal of Financial Economics*, *108*(2), 349-366.
Sufi, A. (2007). Bank lines of credit in corporate finance: An empirical analysis. *Review of Financial Studies*, *22*(3), 1057-1088.
Svensson, J. (2003). Who must pay bribes and how much? Evidence from a cross section of firms. *Quarterly Journal of Economics*, *118*(1), 207-230.
Yu, F., & Yu, X. (2011). Corporate lobbying and fraud detection. *Journal of Financial and Quantitative Analysis*, *46*(6), 1865-1891.
Zeume, S. (2017). Bribes and firm value. *Review of Financial Studies*, *30*(5), 1457-1489.

DECISION SCIENCES INSTITUTE

Estimating the Net Worth of Nonprofit Universities

Yan He
School of Business
Indiana University Southeast
New Albany, Indiana
Email: yanhe@ius.edu

Frank Long
Wuhan University, Wuhan, China
Duke Kunshan University, Kunshan, China
Email: flong@whu.edu.cn; ffl3@duke.edu

ABSTRACT

The paper conducts a DCF valuation for Syracuse University and Indiana University. We find that the conversion of social benefit to net earnings and free cash flows adds value to equityholders. In the best scenario, the net worth is about 2 times the book equity.

KEYWORDS: DCF valuation, Nonprofit universities, Net worth, Social benefit

INTRODUCTION

Nonprofit universities in the U.S. distribute most earnings as scholarship and financial aid to students, pay zero tax, and incur zero cost for financing with donated equity capital. This poses two challenges to financial analyses of such institutions. First, it is difficult to estimate their net worth (market value of equity) due to their low levels of net earnings and free cash flows as well as zero cost of equity. Second, it is difficult to directly compare nonprofits with for-profits in terms of financial performance and net worth.

Recent literature has explored several corporate finance aspects of nonprofit institutions. Adelino, Lewellen, and Sundaram (2015) examine the investment-cash flow sensitivity for nonprofit hospitals and point out that it is similar to the sensitivity for shareholder-owned corporations. Jegers (2011a) shows the existence of financing constraints when nonprofits have no substantial opportunities to increase revenues from fundraising and their managers are unwilling to exert high fundraising efforts. Jegers (2011b) detects that tighter equity constraints are associated with more debts relative to total assets for Belgian nonprofits. Jegers and Verschueren (2006) argue that equity is cheaper than debt for nonprofits, and report that almost half of Californian nonprofits use no debts at all. Bøhren and Josefsen (2013) study the profitability of Norwegian banks and report that stockholder-owned banks do not outperform nonprofit ones. Kalodimos (2017) documents that stronger internal governance of nonprofit hospitals results in better performance as measured by the heart attack survival probability. Gertler and Kuan (2009) investigate the takeover market for hospitals and find that a nonprofit hospital tends to be sold at a lower price to a “like-minded” nonprofit buyer than to a for-profit buyer, based on the measure of Tobin’s *q*.

This paper conducts the Discounted Cash Flow (DCF) valuation of nonprofit universities, a process revealing the relationship between institutional fundamentals and value. Given their problems of low earnings and zero equity cost, we adjust nonprofit net earnings, free cash flows to equityholders, cost of equity, and tax rate to fit into the for-profit model. Hence, the nonprofit net worth can be estimated with the DCF approach. Two organizations, Syracuse University (SU) and Indiana University (IU), are studied in the paper. SU at Syracuse, New York is a nonprofit private organization that does not receive any government appropriations. IU at Bloomington, Indiana, with a few other branches located in the state, is a nonprofit public organization that receives state appropriations. For both, the current time refers to June 2017, and the 10-year historical period ranges from June 2007 to June 2017. The potential investors or acquirers are assumed to be for-profit publicly traded companies in the education industry. We collect the financial reports and other information of SU and IU from their public sources, and obtain the data of the education industry from the website of Professor Damodaran. Two DCF methods, Economic Profit to Equityholders (EP to Equityholders) and Free Cash Flow to Equityholders (FCFE), are applied in the valuation process. The forecasting of future earnings, cash flows, and discount rate is based on the subjective view of the for-profit acquirers. Such analysis does not apply to the situation in which the potential acquirers are governments or nonprofit organizations, since different types of investors hold different perspectives on the value of an acquisition target.

Our study yields several findings. First, we transform nonprofits to for-profits by adjusting a few relevant variables. The transformation generates a considerable market value added to nonprofit equityholders. In the best scenario, the net worth is about 2 times the book equity for both SU and IU. Second, with the net worth breaking even at the book equity, we find that about half of the existing social benefit needs to be converted to net earnings and free cash flows, while the rest is distributed to students. Third, when changing from a nonprofit to a for-profit status, the university would lose some public-and-government-supported fund revenues, and this potential loss would be offset by a potential increase in tuition revenues. A nonprofit private university seems more likely to achieve such a task than a nonprofit public university because the former relies much less on the supported fund revenues than the latter.

This paper attempts to make a few impacts on practitioners. First, it would be easier for government agencies to recognize the value-creating nonprofits in order to allocate funds. Second, it would be easier for public investors to identify the value-enhancing nonprofits in order to make donations. Third, it would be easier for corporations to understand the intrinsic value of nonprofits in order to conduct mergers and acquisitions. Fourth, it would be easier for nonprofits to compare themselves with for-profits.

DCF VALUATION PROCESS

We use two DCF methods to estimate the net worth, by following Pinto, Henry, Robinson, and Stowe (2015), Koller, Goedhart, and Wessels (2015), and Damodaran (2012). Based on the first method, the net worth is calculated as the current book equity plus the present value of expected future EPs to Equityholders. Based on the second method, the net worth is calculated as the present value of expected future FCFEs.

Nonprofit universities provide substantial social benefit, such as granting scholarship and financial aid to students each year. In theory, these organizations are tax free and they distribute all their profits to students as social benefit, leading to zero net earnings as well as zero return on equity. In addition, their equity financing comes from donations, incurring zero cost of equity. With zero return on equity and zero cost of equity, the annual economic profit to equityholders becomes zero, and the market value added to equityholders also becomes zero. Therefore, the net worth is equal to the book value of equity.

When a nonprofit university is purchased by a for-profit company, the tax payment becomes required, the social benefit distribution becomes non-required, and the cost of equity becomes larger than zero. In this transformation, the amount of social benefit distribution is converted to net earnings and free cash flows, generating a market value added to equityholders. The discount rate is defined as the cost of equity of the potential acquirer. The tax calculation is based on the tax rate of the acquirer.

Meanwhile, as it changes itself from a nonprofit to a for-profit organization, a university would lose some public-and-government-supported fund revenues, such as gifts (or contributions), grants and contracts, and state appropriations. We assume that the potential loss in these fund revenues would be offset by the potential increase in tuition revenues. That is, the total revenues would not be affected by this change.

Table 1. DCF valuation process

Panel A. All methods

Current time	June 2017
Past 10 years	June 2007 to June 2017
Future growth pattern	One-stage constant growth rate for the university
Future growth rate (g) First proxy	The first proxy is the geometric average growth rate of the university's adjusted net earnings in the past 10 years.
Future growth rate (g) Second proxy	The second proxy is the geometric average growth rate of the university's revenues in the past 10 years.
Future cost of equity (k)	Proxied by the average cost of equity of the education industry in the past 10 years
BE_0	Book equity of the university at the current time

Panel B. EP to Equityholders method

Adjusted net earnings	(Reported net earnings + Reported scholarship and financial aid*Social benefit converted in %) *(1-tax rate)
Adjusted ROE	Adjusted return on equity Calculated as the adjusted net earnings divided by prior book equity.
Future ROE	Proxied by the average adjusted ROE in the past 10 years.
EP_1	Economic profit to equityholders in future year 1 $EP_1 = BE_0$*(Future ROE – k)
NW_0	Net worth of the university at the current time $NW_0 = BE_0$ + Market value added, where Market value added = $EP_1/(k–g)$

Panel C. FCFE method

FCFE	FCFE = Cash flow from operations – Investment in fixed capital + Net borrowing
Adjusted FCFE	Adjusted FCFE = FCFE – Reported net earnings + Adjusted net earnings
Adjusted FOE	Adjusted FCFE on equity Calculated as the adjusted FCFE divided by prior book equity.
Future FOE	Proxied by the average adjusted FOE in the past 10 years.
$FCFE_1$	Free cash flows to equityholders in future year 1 $FCFE_1 = BE_0$*Future FOE
NW_0	Net worth of the university at the current time $NW_0 = FCFE_1 / (k–g)$

Table 1 presents the DCF valuation process. Panel A provides illustrations for all the methods used, Panel B for the EP to Equityholders method, and Panel C for the FCFE method.

In Panel A of Table 1, the future growth pattern is set up as one-stage constant growth because both the student enrollment and the total revenues of the two universities have climbed slowly in the past. It seems unnecessary to use a two-stage or three-stage pattern. The future growth rate (g) has two proxies: one is based on the average growth rate of the university's adjusted net earnings in the past 10 years, and the other, the average growth rate of the university's revenues in the past 10 years. The future cost of equity (k) is proxied by the average cost of equity of the education industry in the past 10 years.

In Panel B of Table 1, the future return on equity (ROE) is proxied by the average adjusted ROE in the past 10 years, and the adjusted ROE is estimated as the adjusted net earnings divided by the prior book equity. The adjustment of net earnings is made by adding the converted amount of social benefit to the reported net earnings, and by paying taxes. See Equation (1) below for the definition of the adjusted net earnings.

$$\begin{aligned}&\text{Adjusted net earnings}\\&= (\text{Reported net earnings}\\&\quad + \text{Reported scholarship and financial aid*Social benefit converted in \%})\\&\quad * (1\text{-tax rate}).\end{aligned} \tag{1}$$

Specifically, the reported net earnings refer to the increase in net assets, i.e., the bottom line of the income statement. The tax rate means the newly updated marginal tax rate (21%) instead of the old rate (35%), because the purpose of this calculation is to estimate the future ROE. Finally, the net worth at the current time (NW_0) is equal to the current book equity (BE_0) plus the market value added, where the market value added is equal to the projected economic profit to equityholders in future year1 (EP_1) divided by the difference between cost of equity and growth rate, and the economic profit (EP_1) is equal to the current book equity multiplied by the difference between the future ROE and the future cost of equity. See Equation (2) below for the net worth calculation.

$$NW_0 = BE_0 + \text{Market value added,} \tag{2}$$

where Market value added = $EP_1/(k–g)$, and $EP_1 = BE_0*(\text{Future ROE} – k)$.

In Panel C of Table 1, the future FCFE on equity (FOE) is proxied by the average adjusted FOE in the past 10 years, and the adjusted FOE is estimated as the adjusted FCFE divided by the prior book equity. The adjustment of free cash flows is made by subtracting the reported net earnings and adding the adjusted net earnings, as presented in Equation (3) below.

$$\text{Adjusted FCFE} = \text{FCFE} – \text{Reported net earnings} + \text{Adjusted net earnings,} \tag{3}$$

where FCFE = Cash flow from operations – Investment in fixed capital + Net borrowing. Lastly, the net worth at the current time (NW_0) is equal to the projected FCFE in future year1 ($FCFE_1$) divided by the difference between cost of equity and growth rate, as displayed in Equation (4) below.

$$NW_0 = FCFE_1 / (k–g), \tag{4}$$

where $FCFE_1 = BE_0*\text{Future FOE}$.

SUMMARY DATA OF ACQUIRERS AND TARGETS

For-profit publicly traded companies in educational services are regarded as potential acquirers. Table 2 provides summary data of education companies. The average cost of equity is 8.24%, which will be used as the future discount rate in valuing SU and IU. The average percentage of equity out of capital (based on market measures) is as high as 83.73%, implying a financing structure of light debt and heavy equity. We presume that both the cost of equity and the equity to capital ratio will remain constant in the future for the acquirers. Since the tax rate was reduced in January 2018, the new rate (21%) will be used to estimate the future ROE and the future FOE for SU and IU. The average growth in net income (7.51%), the average growth in revenues (7.68%), the average ROE (12.95%), and the average price/book ratio (3.59) are much higher than the current growth in net income (-10.92%), the current growth in revenues (-

3.90%), the current ROE (3.10%), and the current price/book ratio (1.95), respectively. The results suggest better performance historically than currently for education companies. The current price/book ratio (1.95) denotes that the net worth of education companies is nearly 2 times the book equity in 2017.

Table 2. Summary data of education companies

	Current year	10-year average	Future years
Cost of equity	9.42%	8.24%	8.24%
Equity/Capital	74.97%	83.73%	
Marginal tax rate	35%	35%	21%
Historical 5-year average growth in net income	-10.92%	7.51%	
Historical 5-year average growth in revenues	-3.90%	7.68%	
ROE	3.10%	12.95%	
Price/Book ratio	1.95	3.59	

SU and IU are treated as acquisition targets. Table 3 provides summary data of the targets. Panel A reports the current-year data, and Panel B, the 10-year-average data. In Panel A of Table 3, we note that the two universities are large and comprehensive in general. Specifically, SU is smaller than IU in terms of student enrollment (21,970 vs. 94,698), total revenues ($1.0 billion vs. $3.2 billion), total assets ($2.9 billion vs. $5.5 billion), and book equity ($2.0 billion vs. $3.9 billion). The per-student tuition and fees are much higher for SU as a private university than for IU as a public university ($39,907 vs. $15,337).

In Panel B of Table 3, first, we find that both the student enrollment and the total revenues of the two universities grow slowly in the past 10 years. The growth rate of student enrollment is 1.42% for SU and 0.24% for IU. The growth rate of revenues is 2.43% for SU and 2.81% for IU.

Second, pertaining to the income statement, the student tuition and fee revenues out of total revenues are 84.62% for SU vs. 41.86% for IU. The public-and-government-supported fund revenues out of total revenues, including gifts (or contributions), grants and contracts, and state appropriations, are 17.59% for SU vs. 41.82% for IU. Additionally, these fund revenues out of student tuition and fees are 20.89% for SU vs. 100.86% for IU. Therefore, SU relies much less on supported fund revenues than IU. It is noted that SU as a private university receives zero state appropriations. IU as a public university receives considerable amount of state appropriations, which are 19% out of total revenues and 45.99% out of student tuition and fees. The scholarship and financial aid costs out of total revenues are 31.21% for SU vs. 11.99% for IU; and these costs out of student tuition and fees are 36.87% for SU vs. 28.59% for IU. That is to say, SU distributes more social benefit to students than IU, since the former charges much higher per-student tuition and fees than the latter. The instructional costs out of total revenues are similar for the two universities (38.47% vs. 34.34%).

Third, with regard to the balance sheet, the asset turnover (total revenues divided by total assets) is 34.39% for SU vs. 61.43% for IU, suggesting that IU is more efficient in asset usage. The non-current assets contain two major items: the investments which are financial securities such as stocks and bonds, and the fixed assets which are long-term physical assets such as land, land improvements, buildings, and equipment. The investments out of total assets are 45.30% for SU vs. 26.45% for IU, and the fixed assets out of total assets are 39.77% for SU vs. 55.05% for IU. Thus, SU endures more risk exposure to financial markets. The capital structure measures are similar to those of the education companies, including the long-term debts to total assets ratio (15.75% for SU vs. 19.04% for IU), the book equity to total assets ratio

(69.05% for SU vs. 67.14% for IU), and the book equity to capital ratio (81.46% for SU vs. 76.67% for IU). Thus, SU and IU have a capital structure of light debt and heavy equity.

Table 3. Summary data of universities

Panel A. Current-year data (2017)

	Syracuse University	Indiana University
Enrollment of students	21,970	94,698
Total revenues ($million)	995	3,193
Net earnings ($million) (unadjusted)	198	154
Total assets ($million)	2,852	5,474
Book equity ($million)	2,018	3,866
Student tuition and fees per student ($)	39,907	15,337
Panel B. 10-year-average data (2007-2017)		
Annual enrollment growth	1.42%	0.24%
Annual revenue growth	2.43%	2.81%
Student tuition and fees/Total revenues	84.62%	41.86%
Gifts (or contributions)/Total revenues	7.65%	3.49%
Grants and contracts/Total revenues	9.94%	19.33%
State appropriations/Total revenues	0.00%	19.00%
Scholarship & financial aid/Total revenues	31.21%	11.99%
Instructional costs/Total revenues	38.47%	34.34%
Net earnings/Total revenues (unadjusted)	1.88%	6.45%
Gifts (or contributions)/Student tuition and fees	9.16%	8.33%
Grants and contracts/Student tuition and fees	11.73%	46.54%
State appropriations/Student tuition and fees	0.00%	45.99%
Scholarship & financial aid/Student tuition and fees	36.87%	28.59%
Total revenues/Total assets	34.39%	61.43%
Investments/Total assets	45.30%	26.45%
Fixed assets/Total assets	39.77%	55.05%
Long-term debts/Total assets	15.75%	19.04%
Book equity/Total assets	69.05%	67.14%
Book equity/Capital	81.46%	76.67%
ROE (unadjusted)	1.99%	6.42%
FOE (unadjusted)	-1.50%	1.84%

Finally, the net earnings, the net margin (net earnings divided by total revenues), the ROE, and the FOE are unadjusted for social benefit conversion. It is noted that the unadjusted 10-year average ROE is 1.99% for SU and 6.42% for IU, which are much lower than the average ROE of education companies (12.95%). The unadjusted 10-year average FOE is -1.50% for SU and 1.84% for IU, which are also quite low. Since SU and IU dispense some profits as social benefit each year, they end up with low levels of net earnings and free cash flows to equityholders as well as low ROE and FOE ratios. If the unadjusted average ROE and FOE are used to forecast future net earnings and free cash flows, and if the for-profit cost of equity (8.24%) is used as the discount rate, the estimated net worth would be much lower than the book equity.

VALUATION RESULTS: FULL CONVERSION OF SOCIAL BENEFIT

The analysis in this section is based on the best-scenario assumptions summarized as follows. First, the existing social benefit is fully converted to net earnings and free cash flows, i.e., 100% of scholarship and financial aid are not distributed to students any longer; instead, they are converted to profits. Second, the potential loss of the public-and-government-supported fund revenues is offset by the potential increase in tuition revenues. That is to say, the total revenues remain the same.

Table 4. Valuation results: full conversion of social benefit

Panel A. Inputs and assumptions

	Syracuse University		Indiana University	
	Earnings growth	Revenues growth	Earnings growth	Revenues growth
Future growth rate (g)	3.74%	2.43%	4.38%	2.81%
Future cost of equity (k)	8.24%	8.24%	8.24%	8.24%
BE_0 ($million)	2,017.59	2,017.59	3,865.54	3,865.54
% of social benefit converted	100.00%	100.00%	100.00%	100.00%
Panel B. EP to Equityholders method				
Future ROE	14.12%	14.12%	14.22%	14.22%
EP_1 ($million)	118.78	118.78	231.25	231.25
NW_0 ($million)	4,658.39	4,065.13	9,868.44	8,130.68
NW_0/BE_0	2.31	2.01	2.55	2.10
Panel C. FCFE method				
Future FOE	10.62%	10.62%	9.63%	9.63%
$FCFE_1$ ($million)	214.36	214.36	372.33	372.33
NW_0 ($million)	4,765.74	3,695.10	9,665.19	6,867.24
NW_0/BE_0	2.36	1.83	2.50	1.78

Table 4 reports the net worth valuation results for SU and IU, assuming social benefit fully converted to earnings and cash flows. Panel A provides the inputs and assumptions for all methods used. Panel B provides the results for the EP to Equityholders method, and Panel C for the FCFE method. In Panel A of Table 4, the growth rate based on historical adjusted net earnings is 3.74% for SU and 4.38% for IU, and the growth rate based on historical revenues is 2.43% for SU and 2.81% for IU. They are much lower than the average growth rates of education companies (7.51% and 7.68%). Thus, it seems proper to employ a one-stage constant growth pattern in the valuation of SU and IU. In addition, we assume that 100% of social benefit is converted to net earnings and free cash flows.

In Panel B of Table 4, the future ROE (based on adjusted net earnings) is 14.12% for SU and 14.22% for IU, roughly in line with the average ROE of education companies (12.95%). For both universities, since the future ROE is larger than the cost of equity (8.24%), the annual economic profit and the market value added to equityholders are positive, and the net worth exceeds the book equity. As shown, the net worth of SU is estimated to be $4.1 to $4.7 billion, which is 2.01 to 2.31 times the book equity. The net worth of IU is estimated to be $8.1 to $9.9 billion, which is 2.10 to 2.55 times the book equity.

In Panel C of Table 4, the future FOE (based on adjusted free cash flows) is 10.62% for SU and 9.63% for IU. As displayed, the net worth of SU is estimated to be \$3.7 to \$4.8 billion, which is 1.83 to 2.36 times the book equity. The net worth of IU is estimated to be \$6.9 to \$9.7 billion, which is 1.78 to 2.50 times the book equity.

BREAKEVEN: PARTIAL CONVERSION OF SOCIAL BENEFIT

In the previous section, we report that the estimated net worth is higher than the book equity based on the best-scenario assumptions. In this section, we assume that less than 100% of social benefit is converted to profits. That is, one part of existing social benefit is converted to net earnings and free cash flows, and the other part is distributed to students as scholarship and financial aid. With lower adjusted net earnings and lower free cash flows, the net worth would become lower. If the net worth is reduced to the level of book equity (i.e., the breakeven situation), the required percentage of social benefit conversion can be estimated.

Table 5. Breakeven: partial conversion of social benefit

	Syracuse University		Indiana University	
	Earnings growth	Revenues growth	Earnings growth	Revenues growth
Panel A. Inputs and assumptions				
Future growth rate (g)	3.74%	2.43%	4.38%	2.81%
Future cost of equity (k)	8.24%	8.24%	8.24%	8.24%
BE_0 (\$million)	2,017.59	2,017.59	3,865.54	3,865.54
NW_0 (\$million)	2,017.59	2,017.59	3,865.54	3,865.54
NW_0/BE_0	1.00	1.00	1.00	1.00
Panel B. EP to Equityholders method				
Future ROE	8.24%	8.24%	8.24%	8.24%
EP_1 (\$million)	0.00	0.00	0.00	0.00
% of social benefit converted	53.08%	53.08%	34.56%	34.56%
Panel C. FCFE method				
Future FOE	4.50%	5.80%	3.85%	5.42%
$FCFE_1$ (\$million)	90.75	117.04	148.91	209.58
% of social benefit converted	51.17%	61.56%	36.78%	53.95%

Table 5 reports the breakeven results for SU and IU, assuming net worth equal to book equity. In Pane A of Table 5, the growth rates, the cost of equity, and the book equity are the same as those in Panel A of Table 4. For the breakeven assumption, we set the net worth equal to the book equity. Thus, the net worth to book equity ratio is 1.00 accordingly.

In Panel B of Table 5, the future ROE is equal to the cost of equity, and the annual economic profit is zero, implying zero market value added to equityholders. Both the future ROE and EP_1 are lower than those in Panel B of Table 4. In this situation, SU needs to convert 53.08% of existing social benefit to net earnings and free cash flows, and IU, 34.56%.

In Panel C of Table 5, the future FOE is 4.50% to 5.80% for SU and 3.85% to 5.42% for IU. Both the future FOE and $FCFE_1$ are lower than those in Panel C of Table 4. In this situation, SU needs to convert 51.17% to 61.56% of existing social benefit to net earnings and free cash flows, and IU, 36.78% to 53.95%.

CONCLUSIONS

We use the DCF approach to estimate the net worth for two nonprofit universities: SU and IU. Our study transforms nonprofits to for-profits by converting nonprofit social benefit distributions to net earnings and free cash flows and by adopting for-profit cost of equity and tax rate. These adjustments add considerable value to nonprofit equityholders. The following provides a summary of our main results.

First, by using different valuation methods (EP to Equityholders and FCFE) and different growth rates (earnings and revenues), we demonstrate that the net worth to book equity ratio is 1.83 to 2.36 for SU and 1.78 to 2.55 for IU in June 2017. These ratios of SU and IU are by and large consistent with the price/book ratio (1.95) of the for-profit education companies at the same time. Our results are established based on the best-scenario assumptions. That is, the existing social benefit could be fully converted to net earnings and free cash flows, and the potential loss of the public-and-government-supported fund revenues could be offset by the potential increase in tuition revenues.

Second, when the net worth breaks even at the book equity, we find that SU needs to convert 51.17% to 61.62% of existing social benefit to net earnings and free cash flows, and IU, 34.56% to 53.95%, while the rest is distributed to students as scholarship and financial aid.

Third, once acquired by for-profit companies, SU and IU in the worst situation would lose all the fund revenues (gifts or contributions, grants and contracts, and state appropriations). To make up for these losses and to maintain the same levels of total revenues, SU would have to increase tuition revenues by 20.89%, and IU, 100.86%. Therefore, SU relies much less on the public-and-government-supported fund revenues than IU. This is in line with the fact that the former charges much higher per-student tuition and fees than the latter.

Please note that the discussions in this paper do not represent opinions of any universities, nor those of any potential investors. They are only the academic inquiries of the authors, and they may contain errors and biases.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

The Effect of CEO Twitter on Compensation

Linh Thi Thuy Tran
Yuan Ze University
Taoyuan, Taiwan
E-mail address: linh28689@gmail.com

Chien-Ling Lo
Yuan Ze University
Taoyuan, Taiwan
E-mail: cllo@saturn.yzu.edu.tw

Po-Hsin Ho
National Central University
Taoyuan, Taiwan
E-mail address: phho@mail.ntpu.edu.tw

Chih-Yung Lin
National Chiao-Tung University
Hsinchu, Taiwan
E-mail: d95723009@ntu.edu.tw

ABSTRACT

This research investigates whether CEO using twitter, proxied by CEO_twitter influences CEO compensation. First, we find that CEO_twitter is significantly positive related to total compensation (TDC1), confirming CEO_twitter has a real effect on compensation. We also examine whether posting more tweet can curtail the market responses to the earning announcements. In addition, the effect is more salient in the sample of higher information asymmetry. Besides, the effect of CEO Twitter becomes stronger when firms have a high product market competition. Overall, our findings support that CEO with twitter account receive more total CEO compensation.

KEYWORDS: CEO Twitter, Compensation, Abnormal return, Information asymmetry

INTRODUCTION

Information intermediaries, besides traditional ones (Miller 2006; Miller and Skinner 2015, Bushee and Miller 2012), recently explode in new sources of information like Twitter, is of crucial importance for a firm. It can improve the firm's information environment (Chen, Hwang, and Liu, 2017,2018; Bartov, Faurel and Mohanram, 2018; Elliott, Grant and Hodge 2018) so that investors can take some advantages, acquire timely and value-relevant information. Corporate executives are becoming increasingly active on twitter owing to direct-access information technologies (DAITs), which allow users to directly access investors. In the research of Blankespoor, Miller and White, 2014, to examine the effect of

DAIT dissemination on information asymmetry, they focus on sample of information technology firms with active Twitter accounts and find that dissemination via Twitter is associated with lower bid-ask spreads, greater depths, and a higher liquidity ratio after controlling for the information content of the news and other firm as well as market characteristics. It supports that firms use Twitter to spread out news, and this dissemination helps reduce information asymmetry. Firms can use twitter to reduce investors' information acquisition costs and allow more potential investors to process the information.
Further, the results from the study of Chen et al, 2018 suggest that tweets coming out of top executives' personal Twitter accounts are significantly more impactful than those coming out of firm-managed accounts (Chen, Hwang and Liu, 2018). Using Twitter gives CEO any benefit in compensation is an intriguing question which has not been addressed. In this paper, we fill this gap in the literature by examining whether using twitter is associated with higher compensation. Specifically, we explore the following three research questions: (1) Do CEOs with twitter account get higher total compensation? (2) whether posting more tweet can curtail the market responses to the earning announcements (3) the effect is more salient in the sample of higher information asymmetry and high market competition product? (4) Do twitter effects still hold after adjusting for other CEO compensation schemes.

To study our three research questions, we construct our sample of top executives' personal Twitter accounts, we download a list of all CEOs in the Execucomp database between 2006 and 2015. Execucomp covers the S&P 1500 as well as companies that were once part of the S&P 1500 index and that are still trading. We start with the complete list of all CEOs in Execucomp and locate users with active Twitter accounts that have the same first and last names as the CEO in question. We then cross-check the executives' middle names, gender, and company information with user characteristics; we also read tweets to determine whether any account that we find does indeed belong to the executive in question. Through this labor-intensive process, we determine that 336 S&P 1500 CEOs have active personal Twitter accounts and work for firms that have the data necessary to conduct our tests. We obtain all accounting variables and stock prices from the Compustat database and the Center for Research in Security Prices (CRSP). Corporate-governance and CEO-compensation-related variables come from RiskMetrics and ExecuComp.
We document three sets of findings. First, in our main finding we find that CEOs using twitter had higher total compensation (TDC1). That is, with CEOs own at least one personal twitter account, their compensation is higher after controlling for CEO characteristics, firm characteristics and corporate governance factors. The evidence remains strong when we control for industry and year fixed effects.
In further testing, we examine in greater detail whether the effect of using twitter derives from the information asymmetry channel. We use several variables to measure the information asymmetry, including firm age, number of analyst coverage and rating. We argue that, if

CEO with twitter account is viewed as a representation of information asymmetry, then its effect would be outweighed in the sample of higher information asymmetry. Our results provide supportive evidence for this channel, confirming that our main findings of twitter effect are driven primarily by firms with high information asymmetry. Hence, our results suggest that the influence of twitter derives from information asymmetry problem between executives and shareholders. We find supportive evidence that the effect tends to be stronger if the firm is younger, less coverage, and low credit rating.

Next, we analyze the market response of the earning announcements. The dependent variable is the cumulative abnormal return (CAR hereafter), which is the sum of abnormal returns over the windows. The independent variable is standardized unexpected earning (SUE), which is defined as the standardized of the difference between actual earning and the mean of analyst earnings forecasts. To incorporate the role of using twitter, we interact SUE with CEO_twitter in the right-hand side of the regression. Intuitively, we want to explore whether using twitter can affect the market responses to the earning announcements. We run the regression for its subsamples: (1) no twitter; (2) tweet with low frequency; (3) tweet with high frequency.

The regression results show a positive relation between SUE and the CAR, which is plausible. Our key coefficient is the interaction between SUE and CEO_twitter is positive but not significant in the low frequency sample. On average, the market reacts similarly to the news, regardless of whether CEO using twitter or not. However, we find that investors react differently to group with CEOs tweet frequently. Specifically, the interaction between SUE and CEO_twitter is negative and statistically significant in the subsample of high frequency.

Another possible channel is market competition. We further investigate whether the twitter effect would react differently to firms with high market threat. We conduct a subsample analysis by partitioning firms by several proxies for market competition such as CP oriented, product similarity, and product market fluidity (Hoberg and Phillips (2016), Hoberg et al. (2014), and Hoberg and Phillips (2010)) into high and low groups. The results show that the positive association between TDC1 and CEO_twitter occurs more strongly in high market competition group.

As a robustness check of our findings, we consider alternative measures of compensation (delta, vega, firm related wealth; nonequitybase) and test the effect of using twitter on them and find that the results remain supportive of our hypothesis.

Our paper contributes to several ways. First, our paper is related to the studies of social media. Overall, our findings highlight the benefit of using Twitter as a direct-access information - communication channel. These findings make two important contributions. First, they have important implications for the role Twitter plays in executives’ compensation. Most of the literature that addresses issues related to CEO compensation is based on agency theory, our results suggest that individuals use Twitter as an important social network to share information for CEOs's mutual benefit.

Second, our paper contributes to a burgeoning group of literature that focuses on the importance of the managers on a firm's policy such as investment policy and financing policy (Malmendier and Tate, 2005; Malmendier and Tate, 2008; Malmendier, Tate and Yan, 2011; Hirshleifer, Low and Teoh, 2012). The information on CEO Twitter account can help investors in their investment decisions. Thus, Twitter can play a role in making the market more efficient by uncovering additional value-relevant information, especially for firms in weak information environments, and regulatory intervention does not seem warranted.

Third, we contribute to the literature on information asymmetry. The existing studies have examined the information asymmetry issues on disclosure (Brown and Hillegeist, 2007; Shroff et al., 2013), media coverage (Fang and Peress, 2009), cost of capital (Armstrong et al., 2011), and financial reporting frequency and cost of equity (Fu, Kraft, and Zhang, 2012). Our paper also adds to growing literature linking information asymmetry and CEO compensation.

The rest of the paper is organized as follows: Section 2 discusses hypothesis development; Section 3 presents data sources, main variables, and methodology; Section 4 offers empirical results and provides several robustness checks; and Section 5 concludes the paper.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Before the mass adoption of social media applications, to disseminate financial information for the demand of investors, it can't help referring to the role of some internet tools such as: Google search (Da, Engelberg, and Gao (2011), Drake, Roulstone, and Thornock (2012)), Internet bulletin boards (Hirschey, Richardson, and Scholz (2000), Tumarkin and Whitelaw (2001)), message boards, such as Yahoo! or Raging Bull, (Antweiler and Frank (2004), Das and Chen (2007). However, these platforms are limited in term of online interaction in financial market and personal finance (Gallaugher and Ransbotham 2010). To adjust this drawback in recent years, social media appears and explodes in popularity with different forms (eg., Facebook, StockTwits, Twitters) and can have important economic consequence for the using firm and financial market (e.g., Tetlock 2007; Bollen, Mao and Zeng 2011; Laroche, Habibi and Richard, 2012; Chen et al 2014; Lee, Hutton, and Shu 2015; Jung, Naughton, Tahoun, and Wang 2018). Nowadays the power of social media in organizations is significant for not only short-term performance but also long-term productivity benefits inherently connected to firm equity value (Luo, Zhang and Duan 2013). Curtis, Richardson, and Schmardebeck (2016), who focus on the overall social media activity over 30-day rolling windows, find that high levels of activity are associated with greater sensitivity of earnings announcement returns to earnings surprises, while low levels of social media activity are associated with significant post-earnings-announcement drift. Firms increasingly make use of these channels so that the information environment is improved or information asymmetry is reduced, then increase the liquidity of the market for a firm's securities, lower bid-ask spreads and increased trading activity (Diamond and Verrecchia 1991; Leuz and Verrecchia 2000; Verrecchia 2001, Bushee and Leuz 2005; Xu and Zhang 2013, Blankespoor, Miller and White 2014).

Twitter was created in 2006 and undoubtedly is one of the most popular social media disclosure platforms in USA with short format (140-character limitation) and ease of information search. Jung, Naughton, Tahoun, and Wang (2018) collected data from all firms included in the S&P 1500 index and show that Twitter has become the preferred social media platform for companies. Using twitter gives account users opportunity to communicate, share opinions and facilitate open sharing of information in a timely fashion. The text content of Twitter may be day-to-day activities, current interests, and personal mood (Naaman, Boase, and Lai (2010)). Bollen, Mao, and Zheng (2011) find that by text processing techniques, the collective mood states posted on Twitter can help predict changes in the value of the Dow Jones Industrial Average over time. Consistently, using Twitter data, Mao, Wei, Wang, and Liu (2012) show that the daily number of tweets that mention S&P 500 stocks is significantly associated with S&P 500 daily closing price, S&P 500 daily price change and S&P 500 daily absolute price change.

The results in the research of Blankespoor, Miller, and White (2014) indicate that additional dissemination of firm-initiated news via Twitter is associated with lower bid-ask spreads and greater depths, and a higher liquidity ratio after controlling necessary factors such as the information content of the news, the presence of information intermediaries, market conditions, and firm-specific characteristics. Unlike traditional approach of relying on users to access content that is disseminated through media outlets, Twitter allows companies to “push” disclosures directly to users and so directly disseminate news, hereafter helps reduce information asymmetry.

Beside the benefits of company’s own social media pages, a CEO who is an influencer in his or her space can be a big advantage. Users can cultivate their own followings, effectively reduce cost of digital distribution. Making use of Twitter, CEO can write their own stories which influence media narratives rather than press. For example, to become a superstar, the media play a causal role in fostering a celebrity culture and enable the observed changes in CEO behavior, as a consequence, CEO compensation increase following CEO awards (Malmendier and Tate 2009). So why don’t CEOs make use of this free channel to increase CEO status and power within the organization. What left unexplored by these literatures is the question of whether in a company CEO use Twitter account is correlated with CEO compensation, the very question we examine in our paper. Because CEO pay is a topic that has received significant attention both in the popular press and academic journals. Despite a wealth of research, largely grounded in agency or managerialist perspectives (Hartzell and Starks 2003; Chhaochharia and Grinstein 2009; John, Mehran, and Quian 2010, Bugeja, Matolcsy and Spiropoulos 2012; Banker et al 2013). It is significant to have new approaches in the study of CEO rewards. We argue that inasmuch as CEO external directorate networks are strategically valuable to firms, they should be reflected in CEO compensation levels, particularly when the benefits of those ties are of greatest value. Further, the results from study of Chen, 2018 suggest that tweets coming out of top executives' personal Twitter accounts are significantly more impactful than those coming out of firm-managed accounts (Chen, Hwang and Liu 2018). Karadunam 2013, in his preliminary study on the effect social media on personal branding efforts of top level executives of the company, he also emphasizes that making use of social media, CEOs not only can create value for themselves but company also. Based on this considerations, it is hypothesized that:

Hypothesis 1: firms with CEO using twitter are more likely to pay higher compensation

In our paper, we refer to earnings announcements information because information fully reflects the abnormal returns on market efficiency that captures the economic hypothesis (Fama (1998) and large abnormal returns around earnings announcement dates (Kaniel et al. (2012)). Furthermore, the previous studies confirm that earnings announcement influence price changes and abnormal returns (May, 1971; Kane et al., 1984; and La Porta et al., 1997). For example, La Porta et al. (1997) mention that 20% stock returns around the earnings announcement dates. In addition, Lee, Hutton, and Shu (2015) document that the impact of Twitter can weaken the negative price reaction to recall announcements by direct broadcasting firm's intended message to investors without distorting the content. The authors find if properly managed, firms can engage users and customers through social media to mitigate the adverse effect of bad reputation associated with product recalls, a result that potentially extends to other types of corporate events. Also Bartov, Faurel, and Mohanram (2018) argue that twitter can be useful in predicting abnormal return around earnings announcements. They classified two notable roles of Twitter in providing new sources of information as well as disseminating existing information. Especially the opinion from Twitter posts is more prominent in predicting announcement return for high information assymetry environment. Thus, we examining whether CEO twitter effect as information react on earnings announcement dates that reflect the market efficiency and expect our next hypothesis as follow:

Hypothesis 2: the CEO twitter effect are more likely to curtail the market responses to the earning announcements.

The investors may obtain a different level of information with several characteristics, thus, we distinguish original sample into two groups for each characteristic to discuss the information asymmetry. One group of investor has superior information before being spread to the public and another group received after the information is spread to the public

(Tetlock, 2010). Armstrong et al. (2011) argue that information asymmetry impact on the cost of capital when the market competition is low. Tang (2009) find that credit rating information reduces the information asymmetry and allows the lender to access the borrowers' risk profile. Baik, Kang, and Kim (2010) analyze the role of geographic proximity and find that the significant positive correlation of local institutional investors and stock returns is stronger in firms with high information asymmetry. Also, Boone and White (2015) find that institutional investors are associated with analyst following reduces the information asymmetry, trading costs, and enhances monitoring. Blankespoor, Miller, and White (2014) show evidence that Twitter using help reduce information asymmetry by directly spreading out information to investor. Thus, we expect our next hypothesis as follow:

Hypothesis 3: The effect of CEO using twitter on CEO compensation is stronger for firms with high information asymmetry.

According to Akins, Ng, and Verdi (2012), Easley and O'Hara (2004), and Kyle (1985), a degree of private information for investors is lower when firms have a strong competition because competition leads information more quickly that reduce investors to price protect being reflected in the equilibrium price and competition reduces the information asymmetry to uninformed investors. They also mention that pricing is decreases with quality of information. Furthermore, Armstrong et al. (2011) find that firms with lower market competition tend few to no analyst following increases information asymmetry and implied the slope of cost of capital. They also find that information asymmetry with strong market competition degree has no effect on cost of capital.

Hoberg et al. (2014) argue that competitive threats firms to repurchase fewer shares and paying lower dividends and holding more cash and liquid assets. Furthermore, Hoberg and Phillips (2010) also argue that firms face nearer rivals' threats that may need to change their product frequently to maintain firm's profit. In other way, the product similarity is associated with lower profitability. Thus, we expect our next hypothesis as follow:

Hypothesis 4: The effect of CEO using twitter on CEO compensation is stronger for firms with high market competition.

DATA AND SUMMARY STATISTICS

Data and Other Variables

Our main explanatory variable is an executive's Twitter account. It takes the value of one if the executive of firm *i* adopted Twitter as of year *t*, and zero otherwise. To construct our sample of chief executives' personal Twitter accounts, we download a list of all CEOs in the Execucomp database between 2006 and 2015. Execucomp covers the S&P 1500 as well as companies that were once part of the S&P 1500 index and that are still trading. We start with the complete list of all CEOs in Execucomp and locate users with active Twitter accounts that have the same first and last names as the CEO in question. We then cross-check the executives' middle names, gender, and company information with user characteristics; we also read tweets to determine whether any account that we find does indeed belong to the executive in question. Through this labor-intensive process, we determine that 336 S&P 1500 CEOs have active personal Twitter accounts and work for firms that have the data necessary to conduct our tests. We exert great effort in separating out Twitter accounts managed by firms.

The main dependent variable is the level of annual CEO compensation collected from Execucomp databases. CEO compensation is measured using three alternative measures: total annual compensation, annual salary, and annual bonus. The natural logarithm of Execucomp's TDC1 is used to measure total annual compensation. Total compensation (*Total_comp*) includes the CEO's salary, bonus, other annual compensation, total value of restricted stock granted, total value of stock options granted (valued using Black–Scholes), long-term incentive payouts, and all other compensation (in thousands of dollars).

The remaining independent variables are identified from Bugeja, Matolcsy and Spiropoulos 2012; Chhaochharia and Grinstein 2009. They provide controls for governance, CEO, and economic characteristics that are associated with CEO compensation. CEO characteristics include the natural logarithm of the number of years of service of the current CEO (*Log(CEO_tenure)*), dummy variable equal to 1 if the CEO is also the board chairperson (*CEO_chair*) and otherwise is zero, dummy variable equal to 1 if it is the CEO's first year of service at a firm (*CEO_firstyear*) and otherwise is zero, dummy variable equal to 1 if the CEO owns 5% or more of the company's stock (*CEO5pct*) and otherwise is zero.

Firm characteristics are from Compustat and included a firms' book-to-market ratio (*BMV*), company's annual common stock return (*RET*), accounting measure is return on assets (*ROA*), the standard deviation of common stock returns (*Log(Std5RET)*), the standard deviation of return on assets measured over the previous five years (*Log(Std5ROA)*), average total liabilities divided by average total equity (DE).

Corporate governance controls from Risk Metrics included number of board directors (*Bdsize*), dummy variable equal to 1 if the compensation committee is comprised wholly of independent directors(*Indcompcom*) otherwise is zero. Finally, we also control for industry and year-fixed effects. We adjust the standard errors by within-firm clustering (White, 1980).

Summary Statistics

Table 1 presents the summary statistics of *total compensation(TDC1)*, CEOs with twitter account (*CEO_twitter*), and other control variables. The mean value of *CEO_twitter* is 0.0576 with a standard deviation of 0.2331. For CEO characteristics, the mean of *TDC1* is US$ 3003 thousand (natural logarithm of 8.0074) with a standard deviation of 1.0350. The average number of years the CEO held his/her position in a company before he/she signed the bank loan contract is 6 years (natural logarithm of 1.7985) in the sample.

Table 1: Descriptive Statistics

This table presents the descriptive statistics of all the main variables used in this research. The sample period is from 2006 to 2015.

Variables	Mean	p25	p50	p75	Sd
TDC1	8.0074	7.3147	8.0651	8.7572	1.0350
CEO_twitter	0.0576	0.0000	0.0000	0.0000	0.2331
CEO5pct	0.1155	0.0000	0.0000	0.0000	0.3196
CEO_firstyear	0.0683	0.0000	0.0000	0.0000	0.2522
CEO_tenure	1.7985	1.0986	1.9459	2.3979	0.8666
CEO_chair	0.1240	0.0000	0.0000	0.0000	0.3296
Bdsize	9.4386	8.0000	9.0000	11.0000	2.4191
Indcompcom	0.9550	1.0000	1.0000	1.0000	0.2074
BMV	0.5503	0.2827	0.4672	0.7234	0.4044
DE	0.8705	0.1125	0.4645	1.0310	1.7270
ROA	0.0419	0.0125	0.0435	0.0829	0.0896
RET	0.1711	-0.1008	0.1179	0.3481	0.5018
Std5RET	0.1080	0.0687	0.0959	0.1323	0.0560
Std5ROA	0.0509	0.0115	0.0253	0.0576	0.0721

As a univariate test, we split the sample into two groups based on whether CEO using twitter or not Then, we compare our main variables of two groups and conduct a *t*-test of mean equality. Table 2 presents the comparison between two groups. In the group with twitter account, the mean of *TDC1* is 8.2070, which is higher than that in the group of not using twitter by 0.2119. The mean difference between the two groups are statistically significant at 1 percent level.

Turning our firm characteristics, the group with CEOs twitter account tend to have greater return RET and firm risk Std5ROA. However, these firms have less leverage and investment opportunity BMV.

Table 2: CEO, governance, and firm characteristics for CEOs with and without Twitter account.

This table presents the mean comparison of CEO, governance, and firm characteristics for CEOs with and without Twitter account. *t*-test deployed to compare the mean. *, ** and*** denote significance at 10%, 5% and 1% levels, respectively.

Variables	With twitter	Without twitter	Differences	*t*-statistics
TDC1	8.207037	7.995111	.2119261***	6.1798
CEO5pct	.1489362	.1134136	.0355226***	3.2228
CEO_firstyear	.075028	.0678606	.0071674	0.8238
CEO_tenure	1.79071	1.798946	-.0082357	-0.2755
CEO_chair	.0044793	.1314538	-.1269745***	-11.2151
Bdsize	9.2571	9.450177	-.1930766***	-2.0017
Indcompcom	.9461883	.9555449	-.0093566	-1.1315
BMV	.4822057	.5545956	-.0723899***	-4.9636
DE	.6912541	.8817648	-.1905107***	-3.0533

ROA	.0447972	.0417332	.003064	0.9475
RET	.2183034	.1681078	.0501956***	2.7720
Std5RET	.1088312	.1079726	.0008585	0.4245
Std5ROA	.0637015	.0501014	.0136001***	5.2311

We also report the Pearson's correlation coefficient matrix in Table 3. As expected, we find a positive and significant correlation between *TDC1* and *CEO_twitter*. The correlation coefficients between total compensation and other control variables such as return on assets and annual common stock return are also positive and significant at the one-percent level.

Table 3: Correlation matrix

This table presents the Pearson correlation matrix between all the variables. The sample period is from 2006 to 2016. * and ** denote significance level at 5% and 1%, respectively

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)TDC1	1.00						
(2)CEO_twitter	0.05***	1.00					
(3)CEO5pct	-0.18***	0.03***	1.00				
(4)CEO_firstyear	0.02**	0.01	-0.07***	1.00			
(5)CEO_tenure	-0.05***	0.00	0.31***	-0.56***	1.00		
(6)CEO_chair	-0.03***	-0.09***	0.11***	-0.04***	0.10***	1.00	
(7)Bdsize	0.28***	-0.02**	-0.19***	0.04***	-0.12***	0.04***	1.00
(8)Indcompcom	0.09***	-0.01	-0.02**	0.00	0.01	-0.01	0.00
(9)BMV	-0.16***	-0.04***	0.00	0.01	-0.02**	-0.07***	0.08***
(10)DE	0.07***	-0.03***	-0.06***	0.00	-0.03***	0.03***	0.15***
(11)ROA	0.14***	0.01	0.01	-0.03***	0.06***	0.05***	-0.06***
(12)RET	0.12***	0.02**	-0.01	-0.03***	0.02**	-0.04***	-0.03***
(13)Std5RET	-0.20***	0.00	0.08***	-0.01	-0.01	-0.16***	-0.23***
(14)Std5ROA	-0.13***	0.04***	0.04***	0.00	-0.05***	-0.06***	-0.18***
Variable	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(8)Indcompcom	1.00						
(9)BMV	0.00	1.00					
(10)DE	0.00	0.01	1.00				
(11)ROA	-0.01	-0.34***	-0.11***	1.00			
(12)RET	0.00	-0.28***	-0.01	0.18***	1.00		
(13)Std5RET	-0.02*	0.18***	0.02**	-0.26***	0.14***	1.00	
(14)Std5ROA	-0.01	-0.02**	-0.07***	-0.28***	0.06***	0.46***	1.00

EMPIRICAL RESULTS

Main Result: Twitter and CEO Compensation

We use the following regression setting to test our main hypothesis following Vidhi Chhaochharia and Yaniv Grinstein, (2009).

$$TDC1_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where the dependent variable is *TDC1(total compensation)* and the main explanatory variable is *CEO_twitter*. Specifically, $tdc1_{i,t}$ is the natural logarithm of compensation for chief executives officer *i* in year *t*; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm *i* adopted Twitter as of year *t*, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm *i* in year *t-1*. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics *i* in year *t*. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level robust standard errors (White, 1980).

Table 4 presents the effects of using twitter and CEO compensation. Our focus is the coefficient of *CEO_twitter* on *TDC1*. For the robustness of our results, we conduct four specifications in the regression setting. The first one do not control for industry and year fixed effects; the second adds controls for year fixed effect, the third for governance characteristics, CEO characteristics, industry and year fix effects and the fourth for all governance characteristics, CEO characteristics, firm characteristics, industry and year fix effects.

Consistent with our expectation, positive and significant coefficients are observed in all specifications, even though we have controlled for all potential factors, indicating that CEO with twitter account will get higher compensation. Thus, the results support our Hypothesis 1 that *firms with CEO using twitter are more likely to pay higher compensation*. Specifically, the coefficients of *Vega* are positive and significant at the one-percent level from a 0.1430 to 0.3070, indicating that using twitter is associated with higher compensation in a firm.

Table 4: Bank's CEO risk-taking incentives and Bank loan spread

These tables present the ordinary least squares (OLS) regression results of CEO twitter effect on CEO compensation.

$$TDC1_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where the dependent variable is TDC1(total compensation) and the main explanatory variable is CEO_twitter. Specifically, $tdc1_{i,t}$ is the natural logarithm of compensation for chief executives officer i in year t; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm i adopted Twitter as of year t, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm i in year t-1. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics i in year t. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level-robust standard errors (White, 1980). The sample period is 2006-2015. All the variables are defined in Appendix A *, ** and *** denote the significance level of 10%, 5% and 1%, respectively.

	(1)	(2)	(3)	(4)
	TDC1	*TDC1*	*TDC1*	*TDC1*
CEO_twitter	0.3070***	0.1779***	0.1376***	0.1430***
	(7.93)	(4.59)	(3.55)	(3.70)
CEO5pct	-0.3044***	-0.2626***	-0.3237***	-0.3049***

	(-8.40)	(-7.31)	(-9.15)	(-8.72)
CEO_firstyear	-0.0510	-0.0931**	-0.0800*	-0.1506***
	(-1.11)	(-2.05)	(-1.83)	(-3.27)
CEO_tenure	-0.0245*	-0.0556***	-0.0388***	-0.0461***
	(-1.76)	(-4.05)	(-2.95)	(-3.53)
CEO_chair	-0.1485***	0.2395***	0.2508***	0.2207***
	(-4.91)	(6.01)	(6.62)	(5.82)
Bdsize	0.1074***	0.1047***	0.1272***	0.1172***
	(21.67)	(21.69)	(22.08)	(20.32)
Indcompcom	0.4529***	0.3649***	0.3422***	0.3615***
	(10.02)	(8.25)	(8.02)	(8.43)
BMV				-0.2547***
				(-8.34)
DE				0.0504***
				(7.45)
ROA				0.1372
				(0.97)
RET				0.1767***
				(8.44)
Std5RET				-2.6328***
				(-10.38)
Std5ROA				-0.1291
				(-0.75)
Constant	6.8066***	6.6060***	6.9793***	7.6005***
	(97.65)	(74.39)	(28.37)	(43.27)
Control for				
Industry FE	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes
Obs.	10690	10690	10690	10093
Adj R^2	0.1081	0.1540	0.2442	0.2892

Concerning control variables, the results are generally consistent with previous studies (Chalmers et al., 2006; Core et al., 1999, Bugeja, Matolcsy and Spiropoulos, 2012). Compensation is significantly negatively associated with *CEO5pct* and *CEO_firstyear* which consistent with the notion that in the first year of employment CEOs receive significantly lower total compensation and CEOs who own more than 5% of stock are paid significantly less. This finding can be explained by agency costs: CEOs with lower ownership are more likely to extract higher pay inconsistent with maximizing shareholder wealth. Conversely, compensation is significantly positively associated with *CEO_chair.* This may be interpreted as CEOs using their power to extract greater pay, or it may reflect additional talent or effort so CEOs who are also chairpersons receive significantly more pay.

The significant results on the corporate governance measures are generally positive and indicate that total compensation increases with board size and total compensation increases with independent compensation committee. An independent compensation committee is unrelated to CEO remuneration. Overall, the results for the corporate governance measures are similar to those in prior literature (Core et al., 1999; Wade et al., 1990, Bugeja, Matolcsy and

Spiropoulos, 2012). The relationship between firm stock performance and remuneration is positive for total pay. but surprisingly the impact of firm risk on total compensation is significant negative which is shown by the relationship between salary and the standard deviation of returns. Overall, the effects of the control variables exhibit a similar pattern as reported in the existing literature.

Twitter Effect and Cumulative Abnormal Return (*CAR*)

Information fully reflects the abnormal returns on market efficiency that captures the economic hypothesis (Fama (1998) and large abnormal returns around earnings announcement dates (Kaniel et al. (2012)). Furthermore, the previous studies confirms that earnings announcement influence price changes and abnormal returns (May, 1971; Kane et al., 1984; and La Porta et al., 1997). For example, La Porta et al. (1997) mention that 20% stock returns around the earnings announcement dates. In addition, Lee, Hutton, and Shu (2015) document that the impact of Twitter can weaken the negative price reaction to recall announcements by direct broadcasting firm's intended message to investors without distorting the content. The authors find if properly managed, firms can engage users and customers through social media to mitigate the adverse effect of bad reputation associated with product recalls, a result that potentially extends to other types of corporate events. Also Bartov, Faurel, and Mohanram (2018) argue that twitter can be useful in predicting abnormal return around earnings announcements. They classified two notable roles of Twitter in providing new sources of information as well as disseminating existing information. Especially the opinion from Twitter posts is more prominent in predicting announcement return for high information assymetry environment.

To examine the effects of using twitter in cumulative abnormal return (*CAR*), we use the sensitivity of earnings announcement. We run a regression model as follows:

$$CAR[-1;1]_{it} = \beta_1 SUE\_mean_{it} + \beta_2 CEO\_twitter_{it} + \beta_3 SUE\_mean\_Twit_{it} + \gamma' Z_{it-1} + u_j + v_t + \varepsilon_{it} \quad (2)$$

where $CAR[-1;1]_{it}$ is the cumulative abnormal return (from the Fama-French three factor model) in the window [-1;1]; SUE_mean_{it} is the difference between actual and the mean of earning per share (EPS) forecasts of the analysts, divided by the standard deviation of EPS forecast; $SUE\_mean\_Twit_{it}$ is the interaction between SUE_mean and *CEO_twitter;* Z_{it-1} is the vector of control variables; u_j and v_t are industry and year fixed effect, respectively. Control variables include *Asset* (natural logarithm of total assets), *Q* (Tobin's Q), *Leverage* (ratio of long-term debt plus debt in current liabilities to total assets), *BHAR* (the corresponding buy-and-hold abnormal return), *Tangibility* (ratio of net property, plant, and equipment to total assets), *Profitability* (ratio of earnings before interest, taxes, depreciation, and amortization to total assets), *Z-score* (a modified Altman's Z-score), and *CF volatility* (ratio of standard deviation of quarterly cash flows from operations over the four fiscal years prior to loan initiation year to total debt).

Our key independent variable is the interaction between SUE_mean and *CEO_twitter*, which shows how the market reacts to the CEO using twitter of the reporting firm. According to our predictions, this interaction term's coefficient should be negative and significant.

Table 5 presents the regression results. Our key coefficient is the interaction between SUE and CEO_twitter is positive but not significant in the low frequency sample (Model 2). On average, the market reacts similarly to the news, regardless of whether CEO using twitter or not. However, we find that investors react differently to group with CEOs tweet frequently. Specifically, the results support the hypothesis 2 that the interaction between SUE and CEO_twitter is negative and statistically significant in the subsample of high frequency.

Table 5: Earning Announcement Responses: Mean SUE

This table presents regression results for the effects of using twitter in cumulative abnormal return (*CAR*), we use the sensitivity of earnings announcement. We run a regression model as follows:

$$CAR[-1;1]_{it} = \beta_1 SUE\_mean_{it} + \beta_2 CEO\_twitter_{it} + \beta_3 SUE\_mean\_Twit_{it} + \gamma' Z_{it-1} + u_j + v_t + \varepsilon_{it} \quad (2)$$

where $CAR[-1;1]_{it}$ is the cumulative abnormal return (from the Fama-French three factor model) in the window [-1;1]; SUE_mean_{it} is the difference between actual and the mean of earning per share (EPS) forecasts of the analysts, divided by the standard deviation of EPS forecast; $SUE_mean_Twit_{it}$is the interaction between SUE_mean and *CEO_twitter;* Z_{it-1} is the vector of control variables; u_j and v_t are industry and year fixed effect, respectively. Control variables include *Asset* (natural logarithm of total assets), *Q* (Tobin's Q), *Leverage* (ratio of long-term debt plus debt in current liabilities to total assets), *BHAR* (the corresponding buy-and-hold abnormal return), *Tangibility* (ratio of net property, plant, and equipment to total assets), *Profitability* (ratio of earnings before interest, taxes, depreciation, and amortization to total assets), *Z-score* (a modified Altman's Z-score), and *CF volatility* (ratio of standard deviation of quarterly cash flows from operations over the four fiscal years prior to loan initiation year to total debt).Appendix A provides the definitions of variables. In all equations, we report *t*-values based on the robust standard errors adjusted for the clustering at the firm level (White,1980). Appendix provides the definitions of variables. Superscripts *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	no twitter	low frequency	hi frequency
	(1)	(2)	(3)
	CAR[-1;1]	CAR[-1;1]	CAR[-1;1]
CEO_twitter		0.0146	0.0152
		(1.33)	(1.15)
SUE_mean_Twit		0.0000	-0.0072***
		(0.02)	(-4.82)
SUE_mean	0.0024***	0.0043***	0.0081***
	(5.73)	(3.46)	(5.61)
Assets	-0.0023***	-0.0003	0.0015
	(-3.05)	(-0.11)	(0.42)
Q	-0.0027**	-0.0009	0.0004
	(-1.99)	(-0.24)	(0.11)
Leverage	0.0121*	-0.0619*	0.0008
	(1.65)	(-1.73)	(0.02)
BHAR	-0.0036	-0.0053	-0.0046
	(-1.06)	(-0.60)	(-0.54)
Tangibility	-0.0207**	-0.0340	0.0399
	(-2.40)	(-0.89)	(0.81)
Profitability	-0.0024	0.0223	0.0030
	(-0.10)	(0.28)	(0.02)
Z_score	-0.0002	-0.0026	0.0019

	(-0.17)	(-0.51)	(0.20)
CFvolatility	-0.0000	0.0000	-0.0001***
	(-0.47)	(1.31)	(-2.69)
Constant	0.0369***	-0.0275	-0.0716
	(2.74)	(-0.96)	(-1.14)
Control for			
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs	6065	460	385
Adj. R^2	0.0370	0.0135	0.1522

As a robustness check, we replace the SUE_mean by SUE_median in Table 6. Specifically, SUE_median is the difference between actual and the median of analyst forecasts on EPS. We repeat our analysis in the Table 5 by SUE_median. Similar to Table 5, we find that the interaction between SUE_median and CEO_twitter is negative and significant only in the subsample of high frequency of tweet, also support hypothesis 2 (Model 3).

Table 6: Earning Announcement Responses: Median SUE

This table presents regression results for the effects of using twitter in cumulative abnormal return (*CAR*), we use the sensitivity of earnings announcement. We run a regression model as follows:

$$CAR[-1;1]_{it} = \beta_1 SUE\_median_{it} + \beta_2 CEO\_twitter_{it} + \beta_3 SUE\_median\_Twit_{it} + \gamma' Z_{it-1} + u_j + v_t + \varepsilon_{it} \quad (3)$$

where $CAR[-1;1]_{it}$ is the cumulative abnormal return (from the Fama-French three factor model) in the window [-1;1]; SUE_median_{it} is the difference between actual and the mean of earning per share (EPS) forecasts of the analysts, divided by the standard deviation of EPS forecast; $SUE_median_Twit_{it}$is the interaction between SUE_median and *CEO_twitter;* Z_{it-1} is the vector of control variables; u_j and v_t are industry and year fixed effect, respectively. Control variables include *Asset* (natural logarithm of total assets), *Q* (Tobin's Q), *Leverage* (ratio of long-term debt plus debt in current liabilities to total assets), *BHAR* (the corresponding buy-and-hold abnormal return), *Tangibility* (ratio of net property, plant, and equipment to total assets), *Profitability* (ratio of earnings before interest, taxes, depreciation, and amortization to total assets), *Z-score* (a modified Altman's Z-score), and *CF volatility* (ratio of standard deviation of quarterly cash flows from operations over the four fiscal years prior to loan initiation year to total debt).Appendix A provides the definitions of variables. In all equations, we report *t*-values based on the robust standard errors adjusted for the clustering at the firm level (White, 1980). Appendix provides the definitions of variables. Superscripts *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	no twitter	low frequency	hi frequency
	(1)	(2)	(3)
	CAR[-1;1]	CAR[-1;1]	CAR[-1;1]
CEO_twitter		0.0149	0.0154
		(1.35)	(1.17)
SUE_median_Twit		0.0001	-0.0073***

		(0.06)	(-4.79)
SUE_median	0.0024***	0.0044***	0.0082***
	(5.72)	(3.55)	(5.58)
Assets	-0.0023***	-0.0004	0.0014
	(-3.06)	(-0.15)	(0.39)
Q	-0.0027**	-0.0009	0.0004
	(-2.00)	(-0.25)	(0.10)
Leverage	0.0123*	-0.0614*	0.0020
	(1.66)	(-1.72)	(0.05)
BHAR	-0.0036	-0.0055	-0.0048
	(-1.05)	(-0.62)	(-0.56)
Tangibility	-0.0207**	-0.0340	0.0406
	(-2.40)	(-0.89)	(0.82)
Profitability	-0.0023	0.0242	-0.0028
	(-0.10)	(0.31)	(-0.02)
Z score	-0.0002	-0.0027	0.0021
	(-0.16)	(-0.53)	(0.22)
CFvolatility	-0.0000	0.0000	-0.0001***
	(-0.47)	(1.31)	(-2.71)
Constant	0.0369***	-0.0263	-0.0711
	(2.74)	(-0.91)	(-1.13)
Control for			
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs	6065	460	385
Adj. R^2	0.0368	0.0158	0.1540

Twitter Effect and Information Asymmetry

The empirical studies have indicated that information asymmetry cause a gap of information between managers and shareholders. Following the study of (Chae, 2005), (Autore and Kovacs, 2010), and (Leary and Roberts, 2010), we use several proxies to measure the information asymmetry such as, firm age(*Age*), analyst coverage(*Coverage*) and credit rating (*Rating*). They mention that proxies can capture the time-variation and cross-sectional levels in information asymmetry. We separate information asymmetry variables into high and low. The criteria for high information asymmetry as follows: (1) firm age (*Age*) are smaller than median; (2) analyst coverage(*Coverage)* are smaller than median; and (3) firms with low credit rating (*Rating low*); and the low information asymmetry is vice versa. The equation form of model is:

$$TDC1_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where the dependent variable is *TDC1(total compensation)* and the main explanatory variable is *CEO_twitter*. Specifically, $Tdc1_{i,t}$ is the natural logarithm of compensation for chief executives officer *i* in year *t*; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm *i* adopted Twitter as of year *t*, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm *i* in year *t-1*. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics *i* in year *t*. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level robust standard errors (White, 1980).

In Table 7 we hypothesize that the effect should be more significant in cases of high information asymmetry. In particular, CEOs using twitter account in young, low credit rating, or low analyst coverage firm tend to be positively related to compensation. Those results support Hypothesis 3 to confirm that t*he effect of CEO using twitter on CEO compensation is stronger for firms with high information asymmetry*.

Table 7: Market-based information asymmetry

This table presents the results of OLS regression that estimates the influence of twitter effect on CEO compensation by considering the subsample of information asymmetry. We separate information asymmetry variables into high and low. The criteria for high information asymmetry as follows: (1) firm age (*Age*) are smaller than median; (2) analyst coverage(*Coverage)* are smaller than median; and (3) firms with low credit rating (*Rating low*); and the low information asymmetry is vice versa. The equation form of model is:

$$TDC1_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where the dependent variable is TDC1(total compensation) and the main explanatory variable is CEO_twitter. Specifically, $tdc1_{i,t}$ is the natural logarithm of compensation for chief executives officer i in year t; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm i adopted Twitter as of year t, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm i in year t-1. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics i in year t. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level-robust standard errors (White, 1980). The sample period is 2006-2015. All the variables are defined in Appendix A *, ** and *** denote the significance level of 10%, 5% and 1%, respectively.

	High information asymmetry			*Low information asymmetry*		
	Age <Median	*Coverage <Median*	*Rating low*	*Age ≥Median*	*Coverage ≥Median*	*Rating high*
	(1)	(2)	(3)	(4)	(5)	(6)
CEO_twitter	0.1584***	0.1559**	0.1812***	0.1066**	0.0574	0.0532
	(2.82)	(2.08)	(3.70)	(2.06)	(0.42)	(0.54)
CEO5pct	-0.4107***	-0.3449***	0.0582	-0.1480***	-0.0860	-0.0522
	(-8.84)	(-4.66)	(1.04)	(-2.84)	(-0.65)	(-0.21)
CEO_firstyear	-0.1708**	-0.1589***	-0.1653***	-0.1584***	-0.0019	-0.2512**
	(-2.35)	(-3.01)	(-2.60)	(-2.77)	(-0.02)	(-2.48)
CEO_tenure	0.0024	-0.0377	-0.0333*	-0.1063***	-0.0819*	-0.1421***
	(0.13)	(-1.46)	(-1.83)	(-5.93)	(-1.77)	(-4.49)
CEO_chair	0.1705***	0.1301**	0.1463**	0.2714***	0.3244***	0.3793***
	(3.02)	(2.48)	(2.56)	(5.40)	(3.85)	(3.50)
Bdsize	0.1262***	0.1375***	0.0807***	0.1021***	0.0313	0.0134
	(17.86)	(13.66)	(11.98)	(12.74)	(1.18)	(1.23)
Indcompcom	0.3327***	0.3274***	0.3046***	0.3288***	0.5444***	0.0079
	(5.66)	(4.64)	(4.09)	(5.62)	(3.08)	(0.07)
BMV	-0.2632***	-0.2482***	-0.1967***	-0.3067***	-0.1945*	-0.3225***
	(-6.07)	(-4.43)	(-4.62)	(-6.95)	(-1.74)	(-3.25)

DE	0.0215***	0.0442***	0.0175**	0.0680***	0.0617***	0.0542***
	(2.94)	(3.36)	(2.52)	(6.40)	(2.78)	(3.35)
ROA	0.0678	0.1124	0.0393	0.0647	0.2142	-0.4444
	(0.36)	(0.52)	(0.17)	(0.30)	(0.49)	(-0.84)
RET	0.2062***	0.1730***	0.1351***	0.1263***	0.1746***	0.1565**
	(7.89)	(7.61)	(4.97)	(3.89)	(3.64)	(2.03)
Std5RET	-3.0660***	-2.4282***	-1.4260***	-1.6429***	-3.8724***	1.3519
	(-9.23)	(-5.46)	(-4.13)	(-4.31)	(-3.95)	(1.49)
Std5ROA	0.2215	-0.1064	-0.1723	-0.4149	-0.5356	0.2966
	(1.01)	(-0.35)	(-0.56)	(-1.49)	(-0.92)	(0.29)
Constant	7.4025***	7.3181***	7.9667***	7.8349***	7.9901***	8.8482***
	(34.81)	(30.76)	(32.64)	(40.29)	(15.72)	(33.04)
Control for						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	4486	8217	4262	5607	1876	1356
Adj. R^2	0.3108	0.3137	0.2648	0.3135	0.3136	0.3283

Twitter Effect and Market Competition

Next, we extend another channel though how the twitter can affect compensation by examining the possibility of market competition. We use several proxies for market competition such as consumer product-oriented (CP oriented), product similarity, and product market fluidity (Hoberg and Phillips (2016), Hoberg et al. (2014), and Hoberg and Phillips (2010).

We separate market competition proxies into strong and weak. The criteria for strong market competition if product similarity, and product market fluidity are larger than median, CP oriented is smaller than median and the weak market competition is vice versa.

In Table 8 shows the regression results of twitter effect on total compensation with strong and weak market competition. The proxies of CP oriented, product similarity, and product market fluidity show that the effect of CEO using twitter is stronger with firm with high market competition. The results support the hypothesis 4 that the significant positive association between TDC1 and CEO_twitter occurs more strongly in high market competition group.

Table 8: Product market competition

This table presents the results of OLS regression that estimates the influence of twitter effect on CEO compensation by considering the subsample of market competition. We use several proxies for market competition such as consumer product-oriented (CP oriented), product similarity, and product market fluidity (Hoberg and Phillips (2016), Hoberg et al. (2014), and Hoberg and Phillips (2010)). The equation form of model is:

$$TDC1_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where the dependent variable is TDC1(total compensation) and the main explanatory variable is CEO_twitter. Specifically, $tdc1_{i,t}$ is the natural logarithm of compensation for chief executives officer i in year t; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm i adopted Twitter as of year t, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm i in year t-1. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics i in year t. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the t-statistics reported are based on heteroscedasticity at firm-level-robust standard errors (White, 1980). The sample period is 2006-2015. All the variables are defined in Appendix A *, ** and *** denote the significance level of 10%, 5% and 1%, respectively.

	Low product market competition		
	Similarity low	Fluid low	CP oriented high
	(1)	(2)	(3)
	TDC1	TDC1	TDC1
CEO_twitter	0.1098**	0.0949**	0.0781*
	(2.20)	(1.98)	(1.80)
CEO5pct	-0.4584***	-0.3860***	-0.3076***
	(-9.25)	(-8.01)	(-7.87)
CEO_firstyear	-0.1855***	-0.1388**	-0.1269**
	(-3.02)	(-2.35)	(-2.56)
CEO_tenure	-0.0405**	-0.0504***	-0.0377***
	(-2.20)	(-2.73)	(-2.67)
CEO_chair	0.2580***	0.1917***	0.2512***
	(5.13)	(3.71)	(6.22)
Bdsize	0.1290***	0.1155***	0.1149***
	(15.79)	(15.21)	(18.14)
Indcompcom	0.3774***	0.3085***	0.3248***
	(6.80)	(5.15)	(7.12)
BMV	-0.2634***	-0.2824***	-0.2656***
	(-5.55)	(-6.17)	(-8.20)
DE	0.0560***	0.0246**	0.0514***
	(6.52)	(2.56)	(7.02)
ROA	-0.2918	-0.1501	0.2755*
	(-1.52)	(-0.79)	(1.74)
RET	0.1933***	0.2038***	0.1682***
	(6.68)	(7.22)	(7.48)
Std5RET	-3.2362***	-3.0313***	-2.3514***
	(-8.96)	(-8.43)	(-8.62)
Std5ROA	0.2211	-0.4005	-0.2257
	(0.92)	(-1.61)	(-1.15)
Constant	7.4341***	7.6953***	7.5949***
	(34.17)	(43.83)	(41.74)
Control for			
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs	5040	5099	8690
Adj. R^2	0.3321	0.2885	0.2883

	High product market competition		
	Similarity high	Fluid high	CP oriented low
	(4)	(5)	(6)
	TDC1	TDC1	TDC1
CEO_twitter	0.1801***	0.2370***	0.4268***
	(3.08)	(3.99)	(5.19)
CEO5pct	-0.1795***	-0.2350***	-0.2699***
	(-3.74)	(-4.61)	(-3.22)
CEO_firstyear	-0.1322*	-0.1454**	-0.1842
	(-1.95)	(-2.03)	(-1.62)
CEO_tenure	-0.0619***	-0.0511***	-0.0336
	(-3.32)	(-2.79)	(-0.96)
CEO_chair	0.1943***	0.2547***	-0.1049
	(3.44)	(4.69)	(-1.06)
Bdsize	0.1029***	0.1170***	0.1400***
	(12.23)	(14.22)	(12.42)
Indcompcom	0.3331***	0.3378***	0.5695***
	(4.94)	(5.43)	(5.71)
BMV	-0.2065***	-0.2612***	-0.1325
	(-5.08)	(-6.46)	(-1.55)
DE	0.0424***	0.0795***	0.0387***
	(3.95)	(8.34)	(2.61)
ROA	0.5249***	0.4127**	-0.3334
	(2.62)	(2.01)	(-1.03)
RET	0.1641***	0.1355***	0.2226***
	(5.50)	(4.48)	(4.05)
Std5RET	-1.9960***	-2.0714***	-3.7704***
	(-5.54)	(-5.90)	(-6.24)
Std5ROA	-0.8618***	-0.2389	0.1870
	(-3.40)	(-1.05)	(0.58)
Constant	7.9467***	7.5502***	6.2534***
	(22.66)	(26.47)	(23.61)
Control for			
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs	5053	4994	1403
Adj. R^2	0.2888	0.3228	0.4180

Robustness Checks

This subsection provides additional evidence on how CEO twitter account affects other measures of compensation. We replace *TDC1* by different dependent variables: *delta, vega, firm related wealth* and *non-equity base compensation*. Specifically, *delta* stand for manager's change in wealth (expressed in thousands of dollars) for a 1% increase in the firm's stock price. Calculated from managers' complete portfolio of stock and options from Execucomp or Yermack (1995) using the methodology from Core and Guay (2002). The value of the option portfolio is the sum of the Black–Scholes values of the newly granted options and previously granted unexercisable and exercisable options. The Yermack sample contains information on a manager's previously granted stock but not options, so we use information only on options granted in the last year to calculate the managers' options' delta. *Vega* stands for manager's change in wealth (expressed in thousands of dollars) for a 0.01 increase in the annualized standard deviation of firm's stock returns. *Firm-related wealth* is the sum of the value of the stock and option portfolio held by the executive (Daniel, Li and Naveen 2013), *non_equity-based* compensation is total compensation (*TDC1*) minus equity-based compensation (Chhaochharia and Grinstein, 2009).

Table 9 shows the regression results in terms of other measure for compensation. Across all specifications, the coefficients of *CEO_twitter* are significantly positively related to the delta, vega, firm_ related_ wealth and non_equity base. Thus, the evidence still supports our hypothesis 1.

Table 9. Robustness Checks: Alternative Measures of compensation

This table presents regression results for the effect of twitter on alternative measures of compensation. The equation form of model is:

$$Compensation_{i,t} = \alpha_0 + \alpha_1 CEO\_twitter_{i,t} + \beta' F_{i,t-1} + \theta' Z_{i,t} + \gamma_i + \mu_t + \varepsilon_{i,t} \quad (4)$$

where the dependent variable is Compensation (delta, vega, firm related wealth, non_equity base) and the main explanatory variable is CEO_twitter. Specifically, $Comp_{i,t}$ is the natural logarithm of delta,vega, firm related wealth, non_equity base alternatively for chief executives officer i in year t; $CEO\_twitter_{i,t}$ represents CEO with twitter account, takes the value of one if the executive of firm i adopted Twitter as of year t, and zero otherwise. $F_{i,t-1}$ is a vector of control variables (firm characteristics) for firm i in year t-1. $Z_{i,t}$ is the vector of the control variables for corporate governance and CEO characteristics i in year t. γ_i and μ_t represent the fixed effects of industry and year, respectively, and $\varepsilon_{i,t}$ represents the error term of the regression. In all specifications, the *t-statistics* reported are based on heteroscedasticity at firm-level-robust standard errors (White, 1980). The sample period is 2006-2016. All the variables are defined in Appendix A *, ** and *** denote the significance level of 10%, 5% and 1%, respectively.

	(1)	(2)	(3)	(4)
	delta	vega	firm_related_wealth	nonequitybase
CEO_twitter	0.1943**	0.3401***	0.1748*	0.1544***
	(2.15)	(2.74)	(1.94)	(3.34)
CEO5pct	0.2093**	-0.1926	0.2466***	-0.4006***
	(2.54)	(-1.35)	(2.91)	(-8.41)
CEO_firstyear	0.2823***	-0.2332	0.3260***	-0.1297***
	(2.69)	(-1.48)	(3.13)	(-2.58)
CEO_tenure	0.2725***	0.0506	0.2921***	-0.0385***
	(9.55)	(1.31)	(10.19)	(-2.69)
CEO_chair	0.0290	0.1352	0.0022	0.2497***

	(0.38)	(1.47)	(0.03)	(6.05)
Bdsize	0.1460***	0.2000***	0.1345***	0.1146***
	(14.95)	(7.23)	(13.96)	(18.64)
Indcompcom	0.2569***	0.1261	0.1866**	0.3552***
	(2.93)	(1.14)	(2.26)	(7.72)
BMV	-0.7171***	-0.5855***	-0.6775***	-0.2419***
	(-9.82)	(-4.22)	(-9.22)	(-7.57)
DE	0.0364**	0.0547***	0.0363**	0.0536***
	(2.49)	(3.06)	(2.48)	(7.67)
ROA	1.2803***	0.8597	1.3129***	0.0555
	(4.15)	(1.50)	(4.22)	(0.35)
RET	0.2067***	-0.1464	0.2413***	0.1809***
	(4.58)	(-1.62)	(5.29)	(7.96)
Std5RET	-4.6956***	-6.1789***	-3.3527***	-2.8489***
	(-8.83)	(-7.57)	(-6.30)	(-10.03)
Std5ROA	-0.2090	0.0930	-0.1007	-0.1009
	(-0.61)	(0.16)	(-0.29)	(-0.52)
Constant	3.6212***	2.8988***	7.6886***	7.6640***
	(8.18)	(5.62)	(17.37)	(37.74)
Control for				
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Obs	7444	6345	7445	10099
Adj. R^2	0.1697	0.0420	0.1579	0.2605

CONCLUSION

Our study is the first to document and describe how the effect of chief executive officer using Twitter on their compensation. We examine the potential consequences of such behavior for the underlying firm. Twitter was created in 2006 and undoubtedly is one of the most popular social media disclosure platforms in USA with short format (140-character limitation). With the complete list of all CEOs in Execucomp, we locate users with active Twitter accounts that have the same first and last names as the CEO in question. We then cross-check the executives' middle names, gender, and company information with user characteristics; we also read tweets to determine whether any account that we find does indeed belong to the executive in question. Through this labor-intensive process, we determine that 336 S&P 1500 CEOs have active personal Twitter accounts and work for firms that have the data necessary to conduct our tests. Using a sample from 2006 to 2015 of a total 10,093 observations in the U.S., we find an evidence that CEO_twitter is significantly positive related to total compensation (*TDC1*), confirming *CEO_twitter* has a real effect on compensation. In addition, the effect is more salient in the sample of higher information asymmetry. Third, the effect of CEO Twitter becomes stronger when firms have a high product market competition. Besides, we also examine whether posting more tweet can curtail the market responses to the earning announcements Overall, our findings support that CEO with twitter account receive more total CEO compensation. Together, our results point to the growing significance of social media in financial markets and show that social media activity can have important consequences for CEOs that engage in it.

APPENDIX 1: VARIABLES DEFINITION

Variables	Definition	Data source
Dependent variable		
Total annual compensation (TDC1)	Total compensation consists of salary, bonus, value of restricted stock granted, value of options granted (using Black–Scholes), long-term incentive payouts, and other compensation	Execucomp
Main independent variable		
CEO_twitter	CEO with twitter account, takes the value of one if the executive of firm i adopted Twitter as of year t, and zero otherwise	Hand collect
Firm characteristics		
BMV	Book to market ratio	Compustat
DE	Firm leverage, average total liabilities divided by average total equity	Compustat
ROA	Return on assets	Compustat
RET	Company's annual common stock return	Compustat
Std5RET	Standard deviation of common stock returns over the previous five years	Compustat
Std5ROA	Standard deviation of return on assets measured over the previous five years	Compustat
Assets	Natural logarithm of total assets in millions USD.	Compustat
Q	Tobin's Q	Compustat and CRSP
Leverage	Long term debt and debt in current liabilities divided by total assets.	Compustat
Tangibility	Net property, plant and equipment divided by total assets.	Compustat
Profitability	Earnings before interest, taxes, depreciation, and amortization (EBITDA) divided by total assets. This measure is adjusted by industry median of Profitability.	Compustat
Zscore	Modified Altman's Z-score (1.2 × working capital + 1.4 × retained earnings + 3.3×EBIT + 0.999×sales)/Total assets.	Compustat
CF-volatility	Standard deviation of Quarterly cash-flows from operations over the four fiscal years prior to the loan initiation year scaled by total assets.	Compustat
BHAR	Buy and hold abnormal return, $BHAR_{iT} = \prod_{t=1}^{T}(1+R_{it}) - \prod_{t=1}^{T}(1+R_{mt})$, where R_{it} is monthly stock return of firm *i* and R_{mt} is the monthly value-weighted CRSP market return. Similarly, BHAR3y and BHAR5y are defined for periods of three and five years.	CRSP
CEO characteristics and corporate governance		
CEO5pct	An indicator variable equal to 1 if the CEO owns 5% or more of the company's stock	Execucomp
CEO_firstyear	An indicator variable equal to 1 if it is the CEO's first year of service at a firm	Execucomp
CEO_tenure	The natural logarithm of the number of years of service of the current CEO	Execucomp
CEO_chair	An indicator variable equal to 1 which identifies if the CEO is also the board chairperson	Execucomp
Bdsize	Number of board directors	Execucomp

Indcompcom	An indicator variable (Indcompcom) equal to 1 if the compensation committee is comprised wholly of independent directors.	Execucomp
Market competition variables		
CP _oriented	The ratio of advertising expenses to sales revenue	Compustat
Similarity	A degree that firm's product is nearer to rivals	Hoberg and Phillips (2010, 2014)
Fluidity	A degree of rival firm's products change relative to firm's products	Hoberg et al. (2016)
Earning announcement, crash risk, and information asymmetry		
SUE_mean (SUE_median)	Difference between actual and forecast EPS, scaled by standard deviation of forecast EPS by analysts. Forecast EPS is mean or median of analysts' forecasts.	I/B/E/S
CAR[a;b]	Cumulative abnormal returns in event windows [a;b] of the earnings announcements. The estimation windows are [-230; -31] and are required at least 100 observations in the estimation window. Returns in event windows are adjusted by Fama-French three-factor model.	Compustat
Firm age	Firm age is calculated since the first year with non-missing price in CRSP database (CHECK THE REFERENCE THAT WE FOLLOW)	CRSP
Coverage	Number of estimates that constitute EPS Forecast from I/B/E/S Unadjusted Summary Data (WRDS file name is *ibes.statsumu*).	I/B/E/S
Has Rating	Dummy variable, equals to one if the firm has S&P rating in a year	Compustat

* CRSP: Center for Research in Security Prices

APPENDIX 2: SIC CLASSIFICATION

SIC2	Observations	CEO twitter	Without CEO twitter	Standard Industrial Classification
10	60	3	57	Metal, Mining
12	44	2	42	Coal Mining
13	587	9	578	Oil & Gas Extraction
14	39	0	39	Nonmetallic Minerals, Except Fuels
15	135	7	128	General Building Contractors
16	72	0	72	Heavy Construction, Except Building
17	35	0	35	Special Trade Contractors
20	452	21	431	Food & Kindred Products
21	33	4	29	Tobacco Products
22	50	0	50	Textile Mill Products
23	174	0	174	Apparel & Other Textile Products
24	108	5	107	Lumber & Wood Products
25	112	0	112	Furniture & Fixtures
26	184	0	184	Paper & Allied Products
27	150	9	141	Printing & Publishing
28	1139	24	1115	Chemical & Allied Products
29	98	0	98	Petroleum & Coal Products
30	106	1	105	Rubber & Miscellaneous Plastics Products
31	75	0	75	Leather & Leather Products
32	87	3	84	Stone, Clay, & Glass Products
33	197	6	191	Primary Metal Industries
34	215	3	212	Fabricated Metal Products
35	810	76	734	Industrial Machinery & Equipment
36	1162	49	1113	Electronic & Other Electric Equipment
37	382	6	376	Transportation Equipment
38	833	53	780	Instruments & Related Products
39	105	7	98	Miscellaneous Manufacturing Industries
40	45	0	45	Railroad Transportation
41	10	0	10	Local & Interurban Passenger Transit
42	147	0	147	Trucking & Warehousing
44	78	0	78	Water Transportation
45	138	13	125	Transportation by Air
47	61	16	45	Transportation Services
48	444	43	401	Communications
49	802	5	797	Electric, Gas, & Sanitary Services
50	307	25	282	Wholesale Trade – Durable Goods
51	188	6	182	Wholesale Trade – Nondurable Goods
52	37	4	33	Building Materials & Gardening Supplies

53	131	13	118	General Merchandise Stores
54	72	3	69	Food Stores
55	115	7	108	Automotive Dealers & Service Stations
56	274	6	268	Apparel & Accessory Stores
57	66	4	62	Furniture & Home furnishings Stores
58	267	40	227	Eating & Drinking Places
59	256	20	236	Miscellaneous Retail
60	1129	32	1097	Depository Institutions
61	143	15	128	Non depository Institutions
62	390	20	370	Security & Commodity Brokers
63	688	11	677	Insurance Carriers
64	100	5	95	Insurance Agents, Brokers, & Service
65	59	3	56	Real Estate
67	947	31	916	Holding & Other Investment Offices
70	47	0	47	Hotels & Other Lodging Places
72	46	0	46	Personal Services
73	1466	275	1191	Business Services
75	38	0	38	Auto Repair, Services, & Parking
78	56	17	39	Motion Pictures
79	106	12	94	Amusement & Recreation Services
80	310	11	299	Health Services
82	107	19	88	Educational Services
83	11	0	11	Social Services
87	225	19	206	Engineering & Management Services
99	79	7	72	Non-Classifiable Establishments

REFERENCES

Akins, B.K., Ng, J., and Verdi, R.S. (2012). Investor competition over information and the pricing of information asymmetry. *Accounting Review*, 87, pp. 35–58.

Antweiler, W., and Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, 59, pp. 1259–94.

Armstrong, C.S., Core, J.E., Taylor, D.J., and Verrecchia, R.E. (2011). When does information asymmetry affect the cost of capital? *Journal of Accounting Research,* 49, pp. 1–40.

Autore, D.M., and Kovacs, T. (2010). Equity issues and temporal variation in information asymmetry. *Journal of Banking and Finance*, 34, pp. 12–23.

Baik, B., Kang, J.K., and Kim, J.-M. (2010). Local institutional investors, information asymmetries, and equity returns. *Journal of Financial Economics*, 97, pp. 81–106.

Banker, R.D., Darrough, M.N., Huang, R., and Plehn-Dujowich, J.M. (2013). The Relation between CEO Compensation and Past Performance. *The Accounting Review*, 88, pp. 1-30.

Bartov, E., Faurel, L., and Mohanram, P.S. (2018). Can Twitter Help Predict Firm-Level Earnings and Stock Returns? *The Accounting Review*, 93, pp. 25-57.

Blankespoor, E., Miller G.S., and White, H.D. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *Accounting Review*, 89, pp. 79-112.

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science,* 2, pp. 1-8.
Boone, A.L., and White, J.T. (2015). The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics*, 117, pp. 508–533.
Brown, S., and Hillegeist, S.A. (2007). How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies*, 12, pp. 443–477.
Bugeja, M., Matolcsy, Z.P., and Spiropoulos, H. (2012). Is there a gender gap in CEO compensation? *Journal of Corporate Finance*, 18, pp. 849-859.
Bushee, B., and Miller, G. (2012). Investor relations, firm visibility, and investor following. *The Accounting Review,* 87, pp. 867–897.
Bushee, B.J., and Leuz, C. (2005). Economic consequences of SEC disclosure regulation: Evidence from the OTC bulletin board. *Journal of Accounting and Economics*, 39, pp. 233-264.
Chae, J. (2005). Trading volume, information asymmetry, and timing information. *Journal of Finance*, 60, pp. 413–442.
Chalmers, K., Koh, P., and Stapledon, G. (2006). The determinants of CEO compensation: rent extraction or labour demand? *The British Accounting Review,* 38, pp. 259–275.
Chen, H., De, P., Hu, J., and Hwang, B.-H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, 27, pp. 1367-1403.
Chen, H., Hwang, B.H., and Liu, B. (2017). Economic consequences of social media adoption by CEOs and CFOs. *Working paper*.
Chen, H., Hwang, B.H., and Liu, B. (2018). The Emergence of "Social Executives" and Its Consequences for Financial Markets. *Working paper*.
Chhaochharia,V., and Grinstein, Y. (2009). CEO compensation and board structure. *Journal of Finance*, 64, pp. 231-261.
Core, J., and Guay, W. (2002). Estimate the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research,* 40, pp. 613-630.
Core, J., Holthausen, R., and Larcker, D. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*, 51, pp. 371–406.
Curtis, A., Richardson, V., and Schmardebeck, R. (2016). Investor attention and the pricing of earnings news. *Handbook of Sentiment Analysis in Finance (Chapter 8),* pp. 212–232.
Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *Journal of Finance*, 66, pp. 1461–99.
Daniel, N., Li, Y., and Naveen, L. (2013). No asymmetry in pay for luck. *Working Paper*.
Das, S., and Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the Web. *Management Science*, 53, pp. 1375–1388.
Diamond, D.W., Verrecchia, R.E., 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance*. 46, 1325-1359.
Drake, M., Roulstone,D., and Thornock, J. (2012). Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research*, 50, pp. 1001–1040.
Easley, D., and O'hara, M. (2004). Information and the cost of capital. *Journal of Finance*, 59, pp. 1553–1583.
Elliott, W.B., Grant, S.M., and Hodge, F.D. (2018). Negative News and Investor Trust: The Role of $Firm and #CEO Twitter Use. *Journal of Accounting Research*, 56, pp. 1483-1519.
Fama, E.F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics,* 49, pp. 283–306.
Fang, L., and Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, 64, pp. 2023–2052.
Fu, R., Kraft, A., and Zhang, H. (2012). Financial reporting frequency, information asymmetry, and the cost of equity. *Journal of Accounting and Economics*, 54, pp. 132–149.

Gallaugher, J., and Ransbotham, S. (2010). Social media and customer dialog management at starbucks. *MIS Quart. Executive*, 9, pp. 197–212.
Hartzell, J.C., and Starks, L.T. (2003). Institutional Investors and Executive Compensation . *Journal of Finance,* 58, pp. 2351-2374.
Hirschey, M., Richardson, V., and Scholz, S. (2000). Stock price effects of Internet buy-sell recommendations: The Motley Fool case. *Financial Review*, 35, pp. 147–174.
Hirshleifer, D., Low, A., and Teoh, S.H. (2012). Are Overconfident CEOs Better Innovators? *Journal of Finance*, 67, pp. 1457–98.
Hoberg, G., and Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies*, 23, pp. 3773–3811.
Hoberg, G., and Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124, pp. 1423–1465.
Hoberg, G., Phillips, G., and Prabhala, N. (2014). Product market threats, payouts, and financial flexibility. *Journal of Finance*, 69, pp. 293–324.
John, K., Mehran, H., and Qian, Y. (2010). Outside monitoring and CEO compensation in the banking industry. *Journal of Corporate Finance*, 16, pp. 383-399.
Jung, M. J., Naughton, J. P., Tahoun, A., and Wang, C. (2018). Do firms strategically disseminate? Evidence from corporate use of social media*. The Accounting Review, 93, pp.* 225-252.
Kane, A., Lee, Y.K., and Marcus, A. (1984). Earnings and dividend announcements: Is there a corroboration effect? *Journal of Finance*, 39, pp. 1091–1099.
Kaniel, R., Liu, S., Saar, G., and Titman, S. (2012). Individual investor trading and return patterns around earnings announcements. *Journal of Finance*, 67, pp. 639–680.
Karaduman İlkay. (2013). Procedia - *Social and Behavioral Sciences*. 99, pp. 465 – 473.
Kyle, A.S. (1985). Continous acutions and insider trading. *Econometrica*, 53, pp. 1315–1335.
La Porta, R., Lakonishok, J., Shleifer, A., and Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *Journal of Finance*, 52, pp. 859–874.
Laroche, M., Habibi, M.R., and Richard, M. (2013). To be or not to be in social media: How brand loyalty is affected by social media? *International Journal of Information Management*, 33, pp. 76-82.
Leary, M.T., and Roberts, M.R. (2010). The pecking order, debt capacity, and information asymmetry. *J. Financ. Econ*, 95, pp. 332–355.
Lee, F., Hutton, A., and Shu S. (2015). The role of social media in the capital market: Evidence from consumer product recalls. *Journal of Accounting Research,* 53, pp. 367–404.
Leuz, C., and Verrecchia R.E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research,* 38, pp. 91-124
Luo, X., Zhang, J., and Duan, W. (2013). Social Media and Firm Equity Value. *Information Systems Research*, 24, pp. 146-163.
Malmendier, U., G. Tate, G., and Yan, J. (2011). Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies. *Journal of Finance*, 66, pp. 1687–1733.
Malmendier, U., and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance*, 60, pp. 2661–2700.
Malmendier, U., and Tate, G. (2008). Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction. *Journal of Financial Economics*, 89, pp. 20–43.
Malmendier, U., and Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics*, 124, pp. 1593–1638.
Mao, Y., Wei, W., Wang, B., and Liu, B. (2012). Correlating S&P 500 stocks with Twitter data. *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks*, pp. 69–72.

May, R.G. (1971). The influence of quarterly earnings announcements on investor decisions as reflected in common stock price changes. *Journal of Accounting Research*, 9, pp. 119–163.
Miller, G. (2006). The press as a watchdog for accounting fraud. *Journal of Accounting Research*, 44, pp. 1001–1033.
Miller, G.S., and Skinner, D.J. (2015). The Evolving Disclosure Landscape: How Changes in Technology, the Media, and Capital Markets Are Affecting Disclosure. *Journal of Accounting Research*, 53, pp. 221-239.
Naaman, M., Boase, J., and Lai, C.-H. (2010). Is it really about me? Message content in social awareness streams. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, pp.189-192.
Shroff, N., Sun, A.X., White, H.D., and Zhang, W. (2013). Voluntary disclosure and information asymmetry: Evidence from the 2005 securities offering reform. *Journal of Accounting Research*, 51, pp. 1299–1345.
Tang, T.T. (2009). Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement. *Journal of Financial Economics*, 93, pp. 325–351.
Tetlock, P.C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62, pp. 1139–1168.
Tetlock, P.C. (2010). Does public financial news resolve asymmetric information? *Rev. Financ. Stud*, 23, pp. 3520–3557.
Tumarkin, R., and Whitelaw, R. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*, 57, pp. 41–51.
Verrecchia, R. E. (2001). Essays on disclosure. *Journal of Accounting and Economics*, 32, pp. 97–180.
Wade, J., O'Reilly, I., and Chandratat, I. (1990). Golden parachutes: CEOs and the exercise of social influence. *Administrative Science Quarterly*, 35, pp. 587–603.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, pp. 817–838.
Xu, S. X., and Zhang, X. (2013). Impact of Wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quarterly*. 37, pp. 1043-1068.
Yermack, D. (1995). Do corporations award CEO stock options effectively? *Journal of Financial Economics*, 39, pp. 237–269.

DECISION SCIENCES INSTITUTE

Achieving Blue Ocean Strategy by Operations Management

Jenny Tran
Monarch University
Email: jenny.tran@umonarch-mail.ch

Barin N. Nag
Towson University and Monarch University
Email: bnag@towson.edu

ABSTRACT

The emphasis is on Blue Ocean markets and businesses are struggling to create more products with greater value and lower cost. Innovation delivers value in a frugal product and process. Achieving competitive advantage involves efficient management of all aspects of the operation consistent with the applying principles of Operations Management.

KEYWORDS: Operations Management, Blue Ocean, Product Design, Value Added, Low Cost

INTRODUCTION

A current issue that has been addressed by many business leaders is that the business environment is changing as a result of increasing globalization, economic changes, changes in personal preferences, and digitization. The uncertainty that this has caused has made the global market more volatile, complex, and ambiguous. The one constant factor is that there will be a growing demand for products that are of higher quality and of lower cost. This is the essence of the Blue Ocean strategy that aims for products that are of greater value to the customer and are a combination of more for less. The more for less syndrome has transformed from a short-term response to market needs to a long-term strategy. Businesses that realize this reality stand to prosper and those that continue to live in denial will suffer.

It appears that enterprises are in the race to create more blue ocean markets for the future, with far greater speed compared to the past. Firms, whether well-established or in the small-to-medium range, with their varied advantages and disadvantages, are left with few options other than innovating new products and services, Even business models to be focused increasingly on sustainability, have to be customer-centric, and have to have value-innovation goals. The management of the processes of converting input resources into outputs as goods or services, that is known as Operations Management (OM) plays an ever-important role for firms to establish competitive advantages in a business environment that has sharply changed due to penetrating globalization and digitization. Operations hold a vital position in any business unit, in pursuing increasing productivity and putting the organization in a more competitive position in the market, to increase revenues and market share, with greater profitability, and higher levels of quality and customer satisfaction.

Whether with consumers beginning to actively participate in product creation in areas such as product/ service innovation, supply chain management in an expanding international business platform, or managing human resources with a rising percentage of knowledge workers, firms need to utilize Operations Management to improve an R&D approach tailored to customer needs for mass customization, creating sustainable solutions and co-creating values with prosumers, group of consumers participating in the producer process (Prabhu & Radjou, 2015).

All of these efforts have the goal of establishing distinctive competitive advantages. This paper will discuss how competitive barriers removed by the speeding up of changes in the global business and social environment impact the way businesses are run, and the role of Operations Management principles and techniques in aiding firms to stay relevant and ahead of the competition. The business landscape is giving rise to competition from emerging markets, and is attuned to frugal innovation, which forces developed economies to rethink and innovate products and systems and compete in the global market. The process of competition even has an impact on the core business model, forcing innovation in areas that no one had thought to look before.

THE RISE OF FRUGAL ECONOMIES AND DIGITAL PENETRATION IN BUSINESSES

Widespread economic changes, advancements in technology, the rising growth of emerging markets and deep penetration of digital are reasons that give rise to frugal innovation in the developed world. It is of special note that this trend of producing more and better with less resources started from resource-constrained economies, then gradually spread to the more developed countries and has become the new norm in innovation as well as operation management across many industries.

Advanced economies have entered an age of relative austerity in which the idea of frugal living and consuming is becoming popular. In the recent decade, the middle classes in the US, Canada, Europe, Japan, Australia and elsewhere have witnessed their incomes deteriorate and their purchasing power shrink. “Adjusted for inflation, the real median household income in the US increased by merely 19% between 1967 and 2013. According to a 2014 Pew Research survey, only 44% of Americans now define themselves as middle class, a nine-percentage point drop since 2008. Over the same period, the number of Americans who believe that they have fallen into a lower income class has shot up from 25% to 40%. Moreover, since 2009, 95% of all income gains in the US have gone to the top 1% of earners. The richest 20% of Americans now account for over 60% of the country’s consumer spending” (Prabhu & Radjou, 2015, p. 5). With rising inflation and shrinking income, consumers have become more sensitive for pricing when considering purchases. In parts of Europe where recession persists, intensifying poverty is further destroying middle-income purchasing power. Spain and Greece have suffered the most, but more affluent states such as France and Germany are no exception. “For instance, only 58% of Germans now identify themselves as middle class compared with 65% in 1997. In France, between 2008 and 2012, average salaries fell by 24%, while the cost of living rose by 30%” (Prabhu & Radjou, 2015, p. 5-6). As a result of these economic pressures, North American and European consumers are becoming increasingly aware of getting value when they spend and are selecting less expensive products. For example, nearly 30% of European consumers, especially young people who have only experienced recession in their adult lives, are now more attracted in buying a low-cost rather than a premium car. Car purchases by Americans aged 18-34 fell 30% between 2007 and 2012 (Prabhu, Radjou, 2015). Japan sets disposable income below $14,424 as the poverty level. In 2013, in Japan there was 12 percent of the national population under the poverty level (Lu, 2018), consumers are shifting from premium brands to low-priced private-label products in retail stores. Rather than eating out, more Japanese workers now pack their own lunch.

These changes are not temporary but look likely to be persistent. Thomas Piketty, a French economist, forecasts that income inequality in established economies will broaden in the forthcoming decades, as long-term annual growth rates remain stuck below 2%. With inflation overtaking their incomes since 2007, 76% of US adults now believe that in the future their children will be financially worse off than them. And well over half of consumers, surveyed by Booz & Company (now Strategy&), a global management consultancy, reported in late 2012 that they would not revert back to their previous spendthrift behavior even when times improve. Booz calls these frugal buyers “permanently value-sensitive consumers” (Prabhu & Radjou, 2015, p. 6).
It is not only consumers who have become more cost-conscious. Governments throughout the developed world are also tightening budgets. Ageing populations, escalating health-care expenses,

the pensions burden, huge debts and deficits since 2008 have combined to introduce a new spirit of austerity in the US, Europe and Japan. However, within this landscape, the frugal innovation revolution is about more than austerity. Consumers in the developed world are becoming not only more value conscious but also more *values* conscious. They increasingly care about social harmony, worry about ecological degradation and the depletion of natural resources, and want businesses to play their part in making the world better. 81% of consumers say they will make personal sacrifices to address social and environmental issues (Anonymous, 2016). According to the 2014 Edelman Trust Barometer, an annual global study of consumer attitudes, 84% of respondents believe that business can pursue self-interest while doing good work for society. As Carol Cone, who heads Edelman's social business practice argues, that "to increase trust levels among consumers – and earn the 'license to lead' – businesses must learn to simultaneously create more operational and societal value while also impacting less the environment. "Citizen consumers" will vote with their wallets for brands that are socially inclusive and environmentally active." (Prabhu & Radjou, 2015).

A slightly earlier Report, the *2013 Cone Communications Green Gap Trend Tracker,* suggests that 71% of Americans consider the environment when they shop. This is up from 66% in 2008 (Anonymous, 2014). The survey concluded that 7% of individuals consider the environment every time they shop, while 20% consider it regularly and 44% sometimes consider it. Over 80% of Europeans believe that a product's environmental impact is a critical element in their purchasing decisions. 84% of consumers globally report they prioritize socially and environmentally responsible products whenever possible (Anonymous, 2016).

More troublingly for businesses, 90% of millennials (the 70 million or so Americans in their 20s and early 30s), who contribute $180 billion annually to spending, are inclining to shift to more socially and environmentally responsible brands. Despite their tight budgets, millennial consumers also seek for high quality and sustainable products. To gain the trust of these values-conscious consumers, Cone believes that companies must "move beyond transactional thinking towards a better understanding of tangible actions needed to solve critical societal issues necessary to mutual benefit" (Prabhu & Radjou, 2015). As best practices, Cone cites Gap raising its employees' minimum wage; and Starbucks offering employees who work at least 20 hours a week tuition settlement. (Prabhu & Radjou, 2015).

Governments across the developed world are also playing a role in this process. New regulations require businesses to be resource efficient. The US under the leadership of President Obama passed a new law requiring US carmakers to improve fuel efficiency from the current average of 27.5 miles per gallon to 54.5 miles per gallon by 2025. Similarly, in 2012, the European Parliament passed a stricter recycling law that requires by 2020 electronic and electrical goods suppliers and retailers to collect, and potentially recycle, 85% of the electric and electronic scrap they generate. In early 2014 the European Parliament voted in favor of requiring member states to meet stricter binding national targets for 2030 to deal with climate change. The directive includes a 40% reduction in greenhouse gases (compared with 1990 levels) and at least 30% of energy to come from renewable sources. As noted by the European commissioner for the environment, a combination of utilizing environmental benefits and innovative growth opportunities is needed to counter the economic turmoil and rising prices for raw materials of the present time.

Perhaps politicians are aware that consumers' standards are shifting, valuing quality over quantity. Studies show that between 15% and 28% of Americans have willingly reduced their material belongings in favor of greater self-sufficiency, with the goal of leading a humbler and more meaningful life. In Japan, a country famous for intensive working hours, almost 50% of consumers across generations now spend more time at home. Across the developed world – from New York to Paris to Tokyo – consumers now view frugality as a means to *increase,* rather than decrease, their quality of life. It is evident that consumers are looking for a substantive improvement in the quality of life and are not to be satisfied with the appearances of it. The positive side of the post-crisis gloom is that citizens' search for a more balanced lifestyle is helping to create a new economic system – a frugal economy. It represents an improvement of the more excessive trends of 20th century overconsumption and waste.

In addition to the changes in consumers' perception and behavior already noted for the Western world, there are similar growing influences among the largest emerging consumer economies of China and India. While they have to innovate, and grow with huge consumer populations, constraints on resources that have forced developed economies to rethink how to design products/ services have created an awareness of untapped potentials from the market at the bottom of the pyramid. Markets in the developing economies can nurture global business through their sheer size, rate of growth and consumer demands. Everyone knows that the world's poor are distressingly plentiful. Fully 65% of the world's population earns less than $2,000 each per year—that's 4 billion people (Hammond, & Prahalad, 2002). The market for goods and services serving the world's poor – families with an annual household income of less than $6,000 – is massive. The 18 largest emerging and transition countries include 680 million such households, with a total annual income of $1.7 trillion – roughly equal to Germany's annual gross domestic product. Brazil's poorest citizens comprise nearly 25 million households with a total annual income of $73 billion. India has 171 million poor households with a combined $378 billion in income. China's poor residents account for 286 million households with a combined annual income of $691 billion. Surveys show that poor households spend most of their income on housing, food, healthcare, education, finance charges, communications and consumer goods. Multinational corporations have largely failed to explore this global market, even though the rewards for doing so could be significant (Hammond, & Prahalad, 2004).

What these figures indicate is that while assuming that the poor have no money sounds obvious on the surface, it's not entirely true. While individual incomes may be limited, the cumulative buying power of poor populations is quite significantly large. The average per capita income of villagers in rural Bangladesh, for instance, is less than $200 per year, but as a group they are devoted consumers of telecommunications services. Grameen Telecom's village phones, which are owned by a single entrepreneur but used by the entire community, generate an average revenue of roughly $90 a month—and as much as $1,000 a month in some large villages. Customers of these village phones, who pay cash per usage, spend an average of 7% of their income on phone services—a far higher percentage than consumers in existing markets do (Hammond & Prahalad, 2002).

It's also incorrect to assume that the poor are too concerned with fulfilling their basic needs to squander money on luxury goods. In fact, the poor often do buy luxury items besides daily essential goods. For example, in the Mumbai shantytown of Dharavi, 85% of households possess a television set, 75% own a pressure cooker and a mixer, 56% own a gas stove, and 21% have telephones. One of the explanations for this phenomenon is that buying a house in Mumbai, for most people at the bottom of the pyramid, is not a realistic consideration. Neither is getting access to running water. They accept that condition, and rather than saving for tough periods, they spend their income on things that can immediately improve the quality of their lives (Hammond & Prahalad, 2002). These unexpectedly hidden potential markets present an opportunity for any firms that can understand the provision of appropriate solutions.

For instance, as markets become increasingly global, improvements through effective OM could be made in areas such as cost reductions in supply chain and market expansion, with revenue boost and increasing effectiveness as well as the speed of doing business being credited to the development of Internet and other digital technologies. But the expansion of opportunities on an international scale requires firms to adopt relevant global strategies. Three primary strategies that . Nag (2011) introduced in his book, based on Michael Porter's devised set of strategies for businesses to use are (1) Differentiation Strategy, (2) Cost-Leadership Strategy and (3) Response Strategy. A more recently developed strategy mentioned followed the three primary ones is Mass Customization Strategy. A firm can adopt one or a combination of these strategies to achieve its mission. There is no mandatory or set of structures firms must follow, because creating a strategy is an innovation process itself (Nag, 2011). Operations Management activities are based on and guided by organizational strategies. As enterprises drive to establish and maintain competitive advantage and to gain more market share of their target segments, one of the goals organizational strategy tries to achieve is to lay out a process designed with a view to realize the objectives. In this way Operations Management is becoming increasingly vital for firms to compete in the global digital era.

OPERATIONS MANAGEMENT ADJUSTMENTS TO ACHIEVE EFFICIENT MASS CUSTOMIZATION AND SUCCESSFUL PRODUCT AND PROCESS DESIGN

During the 20th century, established corporations such as General Motors (GM) and General Electric (GE) invested lavishly in R&D labs intending to industrialize and centralize the innovation process. These labs consisted of thousands of engineers and scientists charged with pushing technological boundaries in order to invent the next big innovation. These firms' industry dominance enabled them to drive their new products and services out to fairly passive customers. They employed large sales forces and spent weightily on mass marketing, especially TV and print media, to encourage demand. Successful as this model was for decades, its useful cycle is ending. The industrial R&D model, based on science and technology pursuit, is increasingly unfit to the 21st century's fast paced digital economy. In a world of growing financial constraints, resource scarcity and increased competition, including highly empowered, cost-conscious and eco-aware customers, there are several reasons why the industrial R&D model is losing its effectiveness. Some of the disadvantages that can be listed are: 1) High Investment for low value; 2) Time-consuming and inflexible R&D; 3) Rewarding innovation quantity over quality; 4) Complex, expensive and environmentally unfriendly products; 5) Alienating customers through treating them as passive users (Prabhu & Radjou, 2015).

The 20th century witnessed the birth of three significant organizational innovations: the corporate R&D lab (pioneered by Thomas Edison, who founded GE); mass production (developed and refined by Henry Ford); and "big-box" retail and mass distribution (developed by Sam Walton, the founder of Walmart). All three centralized company functions – from R&D and purchasing to manufacturing, sales and marketing – with the goal of generating economies of scale. Although this centralized approach helps mass production, thus reducing per unit costs, it also expends a lot of energy and has become costly to sustain. Even worse, this approach reduced not only consumers, but many employees also to being passive users of products and services, excluding them from involvement in a production process that, typically, would occur distantly from where they live. However, developed world consumers are rapidly and dramatically evolving into creative producers of personalized products and services. Much of this is thanks to 3D printing and do-it-yourself (DIY) platforms such as TechSchop and FabLab, which cut production costs. This new group of consumers participating in the producer process is termed *prosumer.* A new era of distributed manufacturing will be less resource intensive and will yield higher-quality, mass-customized products and services that are also affordable and sustainable (Prabhu & Radjou, 2015).
Due to the fact that many prosumers felt restricted in the real world by physical boundaries, they choose the online environment where they can create applications and communities with an unprecedented level of freedom of speech. These virtual prosumers develop their own features, becoming "knowledgeable consumers, digital users who employ technologies (software applications) for collaboration, have the ability to participate in a product or service conception, design, execution and/or testing and have a certain impact on their social network" (Izvercianu & Seran, 2011).

Companies need to differentiate their products or services to gain market share. Because innovation is considered to be indispensable to competitive advantage (Belardo, & Belardo, 2002) with customer relationships creating the condition for market acceptance, firms need to focus on ways to involve prosumers into co-creating innovation alongside companies. The results of such an activity present a mutual favorable situation, with benefits for both market stakeholders. The traditional focus of a company has always been on prioritizing its own objectives and resources, with relative relationship marketing results like commitment, trust and loyalty. Another new priority for the company to consider focusing on is prosumer orientation. In addition to this, by integrating consumers and especially prosumers in new product or service processes, firms can achieve a wide range of benefits like: new product development (Enkel, 2005) or service design and improvement (Sigala, 2012), which ultimately determine innovation.

Some approaches managers should consider in integrating prosumers into the product or services innovation and creation are:
1) Engage target consumers throughout product and customer life cycles via direct interaction instead of trusting secondary market research. R&D teams would be most efficient in identifying realistic and unmet needs if they are exposed to customers in their natural settings. This continuous observation throughout the product development cycles will expose R&D engineers to relevant customer insights, which might save the whole product design process time and finance as well as other resources in developing products that meet real market demands.
2) Seek affordable, good-enough solutions to the immediate unmet needs faced by customers instead of trying to impress them with technical complexity or widgets, complicating the best simple solution to meet their requirements.
3) Involve customers from the outset, share prototypes with end-users. Firms should actively seek ways to engage consumers in the early stages of product development to ensure that the result is a market-desirable offering. By designing just good-enough, usable products R&D teams could already be sharing their prototypes with early adopters in the market, via some online platform such as Affinova's tools for greater exposure, hence resulting in more relevant feedback to the process.
4) Utilize big data analytics. As discussed earlier, with increasingly connected consumers to the internet and digital channels, mobile phones and digital networking allow researchers to tap into huge amounts of detailed and specific data sets to detect customer needs and trends, so that firms could respond with appropriate solutions (Prabhu, & Radjou, 2015).
In order to support the innovation process, there are considerations for firms to take when it comes to execution agility to improve efficiency and effectiveness as well as keeping the cost down during each innovative stage. With resource scarcity, R&D teams could enjoy the advantages of using dynamic portfolio management tools to identify priorities and focused projects. The portfolio management will reduce the risk of unfocused and wasteful scattering allocated resources to develop products that might not match with real market demands. Another activity to prevent firms from developing over-engineered products is using just-in-time design method, where good enough prototypes are presented to consumers and more functions/ features are only added based on feedback of early users (Prabhu, & Radjou, 2015).

OPERATIONS MANAGEMENT SHIFT FROM ECONOMY OF SCALE IN INDUSTRIAL ERA TO FRUGAL SUPPLY CHAIN MANAGEMENT AND MANUFACTURING

The future role of supply chain strategies will rely on a thorough understanding of today business environment's challenges. Supply Chain Management issues are no longer considered simply a tactical, manufacturing, or logistics issue, but a strategic capability that is critical to the success of the company. This section will discuss some observation of the shifting trends in supply chain management and manufacturing to adapt to today's business.

Local sourcing:

The business world is now witnessing a shift from low cost to local sourcing. Western multinationals are already sourcing from local suppliers in emerging markets as part of localization strategies. Apart from meeting local sourcing requirements, this also allows firms to create more affordable products. This trend is becoming more apparent in mature economies too, where sourcing smaller quantities from smaller firms located near factories and R&D facilities reduces costs and risks.

Motorola, a multinational telecoms company, decided to manufacture all its Moto X smartphones in Texas so that its manufacturing engineers in Texas and its R&D team in Illinois and California could collaborate more closely with local suppliers and respond more quickly to demanding US consumers (Prabhu, & Radjou, 2015).

These large manufacturers are taking their lead from big retailers. In 2010, for instance, Walmart, the US's biggest importer, devoted to doubling the sales of locally sourced products in the US by 2015. Waitrose, a UK supermarket chain, sources nearly 70% of its food from suppliers situated within a 30-mile radius of a location. This is an interesting return to much earlier market places, where the countryside around a large town supplied that town with food through local markets. In modern markets the risks of variations in supply and quality are offset by reduction in transportation and storage costs, the changes being managed by logistics and operations management.

Sharing resources

Rather than letting their production and distribution assets idle, some manufacturers now allow other firms, including rivals, to use their facilities. It has long been common in areas like Africa and India for competing telecoms providers to share mobile-phone towers, and Western telecoms companies are now doing the same. Western firms are also learning from health-care firms in Africa which imitated Coca-Cola's "cold chain" (a temperature-controlled supply chain) as a cost-effective way to preserve life-saving medicine and have it delivered rapidly to remote villages. In many respects, the developments listed are the business-to-business equivalent of the sharing economy, in which companies trade and share supply-chain assets (Prabhu, & Radjou, 2015).

Distributing to the last mile

Fulfilling orders for customers in far-flung locations is a particular challenge. The so-called last-mile challenge exists because it is costly for companies to implement physical distribution (such as bank branches or retail stores) in locations with few users. Innovative distribution models, which make use of dependable locals and networks, are often used in emerging markets. The development of drones is likely to be the next major impactor on this problem (Prabhu, & Radjou, 2015).

Decentralize supply chains with smaller, nimbler factories

Decentralized supply chains, whereby production happens as close as possible to the point of consumption, will help bridge the supply-demand gap. One way to do this is by shifting production from big factories with inflexible manufacturing processes to smaller, nimbler plants that are flexible and versatile. Danone, a French multinational food products company, was capable to build in Bangladesh a yogurt-making micro-factory only 10% the size of Danone's existing factories and much lower cost to build. Impressed, Danone's chairman Franck Riboud asked his R&D and supply chain leaders to find out how Danone could build such low-cost micro-factories in other markets, including in Europe (Prabhu, & Radjou, 2015).

MANAGING HUMAN RESOURCES, JOBS, AND WORK IN A BUSINESS LANDSCAPE OF KNOWLEDGE WORKERS AND AN ENLARGED SERVICE SECTOR

Since the introduction and growth of digital channels and the internet, information is penetrating to whichever customer base that has access to the network. Consumers have become more informed than at any period in the past. In 2015, the International Telecommunication Union (ITU), a United Nations body, predicted that 3.2 billion people would be online, edging towards a half of the world population that is currently standing at 7.2 billion. About 2 billion of those will be in the developing world, the report added (Anonymous, 2015). This number rose from just 400 million internet users worldwide in 2008, an eighth of the 2015 figure. A growing and better-informed consumer base, deep penetration of digital channels and increasing advancement of technology all contribute their share in giving rise to knowledge workers. The management of this rising labor force is a central focus for any firms seeking to establish sustainable competitive advantages.

The change in access to information and knowledge poses significant challenges to the creation of business value. In an industrial economy business value, measured in terms of high rates of productivity and the associated costs, came first from machines and then from labor efforts. According to Drucker (1993), information-based and knowledge-based production had not resulted in higher productivity rates in service and knowledge work. Drucker explains this by, first of all, the immaterial and complex nature of information and knowledge as economic resources: that they cannot be treated and processed in the exact same way as material economic resources. Knowledge and information, Drucker argues, is not productive in and by itself; putting more information into a service does not make a person perform that service more effectively as such. The same is true in the case of the knowledge worker; his specialized knowledge is not productive in and of itself. Knowledge only becomes productive by fusing different kinds of specialized knowledge into something that makes a difference (Drucker, 1993). Also, and according to Drucker perhaps even more problematic, knowledge and information is intimately related to the knowledge worker, who now "collectively own the means of production" (Drucker, 1993).

Drucker believed that managerial success in a new economy rests on the aptitude to build a new organizational and managerial system that is able to differentiate productive activity from the busy-ness of people (Bang, Cleemann, & Bramming, 2010). Drucker proposes that a central premise must be to return to the fundamental strength of a business, meaning: "Accountability and measurability. It is the discipline of test, productivity measurements, and profitability requirement. Where these are lacking, businesses are essentially out of their depth" (Drucker, 1973). In the later part of the twentieth century, building on this agenda, accountability and measurability become the premises for business value, with work organized around the factory or firm as having a one task-purpose. However, due to the immaterial nature of information and knowledge, the "traditional factory" - the one task-purpose organization - cannot depend on traditional economic factors such as consumption and investment as indicators of business value.

Therefore, Drucker suggests that the focus of the modern factory must be at the level of costs: minimization of costs becomes the visible aspect of how to measure economic results, as costs are the only known factor in the organization's economy. A tight focus on costs, implying a continuous strive to lower costs, is key for the assessment of "performance" and "results". Consequently, Drucker suggests that all work is reorganized around the *task* and its flow, and to measure the costs of the organization by "time". It "has to be time" Drucker (1992) says, as the only thing that is "variable and controllable is how much time a given process takes. And benefit is whatever reduces that time" (Drucker, 1992). In effect, the production, the organizational structure, its management and employees all become factors that are organized around the task's time flows.

Due to growing automation in all business areas, the more it is likely to do a certain task the same way twice, the harder it is for companies to differentiate their offerings from the bloody competition. Innovation becomes the answer to this stringent problem, and it is highly reliant on intellectual capital. Thus, human capital and customer capital become primary sources of innovation. An extension of the concept that human capital refers to employee capabilities is that customer capital represents the relationship with important stakeholders (Potra, Izvercian, 2017).

Half a century ago, Peter Drucker presented the term “knowledge worker” to describe a new class of employee whose fundamental production was no longer capital, land or labor but, rather, the constructive use of knowledge. Today, these knowledge workers, a group that might be referred to as professionals, represent a large and growing percentage of the employees of the world biggest corporations. As stated by Bryan and Joyce (2005), in industries such as financial services, health care, high tech, pharmaceuticals, and media and entertainment, professionals now account for over 25 percent of the workforce, and in some cases undertake some of the key line activities. New business ideas are continuously generated among these groups of employees. They make it feasible for companies to deal with today's speedily-changing and ambiguous business environment, and they generate and administer the intangible assets that are the main way many companies across industries now create value. Dynamic professionals make big enterprises competitive. A few startups which have grown into giants in knowledge-based industries that attract global best talents are GAFAs (Google, Amazon, Facebook, Apple) (Galloway, 2017), Airbnb, Uber (Stone, 2018) and many more if somewhat smaller organizations.

Besides Drucker, another notable representative of this approach is Jack Vinson, who sees a knowledge worker as someone who depends on his knowledge and ability to learn, and who works with his brain (Vinson 2009). Lowe's (2002) definition of knowledge worker is limited to those with a university degree. Thomas Davenport (2005) sees knowledge workers as people with great levels of expertise, education, or intensive experience. Davenport states that the fundamental purpose of a knowledge worker's job involves the creation, distribution, or application of knowledge. Knowledge workers utilize their knowledge and think for a living (Davenport, 2005).

There is intangibility in the specifics of a knowledge workers' major tool and resource - knowledge. Knowledge consists of two dimensions, explicit and tacit. The explicit dimension is easy to solidify and share, by language, pictures and notes. The tacit dimension is associated to practical activity; it is greatly personal, partly or fully subconscious. It cannot be detached from its human owner. Due to tacit knowledge, upon which judgement is based, knowledge as a whole is intangible. The intangibility of knowledge makes knowledge workers special and difficult to manage. The most important part of the work of knowledge workers occurs in their heads even though the final result may be materialized. It cannot be observed, supervised or measured.

Many knowledge workers state that their finest ideas and solutions were conceived outside their professional working environment when they were relaxing, not officially at work. Even more, the results of the work of knowledge workers may differ from the short and long-term perspectives, which causes problems with standards, measurement and evaluation.

Literature offers some but not many concepts on how to manage knowledge workers. A literature review suggests that there is lot of room for future research on this topic. As a new rising and increasingly popular percentage group in the labor force, managing knowledge workers requires different approaches, tools and methods than managing traditional non-knowledge workers. As Peter Drucker conveyed, knowledge workers are employees who would oppose the command-and-control model that business took from the military 100 years ago (Brinkley, Fauth, Mahdon, Theodoropoulou, 2009). Managers cannot monitor the progression of the work of knowledge workers, which makes control or supervision challenging. Knowledge workers understand their work better than their managers and they are capable of and generally willing to be accountable for autonomous decision-making.

At the turn of the century Newell (2000) proposed six key requirements that are necessary to manage knowledge worker group productively: autonomy, achievement, keeping up to date, professional identification, participation in missions and goals, support and sharing. At much the same time Harman and Brelade (2000) advise managers to focus on six other factors important for knowledge workers: networking and broad contacts externally and internally, respect for individuals, creativity and innovation, trust, knowledge sharing and sound underlying systems and development. Suff and Reilly (2005) observe that knowledge workers are likely to thrive in some cultures, but also there are situations that inhibit their performance. Organizations should seek to develop the right climate. For example; is information sharing encouraged in practice or discouraged? Is direct involvement favored over indirect representative involvement? Is risk taking supported, tolerated or frowned upon. Knowledge workers will be sensitive to these organizational signals and will respond accordingly. They will not take risks if they think they will be penalized. They will not share information if they think it pointless. Knowledge workers will look after their careers and exhibit behaviors that help them further their ambitions in the context of what the organization appears to value" (Mládková, 2012).

A special task of the management of knowledge workers is the management of their performance. The CIPD (2004) emphasizes the need for a performance management system that knowledge workers feel comfortable with. The study says that knowledge workers are quite sensitive to performance management schemes because of the high degree of ownership they have over their skills and knowledge, which means they incline to take great pride in realizing progressive standards of performance. It recommends that knowledge workers be involved in the development of such performance evaluation structures, as knowledge-intensive outcomes tend to be vague and are often problematic to measure (Suff & Reilly, 2005).

Probably the most complex approach to the management of knowledge workers and their productivity is offered by the research of the Gallup Organization on motivation and management of so-called *talented employees* (Mládková, 2013). Research for over 25 years focused on employees' performance and loyalty to their organizations. The term 'talented people' is close to our term 'knowledge workers' and the results of the research are entirely relevant to this group of employees (Buckingham, & Coffman, 2005).

Analysis of the hugely extensive volume of data determined twelve key factors, summarized through twelve questions, that affect the behavior of an employee in an organization. These factors are needed to attract, develop and keep knowledge workers in the company because they are considered as crucial for their labor productivity. Factors are presented as questions knowledge workers ask and answer continuously. These twelve questions are:

- Do I know what is expected of me at work?
- Do I have the materials and equipment I need to do my work correctly?
- At work, do I have the opportunity to do what I do best every day?
- In the last seven days, have I received recognition or praise for good work?
- Does my supervisor, or someone at work, seem to care about me as a person?
- Is there someone at work who encourages my development?
- At work, do my opinions seem to count?
- Does the mission and purpose of my company make me feel like my work is important?
- Are my co-workers committed to doing quality work?
- Do I have a best friend at work?
- In the last six months, have I talked with someone about my progress?
- At work, have I had opportunities to learn and grow? (Buckingham & Coffman 2005).

The research also displayed great impact by direct managers on the performance of knowledge workers. (Buckingham, & Coffman, 2005; Mládková, 2012). The research on knowledge workers and their management shows that knowledge workers are proud experts whose performance is influenced more by good co-workers, the availability of contacts and knowledge than by HR policies and the material benefits that their organization offers them. They highly value it if their manager continuously communicates with them, if s/he creates an appropriate environment for their work and aligns their objectives with the objectives of the organization. They do not wish managers to interfere in their work but do not dislike being appropriately controlled.

CONCLUSION

Economic changes are happening, and happening globally, with greater speed than ever before. Firms from small to medium to large corporations are under severe pressure to keep up not only with innovation to stay relevant in their competitive environment but also being agile enough in operations to adapt to a fluctuating business landscape. Widening inequalities in income gaps in both developed and emerging markets present both opportunities as well as pressure for incumbent businesses to cope with. Increasing globalization, growth and penetration of a digital infrastructure, a growing percentage of knowledge workers and changing trends in consumers' behaviors all require managers of Operations Management to prepare not only for short-term competitive pressures but, more importantly, creating companies capable of innovating for long-term strategic competition with resource scarcity. Though there has been much research and suggested concepts for firms to utilize their operations management to adapt to changes, there are still a lot of space for creativity and future exploration. Operations Management must always be necessarily evolving if it is to respond effectively to the changing business landscape. It is not as it was in past decades. With today's sharply volatile business landscape, organizations and their executives need to be prepared with the most agile and lean structure, tangible as well as intangible resources to cope with changes and secure competitive advantages. Operations Management has the capacity to do that if it has the will to embrace frugal innovation.

REFERENCES

Anonymous (2014). US Environmental Concerns: Developments and Implications The United States' green consumerism trend presents significant opportunities for those Hong Kong enterprises that are environmentally-friendly and can clearly demonstrate their eco-compliance. *HKTDC Research*. Retrieved from https://hkmb.hktdc.com/en/1X09Z7YH/hktdc-research/US-Environmental-Concerns-Developments-and-Implications

Anonymous (2015). Internet used by 3.2 billion people in 2015. *Technology, BBC.* Retrieved from https://www.bbc.com/news/technology-32884867

Anonymous (2016). Study: 81% of Consumers Say They Will Make Personal Sacrifices to Address Social, Environmental Issues. *Sustainable Brands*. Retrieved from https://sustainablebrands.com/read/stakeholder-trends-and-insights/study-81-of-consumers-say-they-will-make-personal-sacrifices-to-address-social-environmental-issues

Bang, A., Cleemann, C. M., Bramming, P. (2010). How to create business value in the knowledge economy: Accelerating thoughts of Peter F. Drucker. *Management Decision*, 48 (4), 616-627.
Belardo, S., & Belardo, A. W. (2002). Innovation Through Learning: What Leaders Need to Know in the 21st Century. Whitestone, New York, NY.

Brinkley, I., Fauth, R., Mahdon, M., & Theodoropoulou, S. (2009). Knowledge Workers and Knowledge Work. *A Knowledge Economy Programme Report.* Retrieved from http://www.theworkfoundation.com/Assets/Docs/Knowledge%20Workers-March%202009.pdf

Bryan, L.L., & Joyce, C. (2005). The21st-century Organization. *McKinsey Quarterly*, New York (3), 24-33.

Buckingham, M., & Coffman, C. (2005). First, Break All the Rules. Simon & Schuster, London, UK.

Davenport, T. (2005). Thinking for Living. HVB School Publishing.

Drucker, P. F. (1973). Management – Tasks, Responsibilities, Practices. Harper & Row, New York, NY.

Drucker, P. F. (1993). Post-Capitalist Society. *Harper Business,* New York, NY.

Enkel, E., Perry-Freje, J., & Gassman, O. (2005). Minimizing market risk through customer integration in new product development: learning from bad practice. *Creativity and Innovation Management,* 14 (4), 425 – 437.

Galloway,S.(2017).Thefour:ThehiddenDNAofAmazon,Apple,Facebookand Google. Great Britain: Bantam Press.

Hammond, A. L, & Prahalad, C. K. (2002). Serving the world's poor, profitably. *Harvard Business Review*, 80 (9), 48-57.

Hammond, A. L, & Prahalad, C. K. (2004). Selling to the Poor. *Foreign Policy*, 142, 30-37.

Harman, C., & Brelade, S. (2000). Knowledge Management and the Role of HR. Financial Times, Prentice Hall.

Lowe, G. S. (2002). Leveraging the skills of Knowledge Workers. Isuma, 79 - 85.

Lu, J. (2018). World Update: Top 10 Facts About Japan Poverty. The Borgen Project. Retrieved from https://borgenproject.org/top-10-facts-about-poverty-in-japan/

Mládková, L. (2012). Management of Knowledge Workers. *European Conference on Knowledge Management*, 766-XXV.

Mládková, L. (2013). Knowledge Worker From the Perspective of Their Managers. *European Conference on Knowledge Management*, 446-452.

Nag, B. N. (2011). *Managing Operations in Manufacturing, Services and e-Business.* Deer Park, Linus Publication, Inc.

Newell, D. (2000). How to Retain Technical Professional. People Management.

Potra, S., Izvercian, M. (2017). Value Co-Creators in Today's Open Innovation. How Are They Like?. *European Conference on Intellectual Capital*, 254-260.

Prabhu, J., Radjou, N. (2015). Frugal Innovation: How to Do More with Less. The Economist.

Seran, S. A., & Izvercian, M. (2011). Defining and Involving the Virtual Prosumer Through Participative Marketing Strategies in the MPT Master Program. Todesco Publishing House, Cluj Napoca.

Sigala, M. (2012). Web 2.0 and customer involvement in new service development: a framework, cases and implications in tourism. *Tourism and Hospitality: Theory, Practice and Cases,* 44 (3).

Stone, B. (2018). The Upstarts: Uber, Airbnb and the Battle for the New Silicon Valley. *Great Britain: Bantam Press*.

Suff, P., & Reilly, P. (2005). In the Know: Reward and Performance Management of Knowledge Workers. *Institute of Employment Studies*. Retrieved from http://www.employment-studies.co.uk/pdflibrary/mp47.pdf

DECISION SCIENCES INSTITUTE
Conditions under which Synergistic Innovations Appear

Jeffrey Yi-Lin Forrest
Slippery Rock University
Email: jeffrey.forrest@sru.edu

Yaoguo Dang
Nanjing University of Aeronautics and Astronautics
Email: iamdangyg@163.com

Larry McCarthy
Slippery Rock University
Email: larry.mccarthy@sru.edu

Sifeng Liu
Nanjing University of Aeronautics and Astronautics
Email: sfliu@nuaa.edu.cn

Yong Liu
Jiangnan University
Email: clly1985528@163.com

ABSTRACT

We investigate holistically the questions: How can producer-side synergies be created by employing the strategy of economies of scope? And how can demand-side synergies be developed by making use of simultaneous consumer utilities and multi-sided markets? We describe how resources interact with each other and how simultaneous consumer utilities, two-sided markets and consumers' willingness to pay react to each other. By modeling business firm, resource, innovation, diversification, synergy, segment of consumers, etc., as interacting systems, we establish 8 propositions through examining how relevant systems exert forces on each other. This work provides practical managerial guidance for creating and capturing values.

KEYWORDS: competitive advantage, economy of scope, resource, simultaneous consumer utility, two-sided market

INTRODUCTION

Recent researches in strategic management suggest and demonstrate the need for firms to develop the necessary organizational culture and capability to effectively ride waves of transient competitive advantages (e.g., McGrath, 2013). For a firm to do so successfully it can look either inwardly at its nascent, heterogeneous resources and dynamic capabilities to see what values it can create for consumers (e.g., Barney, 1991; Eisenhardt & Martin 2000), or outwardly at product markets and consumer demands to find out where consumer synergies are located and what can be potentially developed (Drucker, 1954). No matter whether to look inwardly or outwardly, one way for a firm to possibly find potentials of different competitive advantages is to create economies of scope by diversification at either the producer side (Porter 1985; Santalo & Becerra 2008) or the demand side (Ye et al, 2012) or both. If a firm adopts this approach, then the firm has to address the following natural questions: (1) how can producer-side synergies be created when using the strategy of economies of scope? (2) How can demand-side synergies be developed by making use of simultaneous consumer utilities and multi-sided markets, after all consumers determine the success of companies (Penrose, 1959)?

Aiming at addressing these two questions, this paper develops a cohesive theory on how to potentially develop synergistically innovative ideas at either the producer side or the demand side. In this theoretical development, we rely one-hundred percent on logic reasoning, systems thinking and the intuitive yoyo model (Lin, 2009) without analyzing any set of data or anecdotes. Considering the fact that data mining and anecdotal analysis can only help uncover potential facts because of their constraints inherently existing in the available data/anecdotes and limitations existing in the analysis tools, results and conclusions established in this paper are more reliable and practically useful than those based on data and anecdotes.

In terms of the structure of our cohesive theory, we first spell out the basic assumptions and meanings of particular terms. Then by using the reasoning of logic and the methodology of systems science we develop formal propositions, which address various issues related to the two questions given above.

The rest of this paper is organized as follows. The next section reviews the relevant literature and demonstrates the contributions of this work. The following section provides the basics of systems science and methodology in order to make this presentation self-contained and the necessary assumptions and terminology in order to make the following logic reasoning rigorous. Then developed are a few insights on synergistic innovation on the producer side, and similar issues on the demand side. And this paper is concluded in the last section with managerial recommendations and problems for future research.

LITERATURE REVIEW

Because this paper contributes simultaneously to the literatures of innovation and strategy, in this section, we will look at the relevant literatures individually and show where our contributions lie in each case.

For the literature of innovation, over two hundred years ago Adam Smith (1776) acknowledges the importance of innovation. The topic of innovation of the producer side has been well investigated through the years by many scholars (Aas et al, 2015; Adner & Levinthal, 2001; Damanpour, 1991; Visnjic et al, 2016). In the business world, innovation is characterized as the introduction of original products and processes (Becheikh et al, 2006) that enable firms to competitively enter or create new markets (Smith & Tushman 2005). And it has been treated by many as one major factor of long-term performance (Kanter 2001). Veugelers and Cassiman (1999) find that high perceived risks and costs and low appropriation do not discourage innovation, but determines how the innovation sourcing strategy is chosen. They find that small firms are more likely to restrict their innovation strategy to an exclusive make or buy strategy, while large firms are more likely to combine both internal and external knowledge acquisition in their innovation strategy. With economic globalization and the advent of internet-based technologies, firms experience increasing pressures to become more competitive and innovative than ever before (Caputo et al, 2016; Buffington 2016; Zollo et al, 2016).

Contributing to this activity area of research, this paper establishes the following general result: by adding to a successful set X of resources or replacing some resources in X by other resources, a firm can potentially create to positive synergies.

For the literature of strategies, resource-based approaches tend to look inwardly at how heterogeneous resources and dynamic capabilities of firms can create value for consumers (e.g., Barney, 1991; Eisenhardt & Martin, 2000; Nason & Wiklund, 2018) through bundling internal resources in different ways (e.g., Sirmon et al, 2008). Because the concept of resources offers a framework for integrating multiple, dissimilar resources to explain their synergistic, differential effects on firms' performance and the contingencies associated with each effect (Fang et al, 2011), these approaches greatly contribute to knowledge of strategies (e.g., Lockett et al. 2009). At the same time, Barney (2001) and Priem and Butler (2001a, b) recognize that when looking instead outwardly, demand issues need to be addressed equally seriously, because after all consumers determine the success of firms (Drucker, 1954; Penrose, 1959).

Along this suggestion of outward looking at the demand side, scholars examine various strategic issues, such as how technological innovations and competitive advantages are affected by market demand (e.g., Adner & Zemsky 2006; Tripsas, 2008), how value can be created and captured by employing consumer-focused strategies (e.g., Priem 2007, Gans et al. 2008, Adner & Snow 2010), and how entrepreneurial innovation can be influenced by consumers (Sawhney et al. 2005, Shah & Tripsas 2007). More specifically, Ye et al. (2012) study how bundled preferences of consumers can be employed as a basis for developing firms' strategies in an effort that is parallel to the recent producer-focused synergy studies on economies of scope (e.g., Crossland & Smith 2002; Gary, 2005; Tanriverdi & Venkatraman, 2005).

By attending to both the demand side and the producer side of the literature of strategies, this paper establishes the following main results, among others, that are generally true: (1) if a firm makes its resources available to all business units in its attempt to produce an economy of scope, the firm will then be able to innovate synergistically. (2) Mobilizing inconsistent resources generally produce negative synergies. And (3) If a set of core resources is behind a firm's portfolio of offers, it is likely for the firm to focus on exploiting its existing competitive advantages instead of developing new advantages.

Other than contributions in the previously listed individual literatures, what is more important is that this paper holistically incorporates the demand side simultaneously into the studies of strategic management and innovation. At the same time, by using logic reasoning, as a well-tested method in science and mathematics, this work integrates different streams of research into an organic whole. On top of that a theoretical foundation is developed so that strategic relatedness in providing simultaneous consumer utilities of the demand side (Sakhartov & Folta, 2014; Ye et al, 2012) and resource-based relatedness in diversification of the supply side (Farjoun, 1998; Kim & Finkelstein, 2009; Villasalero, 2017) can be identified. Additionally, we generally theorize some of the fundamental properties of two-sided markets uncovered previously by Rochet and Tirole (2006), Sun and Tse (2009), Ye et al (2012), and others, either empirically or anecdotally. (For a similar but different approach, see Azevedo and Leshno, 2016).

Another important contribution this work makes is the introduction of systems science (especially, the systemic yoyo model) into the study of synergistic innovations. Although written statements, data analysis, and systemic expressions are essentially different forms of articulations, there are major differences. In particular, written statements, no matter how logically and precisely they are transcribed, are often inconclusive due to their nature of linearity and inability of simultaneously controlling the effect of several arguments. Although methods of data analysis overcome these weaknesses of written statements, the essence of these methods is to extrapolate the known past into the unknown future. That explains why methods of data analysis cannot be successfully employed to predict drastically different future (Lin & OuYang, 2010). In comparison, systems science is fundamentally different from written statements and data analysis. It looks into the internal structure of the organization of concern, where the structure and the organization come into being simultaneously. Hence, conclusions of investigations based on systems methodology are much closer to the real situation under investigation than those produced otherwise.

PREPARATION

To make this paper self-contained, this section introduces the relevant concepts and the yoyo model of systems science, the basic assumptions and terminology.

Relevant Elements of Systems Science and Methodology

To make this presentation self-contained, this subsection briefly introduces the relevant basics and facts of systems science and methodology, in order to help keep scholars in works related to this paper updated with many technical details of systems science.

Systems can be seen everywhere in real life, especially in investigations of issues related to economies and business decision making. For example, each individual is a biological system that is made up of many smaller systems. He/she is a part of many social and economic systems, such as families, neighborhoods, communities, etc., and he/she interacts with various kinds of systems, such as an automobile, a smart device, retail businesses, the company he/she works for, etc. Through the wide-range existence of such individuals, these large-scale social and economic systems constantly interact with one another. Hence, when we study social phenomena and economic problems involving organizations, other than using numbers and variables, which is mostly reflected in the literature of social sciences, there is a real need to employ the concept of systems and relevant methods to study how organizations evolve and how they interact with each other in order to produce insightful understandings and conclusions that can practically lead to tangible conclusions in addressing real-life challenges.

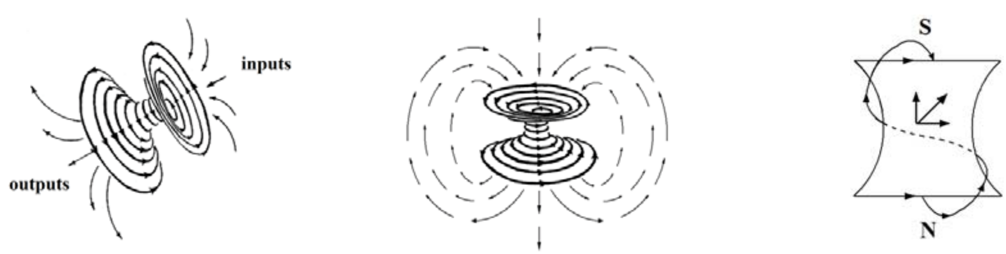
Historically, the concept of systems has been directly or indirectly introduced by scholars from different disciplines. For example, in the area of economics Rostow (1960) wrote that: The classical theory of production is formulated under essentially static assumptions … to merge classical production theory with Keynesian income analysis … introduced the dynamic variables: population, technology, entrepreneurship, etc. But … do so in forms so rigid and general that their models cannot grip the essential phenomena of growth … We require a dynamic theory … which isolates not only the distribution of income between consumption, savings, and investment (and the balance of production between consumers and capital goods) but which focuses directly and in some detail on the composition of investment and on developments within particular sectors of the economy.

In the area of biology von Bertalanffy (1924) points out that because the fundamental character of living things is their organization, the customary investigation of individual parts and processes cannot provide a complete explanation of the phenomenon of life. And others, such as Klir (1985), Lin (2009), Porter (1985), etc., also demonstrate how powerful holistic thinking and relevant methodology could be in producing conclusions regarding organizations and their behaviors that are realistically reliable and practically usable. That is, business entities, economies, and markets, as organizations, do have their individually different internal structures and interact externally with each other. As a matter of fact, since the 1920s (von Bertalanffy, 1924), such a holistic view of nature, organizations, and social events has permeated the entire spectrum of knowledge (Lin, 2009).

Based on what is discussed in the previous paragraphs, it can be seen readily that numbers and systems are concepts respectively abstracted out of the physical world, but from two different points of view. In particular, when an organization is seen as a set of non-interacting elements, numbers come into play, say, *n* employees, *m* buildings, etc. On the other hand, as soon the organization is treated holistically with internal structure systems emerge, where such elements of the organization of concern as employees, capital, properties, talents, etc., form an organic whole through various relationships. Absent of these relationships the organization does not exist. In other words, all business related disciplines are essentially about relationships of organizations (or systems), be they small economic agents, firms (large or small), markets (local, national or international), industries, or economies, and how organization interact with each other.

To summarize, the concepts of numbers and systems are very different in the following two main aspects: 1) the former is a small scale local concept, while the latter a large-scale organizational concept (Lin, 1988; 1999); and 2) numbers exist only post existence, while systems emerge at the same time when physical or intellectual existence comes into being (Lin, 2009).

The first difference tells the reason why systems methodology is a more appropriate tool than the classical theories developed on numbers and variables for the investigation of economic entities when their internal structures cannot be ignored. The second difference explains why business scholars and practitioners still cannot successfully make advanced predictions for imminent economic disasters (Lin & OuYang, 2010).



(a) Eddy motion model of the general system (b) The meridian field of the yoyo model (c) The typical trajectory of how matters return

Figure 1. The systemic yoyo model in 3-Dimensional Euclidean Space

As for systems science, it is simply the totality of studies of various kinds of systems. In the past century, the methods and thinking logic of systems science have been widely employed in scholarly works (Klir, 2001). Similar to how the Cartesian coordinate system – consisting of the crossing of two or more number lines – plays its role in the development of the traditional science and mathematics (Kline, 1972), in systems science the role is played by the systemic yoyo model (Lin, 2007), shown in Figure 1.

Specifically, the blown-up theory (Wu & Lin, 2002), a general theory of development, and discussions on whether or not the world can be seen from the viewpoint of systems (Lin, 1988; Lin, et al, 1990) have led to the development of this yoyo model for each object and every system imaginable. Speaking differently, each system can be theoretically seen as a multi-dimensional entity that spins about its axis. If such a spinning entity is fathomed in the 3-dimensional space within which we live, such a structure as artistically shown in Figure 1(a) appears. The input side pulls in *things*, such as human resources, materials, information, investment, profit, etc. After funneling through the "neck", *things*, such as products and services, are spit out from the output side. Some of the things, spit out as outputs, never return to the input side and some will (Figure 1(b)). Due to its general shape, such a structure is referred to as a yoyo.

What this systemic model says is that each physical or intellectual entity in the universe, be it a tangible or intangible object, a living being, an organization, a market, an economy, etc., can all be treated as a realization of a certain multi-dimensional spinning yoyo with an eddy and meridian field around. This yoyo remains in a spinning motion as depicted in Figure 1(a). When it stops its spinning motion, it will no longer be an identifiable system. What Figure 1(c) tells is that the interaction of the eddy field, which spins perpendicularly to the axis of spin, of the model, and the meridian field, which rotates parallel to axis of spin, makes *things* that are either new to the yoyo body or returning to the input side travel along a spiral trajectory.

Based on this systemic yoyo intuition, each market competition can be seen as a collection of yoyo fields struggling against each other for survival and for growth; each market, large or small, can be modelled abstractly as an ocean of yoyo fields that push against each other and pull similar fields together to form much larger and more powerful fields.

In order to actually develop useful insights and conclusions, let us follow the approach given by Lin (1999), where each system is described as a mathematical structure from which theoretically important and practically useful theorems can be established: S is a (general) system if S is an ordered pair (M, R) of sets, where R is a set of some relations on the set M. Each element in M is called an object of S, and M and R are called the object set and the relation set of S, respectively. A system S is discrete if $R = \emptyset$ or $R = \{\emptyset\}$ and $M \neq \emptyset$; it is trivial if $M = \emptyset$, where $\emptyset$ stands for the empty set.

Basic Assumptions and Terminology

To make logical reasoning in this paper valid, assume that each business firm exists for the purpose of satisfying a particular market niche through generating a positive cash flow, which can be profits from the marketplace, or investments from various investors, or both. In other words, the firm has to actively pursue after certain endeavors in order to meet some market demand(s).

By consumers, we mean the end users of products and/or services. By resource, it means an asset of a firm, be it tangible or not (Harmancioglu et al, 2009), which that firm can use to design and implement its strategies (Barney & Arikan, 2001). In other words, a resource stands for something, be it physical, financial, intellectual, or organizational, the firm can mobilize to accomplish its business goals. And, a resource can generally provide alternative services that differentially increase either the efficiency inside the firm or consumer value creation outside the firm.

By synergistic innovation, it stands generally for such an innovative strategy of a focal firm that systemically deploys and redeploys known elements that exist either within or outside that firm to produce additionally economic benefits. In particular, known elements can be resources - either tangible or intangible, established management routines, production processes, consumer preferences, market demands, etc.; and the word "systemically" means that a focal firm can develop a meaningful system or a set of systems that organically combine the known elements to produce new products with much improved functionalities, and /or new services that meet a forever evolving market demand or consumer preference. The following are some good examples of synergistic innovations in action on either producer-side or consumer-side synergies.

- Different combinations of known technological processes can produce positive synergy (Crossland & Smith, 2002).
- When knowledge across business units of a focal firm is shared, benefits from transferred skills can be anticipated (Hitt et al, 2001).
- When economies of scope and skill transfers are combined, the social network of a focal firm can greatly contribute to that firm's innovativeness (Tsai, 2001).
- Evolving consumer tastes stimulate firms to embark on technological development and innovation (Adner & Zemsky, 2006).
- The benefits consumers enjoy when their knowledge learned from using one product can literally reduce or eliminate the learning time required to use another product (Priem, 2007; Tanriverdi & Lee, 2008).
- The appearance of a disruptive technology can help reveal a new dimension or confirm an existing dimension of consumer preferences for various industries (Adner & Snow, 2010).

SYNERGISTIC INNOVATION AT THE PRODUCER SIDE

This section establishes relevant concepts and elementary properties regarding synergistic innovations at the producer side (Priem & Butler, 2001a, b). The concept of synergistic potential arising from resource relatedness has been introduced since at least the early 1970s (e.g., Rumelt, 1974; Palich et al, 2000) when Rumelt noted that firms, which diversify through employing respectively related resources, generally outclass those firms, which also diversify, but only doing so by utilizing unrelated resources.

Deployment and Redeployment of Resources

By synergistic innovation at the producer side, it means that a focal firm designs and employs an innovative strategy that systemically deploys available resources in such a way that the firm can produce an economy of scope in terms of products and/or services. By producer synergy, it means that through product diversification as a way to widen a firm's stock of assets and skills, the firm is able to expand the perimeter of its value activities (Porter, 1985, p. 380). By synergistic potential arising from resource relatedness, it means the potential of producing new products and/or services, which helps increase the economy of scope of a firm, by innovatively mobilizing related resources. When a mobilization of certain resources leads to positive economic benefits, then we say that the resources create a positive synergy. Otherwise, we say that the mobilization produces a negative synergy.

Proposition 1. If a firm is able to deploy and redeploy resources repeatedly in its effort to produce an economy of scope by making resources, be they tangible or intangible, available to all business units, then there is a potential for synergistic innovations to appear from the effort of diversification.

To see why this proposition holds true, let us model the firm of concern as a system $S = (M, R)$ such that the object set M consists of the resources the firm controls and each relation r in the relation set R stands for a product produced by mobilizing a subset of the resources in M. Now, the given assumption means that freedom and encouragement are given for the employees of the firm to establish possible associations among elements in any subset of M, where each element in M can be repeatedly used in as many associations as possible. That of course implies that a potential of designing and producing many additional products and/or services appear. This logic reasoning is well depicted in Figures 2(a) and 1(b), where the rectangle stands for the object set M, and in Figure 2(a) *A* and *B* are the two existing products of the firm, which are produced respectively by associating resources in their enclosed areas. When employees are encouraged to associate resources any way imaginable, additional enclosed areas, such *C* and *D* in Figure 2(b), can be produced.

(a) Existing associations of resources (b) Additional associations are created

Figure 2. Additional associations of resources are encouraged

This proposition can be shown rigorously by using the following theorem of systems science:

Theorem 1. If $S = (M, R)$ is a multi-level system with a finite object set M such that

- Each object system of S has a finite object set,
- Each object that appears in any relation in the system S or an object system of S only appears once in that relation; and
- S has at least one chain of object systems of more than one level and each object system on the chain contains more than one object,

and $M(S)$ the set of all fundamental objects in S, then $|M| < |M(S)|$.

The proof is very technical and is omitted.

Now, let us model the focal firm with business units strictly defined and resources divided according to the unit divisions as a system $S = (M, R)$, where each object m in M stands for the bundle of resources available only to a particular business unit and each relation r in R as a product or service of the firm. Then the set $M(S)$ of all fundamental objects of the system S

stands for the totality of resources the firm has control over. Now, the conclusion of Theorem 1 says that $|M| < |M(S)|$, which means that the number of available resources is greater than that of business units. Therefore, there are more ways to define relations (products) on *M(S)* than on *M*. That proves Proposition 1.

In terms of the literature, comparing to Proposition 1, a more specific conclusion is given by strategic management scholars. For example, it has been argued (e.g., Porter, 1985) that positive synergies from diversification can emerge if a firm is able to share tangible functions and intangibles, such as knowledge and skills, among business units.

In terms of when a mobilization of certain resources can lead to negative synergies, we have the following result.

Proposition 2. Assume that A and B represent two opposing business goals of a focal firm and that X and Y are mutually exclusive sets of resources that firm has control over, satisfying that resources in X support the realization of goal A and those in Y help actualize goal B. Then a mobilization of any resource in X and any resource in Y will generally produce negative synergies.

To illustrate how two business goals of a focal firm could be opposing to each other, let us consider the following situation (McGrath, 2013):

- A = the goal of implementing the strategy of exploiting the competitive advantages proven successful over time; and
- B = the goal of implementing the strategy of constantly designing and developing new competitive advantages.

Based on this illustration it can be seen that the set of resources employed to materialize goal A and the set of those used for goal B are not the same. All the involved resources can be written into three mutually disjoint sets X, Y, and Z, other than X and Y, as defined in the proposition, set Z includes all those resources that are needed for materializing both business goals A and B, Figure 3. The elements in set Z could include such commonly deployable resources in multiple directions as strong relationships that lead to information sharing, risk taking, and adoption of innovations (Dutta, et al, 1999). Now, it is ready to see that a joint mobilization of the special resources in X and those in Y generally leads to undesirable effects or negative synergies.

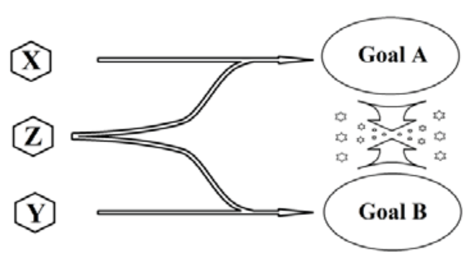


Figure 3. How resources are divided corresponding to business goals A and B

Inconsistent Resources and Core Resources

Two resources x and y are said to be inconsistent, if mobilizing these resources jointly leads to negative synergy. For example, any resource from set X and any resource from set Y in Proposition 2 will be inconsistent, because jointly they lead to undesirable outcome(s). Now, in terms of when a mobilization of certain resources can potentially lead to positive synergies, we have the following conclusion.

Proposition 3. Assume that X represents the set of resources that were used to generate a positive synergy in a focal firm's effort to create an economy of scope. Then by either adding another set of available resource into X or replacing some resources in X by another set of available resources the firm can potentially lead to positive synergies, if none of the new resources under consideration is inconsistent with the other remaining resources in X.

This result holds true, because altering set X is generally the way for the focal firm to produce additional versions of the product initially developed on X with varying degrees of functionality. Systemically, either adding another set of consistent resource into X or replacing some resources in X by another set of consistent resources mean that the new resources will alter the yoyo structure that underlies the original association of the resources in X in terms of spinning strength, the orientation of the spinning axis, the area the spin field covers, etc. The resultant yoyo structure of the updated set X of course means a different product or service, which is either very different from or somehow related to the original one. Depicting this situation, Figure 4 also demonstrates why the condition that the set of new resources used to alter X has to be consistent with the remaining resources in X, otherwise, the resultant yoyo body can be potentially broken into pieces, as what Proposition 2 shows. In Figure 4, the original X evolves into either X' or X", where some of the round dots are replaced by diamonds in the former case, while by starts in the latter case.

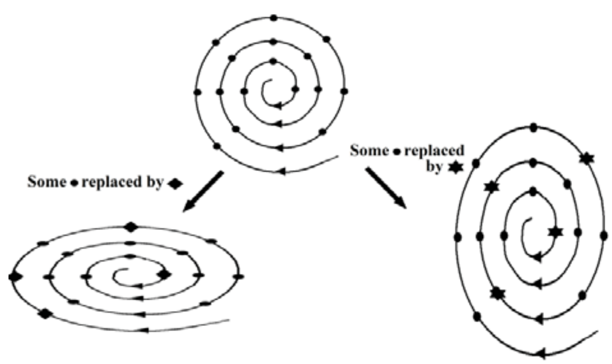


Figure 4. Altering an existing systemic field by changing its composition

A set of resources is said to be relatable, provided that the association of any two resources in the set does not lead to negative synergy, although the association might not produce a positive synergy, either.

Proposition 4. The most natural portfolio of products and/or services of a firm is bound together by a set of related resources that generate positive synergies.

From the point of view of development (Wu & Lin, 2002), any portfolio of products and/or services must start with the very first product or service. When a firm develops its first product or service that generates a positive cash flow (seen as a positive economic benefit), either from the marketplace or investors, there must be a set X of resources the firm has control or influence over that have jointly produced positive synergies for the firm. So, Proposition 3 indicates that a natural way to develop different products and/or services will be finding growth points out of the set X of once-already-proven resources through modifying X by either replacing some resources from X with new and consistent ones or adding additional consistent resources into X. That end is vividly shown historically by how industrial revolutions evolve over time and how one follows another (Forrest et al, 2018); and it is partially recognized by Robins and Weirsema (2003, p. 45) when they remarked that “contemporary view(s) of related diversification might be summarized by … that … a portfolio of businesses is bound together by some shared strategic resources or capabilities”.

As for why such a portfolio of products and/or services is considered most natural, it is because finding an organic association of another set of resources to produce positive economic benefits generally is much harder, which is equivalent to starting up a new business. Other than opportunities of luck, this is much more difficult to accomplish than simply finding growth points on top of the set X as described in the previous paragraph (Salamzadeh & Kawamorita, 2015).

Systemically, this result is quite intuitively straightforward: After a yoyo field emerges, as a new born business entity, in the turbulent ocean (a model of the business world) of countless incumbent yoyo fields that intensively combat against each other first for survival and then for further development, the best and most natural way for it to survive and to grow is to make itself stronger through strengthening what the new born has already achieved and acquiring additional elements so that new and enhanced associations of various resources will help the yoyo field of the new born spin faster and more powerfully than ever before.

Proposition 5. If a firm’s portfolio of products and/or services is bound together by a set of core resources, then there is a good chance for the firm to continuously exploit its established competitive advantages instead of constantly developing new advantages.

This result is a corollary of the previous propositions. In particular, if there is a set of core resources that underlies a firm’s portfolio of products and/or services, the superior value of those resources will be widely acknowledged by the stakeholders of the firm if not yet the case already. So, there will be a momentum for the firm to continue its current state of affairs based on the following first law on state of motion.

First law on state of motion (Lin, 2009): Each imaginable and existing entity in the universe is a spinning yoyo of a certain dimension. Located on the outskirt of the yoyo is a spin field. Without being affected by another yoyo structure, each particle in the said entity’s yoyo structure continues its movement in its orbital state of motion.

The scenario can be seen from a different angle: after putting in tremendous amounts of time and efforts in trying various combinations of resources to get a business operation going, the founders of the firm, as systems themselves with life cycles, can be most likely tired both mentally and physically. That explains why under such circumstances, there is a good chance for the firm to continuously exploit its established competitive advantages instead of constantly developing new advantages, unless new stimuli, either from competitors or from the changing market condition, appears to redirect the attention and energy of the founders and their successors.

SYNERGISTIC INNOVATIONS AT THE DEMAND SIDE

Since over 50 years ago when Penrose (1959) noted that when paying attention to the needs of consumers, companies grow, studies in strategic management have advocated for paying additional attention to the demand side in decision making (e.g., Adner & Zemsky, 2006; Priem, 2007). Riding on this trend of thoughts, this section investigates the concepts and elementary properties of synergistic innovations at the demand side (Ye et al, 2012) by placing end users of products and services in the square center.

When Consumers Are Willing to Pay Additional

Similar to how the concept of producer synergy is defined earlier in the previous section we have the following (Ye et al, 2012): By consumer synergy, it stands for an increased value created for consumers and captured by a firm through employing combination(s) of products and/or services that jointly expand consumer utilities beyond the sum of the consumer utilities offered individually by the products and/or services. For illustration purpose, let us look at an excellent example constructed by Ye at al (2012).

A state university in northern USA has a population of over 29,000 students, most of whom are unmarried undergraduates living on tight budgets. As most of the rental properties around the campus do not offer laundry facilities, self-service, coin-operated laundromats near the university compete for the students' business. Another successful kind of college-town business, attractive mostly to college women, is indoor tanning. Instead of offering these services separately, two entrepreneurs provide their independent, single-location businesses, each of which jointly provide coin-operated self-service laundry facility and indoor tanning service. Such business combination creates consumer synergies from diversification of commonly available assets – same or similar equipment –without much reduction in the overhead costs (from completely different sets of equipment) and labor costs (the combined business types require considerably different production and customer service skills). However, this business combination creates tremendous amount of demand-side synergies due to the following reasons:

(1) Students can start washing their clothes and then tan their skin simultaneously, saving the otherwise unproductive and costly waiting time; and
(2) Students are provided with a place for the potential of establishing relationships because the location of combined services provides opportunities for both males and females to meet, where women's greater participation in the tanning service attracts a larger number of men to the laundry service.

Consequently, the created consumer (or demand-side) synergies help respectively support higher fees collected from the laundry service than other standalone stores and maintain competitive prices for the much increased demand for the tanning service. In other words, a greater value is respectively created for consumers and captured by business providers because of the resultant consumer synergies from product-and-service diversification.

In general, if two or more market demands can be met simultaneously, such as the situation described in reason (1) above, while saving valuable resources for consumers, such as time and/or money, then we say that the diversification effort of meeting different market demands has created a simultaneous consumer utility effect. If two services, each of which attracts a different segment of consumers, although the segments might be overlapping, are collocated so that the increased traffic flow due to the joint offer of services from one segment of consumers attracts an increasing traffic flow from the other segment of consumers, while the latter does not have any adverse effect on the former, as described in reason (2) above, then we say that the diversification of services has created a two-sided market effect.

Proposition 6. When collocating diverse products and/or services leads to a simultaneous consumer utility effect, consumers will be willing to pay additionally for a product or service beyond the market norm.

First, when a simultaneous consumer utility effect is created by collocating diverse products and/or services, by definition it means that consumers have saved some of their valuable resources. That saving generally is convertible into consumers' willingness to pay more than what is expected on the average in the marketplace. When a two-sided market effect is developed by collocating diverse products and/or services, the alternatively reinforcing effects on the traffic flows of the different segments of consumers make the demand for the joint offer of products and/or services rise. That in turn supports rising prices.

Systemically speaking, each simultaneous consumer utility effect can be described by employing the yoyo model in Figure 5. In this systemic depiction, three market demands A, B and C can be satisfied by either three different business entities X, Y and Z separately or by one business firm whose yoyo field encompasses all those of the three entities X, Y and Z. The idea is kind like, but not exactly the same as that of one-stop shopping. The reason why consumers are willing to pay extra is demonstrated by the fact that the conflict zones existing in areas between the yoyo fields X, Y and Z no longer exist, if these small yoyo fields can be combined into a much larger field that encompasses those of X, Y and Z.

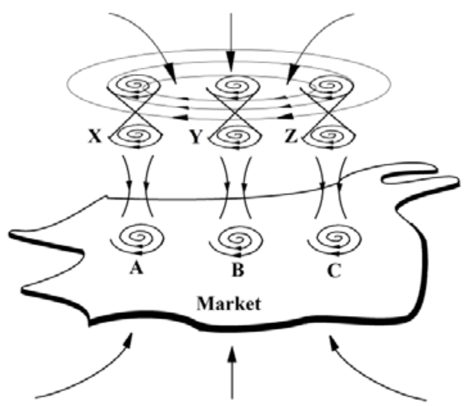


Figure 5. Meeting three market demands simultaneously

Proposition 7. When collocating diverse products and/or services leads to a two-sided market effect, consumers will be willing to pay additionally for a product or service beyond the market norm.

For two-sided market effects, the concept can be intuitively seen systemically in Figure 6. In this systemic intuition two different market segments A and B of consumers are identified, where consumers in B are closely attached to those in A who at the same time are particularly attracted to the service provided by firm X while can also take advantage of the service of Y. From such a realization of potentially creating a two-sided market through diversified offers of products and/or services, the business ideas of both X and Y are combined into one business practice, whose yoyo field encompasses those of both X and Y. This new business firm will attract the consumers in A, through which those consumers in B are also pulled over from the corresponding competitors.

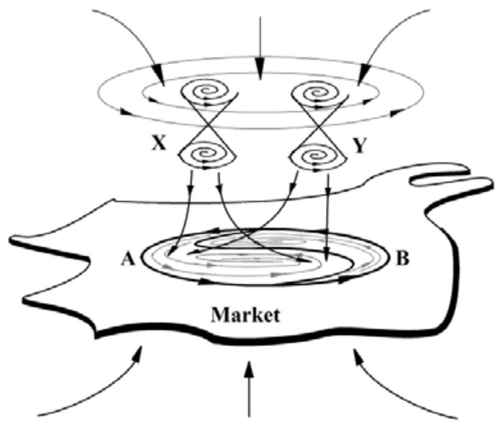


Figure 6. How a two-sided market effect appears

Corporate Growth and Improving Firm Performance

When consumers are willing to pay extra beyond the expected market fees for a product or service due to the reason that diverse products and/or services are provided to consumers as bundles, we say that the diversification has produced a demand-side or consumer synergy. Because at any given moment of time each human being has multiple needs from across different industries, Propositions 6 and 7 imply that demand-side synergy can be obtained via inter-industry diversification of seemingly unrelated products and services from the viewpoint of producers.

Proposition 8. Corporate growth and improving performance can be achieved simultaneously through diversification at either the producer side, or the consumer side, or both.

At the producer side, diversification or creating an economy of scope is most likely achieved in real life by making resources available to all business units (Proposition 1) so that these resources can be deployed and redeployed repeatedly for different purposes, leading to a rich portfolio of products and/or services, which are more or less related due to the resources employed (Proposition 4). From Proposition 1, it follows that diversification at the produce side leads to the potential for synergistic innovations to appear, which helps to grow the business and to improve the performance. This end is like what Porter (1985) argued that firms should consider the ‘fit’ of related diversification for better performance.

From the perspective of demand-side synergies, a synergistic fit of diversified offers can occur on the consumer side (Proposition 6), leading to increasing top lines of a firm. Hence, both corporate growth and improving performance can be simultaneously achieved through diversification.

By combining the previous analyses, it shows that if diversification occurs at both the producer and the consumer side, the chance for corporate growth and improving performance increases tremendously.

CONCLUSIONS

The present world of market economies has been turbulent. To successfully survive and grow in such a constant turbulence, a firm has to continuously look for potentials of different competitive advantages (McGrath, 2013). One way of doing so is to create economies of scope by diversification at either the producer side (Porter 1985; Santalo & Becerra, 2008) or the demand side (Ye et al, 2012) or both. If a firm adopts this approach of diversification, then the firm faces the challenge of how to develop the necessary ideas for diversification. It is around our attempt to address this challenge that this paper establishes a series of both theoretically and practically significant results.

By employing the thinking logic of systems science, this paper is able to develop urgently needed conclusions generally useful for managerial problem solving and decision making (Lockett et al, 2008) and to address some key issues in strategic management. In terms of applications, this paper provides practically significant recommendations for entrepreneurs, managers, and retailers, respectively.

In particular, for managers, entrepreneurs and retailers who focus on the demand side by considering collocating commonly available products and/or services, this paper shows that other than an intimate understanding of consumers' needs and the creativity to envision a novel collocation of commonly available resources (Ye at al, 2012), superior consumer benefits can still be created by offering consumer groups utilities that they were not previously aware. In particular, this paper shows why diversification helps firms with corporate growth and performance (Proposition 8); and when consumers are willing to pay extra for their needs (Propositions 6 and 7). Beyond theoretical results, this paper also provides practically applicable recommendations. For example, when strategically planning for collocations of commonly available products and/or services, a manager/entrepreneur should always keep in mind the potential of complementary consumer benefits (Propositions 6 and 7), where consumer interactions can create another dimension of growth and satisfaction and where a two-sided or even multi-sided market can emerge.

For managers of the producer side who intend to produce an economy of scope, beyond explaining why the idea of economies of scope will help with corporate growth and performance (Proposition 7), this paper also shows how to potentially create synergistic innovations by

- Making their limited resources available to all business units within their companies,
- Discovering what new ideas could be introduced through combining different resources,
- Naturally formulating a portfolio of diversified products and/or services,
- Staying alert about when their firms could likely become exploiters of existing competitive advantages instead of explorers of new advantages in the turbulent consumers' markets, and
- Avoiding negative synergies by not jointly mobilizing inconsistent resources.

To conclude this presentation, let us now look at the limitations of this paper. One limitation stems from the assumption of why a firm actually exists. In fact, firms in the business world behave widely differently from each other depending on why and how they are established. That will explain why results established will not be universally applicable to many actual situations. Although our attention in this paper is on how both conceptually and practically a focal firm can synergistically collocate available resources at the producer side or commonly available products and/or services at the demand side in its effort to create economies of scope, we did not consider the limitations that firm is subject to. These listed and unlisted limitations of this work should help point to future research problems in order to develop additional practically useful insights beyond what have been established in this paper.

REFERENCES

Aas, T.H., Breunig, K.J., Hydle, K.M., & Pedersen, P.E. (2015). Innovation management practices in production-intensive service firms. *International Journal of Innovation Management, 19*(05), DOI: 10.1142/S1363919615500553.

Adner, R., & Levinthal, D. (2001). Demand heterogeneity and technology evolution: Implications for product and process innovation. *Management Science*, *47*(5), 611–628.

Adner, R., & Snow, D. (2010). Old technology responses to new technology threats: Demand heterogeneity and graceful technology retreats. *Industrial Corporate Change, 19*(5), 1655–1675.

Adner, R., & Zemsky, P. (2006). A demand-based perspective on sustainable competitive advantage. *Strategic Management Journal, 27*(3), 215–239.

Azevedo, E.M., & Leshno, J.D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy, 124*(5), 1235-1268.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

Barney, J., & Arikan, A. (2001). The resource-based view: Origins and implications. In M. Hitt, R. Freeman, & J. Harrison (Eds.), *Handbook of Strategic Management* (pp. 124–185). Oxford, UK: Blackwell.

Barney, J.B. (2001). Is the resource-based “view” a useful perspective for strategic management research? Yes! *Academy of Management Review, 26*(1), 41–56.

Becheikh, N., Landry, R., & Amara, N. (2006). Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993 to 2003. *Technovation, 26*(5), 644–664.

Buffington, J. (2016). The future of manufacturing: An end to mass production. In Buffington, J., (2016). *Frictionless Markets* (pp. 49–65). Berlin: Springer.

Caputo, A., Marzi, G., & Pellegrini, M.M. (2016). The internet of things in manufacturing innovation processes: Development and application of a conceptual framework. *Business Process Management Journal, 22*(2), 383–402.

Crossland, P., & Smith, F.I. (2002). Value creation in fine arts: A system dynamics model of converse demand and information cascades. *Strategic Management Journal, 23*(5), 417–434.

Damanpour, F. (1991). Organizational innovation: A meta-analysis of effects of determinants and moderators. *The Academy of Management Journal*, *34*(3), 555–590.

Drucker, P.F. (1954). *The Practice of Management*. New York: Harper & Row.

Dutta, S., Narasimhan, O., & Rajiv, S. (1999). Success in high-technology markets: Is marketing capability critical? *Marketing Science*, *18*(4), 547–568.

Eisenhardt, K., & Martin, J. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, *21*(10/11), 1105–1121.

Fang, E., Palmatier, R., & Grewal, R. (2011). Effects of customer and innovation asset configuration strategies on firm performance. *Journal of Marketing Research*, *48*(3), 587–602.

Farjoun, M. (1998). The independent and joint effects of the skill and physical bases of relatedness in diversification. *Strategic Management Journal, 19*(7), 611–630.

Gans, J.S., MacDonald, G., & Ryall, M.D. (2008). The two sides of competition and their implications for strategy. Working paper, Melbourne Business School, Carlton, VIC, Australia. https://works.bepress.com/michael_ryall/15/.

Gary, M.S. (2005). Implementation strategy and performance outcomes in related diversification. *Strategic Management Journal, 26*(7), 634–664.

Harmancioglu, N., Droge, C., & Calantone, R. (2009). Strategic fit to resources versus NPD execution proficiencies: What are their roles in determining success? *Journal of the Academy of Marketing Science*, *37*(3), 266–282.

Hitt, M.A., Bierman, L., Shimizu, K., & Kochhar, R. (2001). Direct and moderating effects of human capital on strategy and performance in professional service firms: A resource-based perspective. *The Academy of Management Journal, 44*(1), 13–28.

Kanter, R.M. (2001). *Evolve! Succeeding in the Digital Culture of Tomorrow*. Boston: Harvard Business School Press.

Kim, J.Y., & Finkelstein, S. (2009). The effects of strategic and market complementarity on acquisition performance: Evidence from the U.S. commercial banking industry, 1989–2001. *Strategic Management Journal, 30*(6), 617–646.

Kline, M. (1972). *Mathematical Thought from Ancient to Modern Times*. Oxford: Oxford University Press.

Klir, G. (1985). *Architecture of Systems Problem Solving*. New York, NY: Plenum Press.

Klir, G. (2001). *Facets of Systems Science*. New York: Springer.

Lin, Y. (1988). Can the World be Studied in the Viewpoint of Systems?? *Mathl. Comput. Modeling*, *11*, 738–742.

Lin, Y. (1999). *General Systems Theory: A Mathematical Approach*. New York: Kluwer Academic/Plenum Publishers.

Lin, Y. (2007). Systemic yoyo model and applications in Newton's, Kepler's laws, etc. *Kybernetes: The International Journal of Cybernetics, Systems and Management Science*, *36*(3-4), 484–516.

Lin, Y. (2009). *Systemic Yoyos: Some Impacts of the Second Dimension*. New York: CRC Press (an imprint of Taylor and Francis).

Lin, Y., & OuYang, S.C. (2010). *Irregularities and Prediction of Major Disasters*. New York: CRC Press, an imprint of Taylor and Francis.

Lockett, A., O'Shea, R.P., & Wright, M. (2008). The development of the RBV: Reflections from Birger Wernerfelt. *Organization Studies, 29*(8–9), 1125–1141.

Lockett, A., Thompson, S., & Morgenstern, U. (2009). The development of the resource-based view of the firm: A critical appraisal. *International Journal of Management Review, 11*(1) 9–28.

McGrath, R.G. (2013). *The end of competitive advantage: How to keep your strategy moving as fast as your business*. Boston: Harvard Business Review Press.

Nason, R.S., & Wiklund, J. (2018). An assessment of resource-based theorizing on firm growth and suggestions for the future. *Journal of Management, 44*(1), 32-60.

Palich, L.E., Cardinal, L.B., & Miller, C.C. (2000). Curvilinearity in the diversification-performance linkage: An examination of over three decades of research. *Strategic Management Journal, 21*(2), 155–174.

Penrose, E.T. (1959). *The Theory of the Growth of the Firm*. John Wiley & Sons, New York.

Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press.

Priem, R.L. (2007). A consumer perspective on value creation. *The Academy of Management Review, 32*(1), 219–235.

Priem, R.L., & Butler, J. E. (2001a). Is the resource-based "view" a useful perspective for strategic management research? *The Academy of Management Review, 26*(1), 22–40.

Priem, R.L., & Butler, J. E. (2001b). Tautology in the resource-based view and the implications of externally determined resource value: Further comments. *The Academy of Management Review, 26*(1), 57–66.

Robins, J.A., & Wiersema, M.F. (2003). The measurement of corporate portfolio strategy: Analysis of the content validity of related diversification indexes. *Strategic Management Journal, 24*(1), 39–59.

Rochet, J.C., & Tirole J. (2006). Two-sided markets: A progress report. *RAND Journal of Economics, 37*(3), 645–667.

Rostow, W.W. (1960). *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge, UK: Cambridge University Press.

Rumelt, R.P. (1974). *Strategy, Structure, and Economic Performance*. Cambridge, MA: Harvard University Press.

Sakhartov, A. V., & Folta, T. B. (2014). Getting beyond relatedness as a driver of corporate value. *Strategic Management Journal, 36*(13), 1939-1959.

Salamzadeh, A., & Kawamorita, K.H. (2015), Startup companies: Life cycle and challenges. *The 4th International Conference on Employment, Education and Entrepreneurship* (EEE), Belgrade, Serbia, 2015. Available at SSRN: https://ssrn.com/abstract=2628861 or http://dx.doi.org/10.2139/ssrn.2628861

Santalo, J., & M. Becerra. (2008). Competition from specialized firms and the diversification-performance linkage. *Journal of Finance, 63*(2), 851–883.

Sawhney, M., Verona, G., & Prandelli, E. (2005). Collaborating to create: The Internet as a platform for customer engagement in product innovation. *Journal of Interactive Marketing, 19*(4), 4–17.

Shah, S.K., & Tripsas M. (2007). The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal, 1*(1–2), 123–140.

Sirmon, D.G., Gove, S. & Hitt, M.A. (2008). Resource management in dyadic competitive rivalry: The effects of resource bundling and deployment. *The Academy of Management Journal, 51*(5), 919–935.

Smith, A. (1776). *The Wealth of Nations*, books I-III, (1986 printing). London: Penguin Books.

Smith, W.K., & Tushman, M.L. (2005). Managing strategic contradictions: A top management model for man-aging innovation streams. *Organization Science, 16*(5), 522–536.

Sun, M., & Tse, E. (2009). The resource-based view of competitive advantage in two-sided markets. *Journal of Management Studies, 46*(1), 45–64.

Tanriverdi, H., & Lee, C.H. (2008). Within-industry diversification and firm performance in the presence of network externalities: Evidence for the software industry. *The Academy of Management Journal, 51*(2), 381–397.

Tanriverdi, H., & Venkatraman, N. (2005). Knowledge relatedness and the performance of multibusiness firms. *Strategic Management Journal, 26*(2), 97–119.

Tripsas, M. (2008). Customer preference discontinuities: A trigger for radical technological change. *Managerial Decision Economics, 29*(2–3), 79–97.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *The Academy of Management Journal, 44*(5) 996–1004.

Veugelers, R., & Cassiman, B. (1999). Make and buy in innovation strategies: Evidence from Belgian manufacturing firms. *Research Policy, 28*(1), 63–80.

Villasalero, M. (2017). A resource-based analysis of realized knowledge relatedness in diversified firms. Journal of Business Research, 71(Feb.), 114-124.

Visnjic, I., Wiengarten, F., & Neely, A. (2016). Only the brave: Product innovation, service business model innovation, and their impact on performance. *Journal of Product Innovation Management, 33*(1), 36–52.

von Bertalanffy, L. (1924). Einfuhrung in Spengler's werk. *Literaturblatt Kolnische Zeitung*, May.

Wu, Y., & Lin, Y. (2002). *Beyond Nonstructural Quantitative Analysis: Blown-ups, Spinning Currents and Modern Science*. River Edge NJ: World Scientific.

Ye, G.L., Priem, R.L., & Alshwer, A.A. (2012). Achieving demand-side synergy from strategic diversification: How combining mundane assets can leverage consumer utilities. *Organization Science, 23*(1), 207-224.

Zollo, L., Marzi, G., Boccardi, A., & Ciappei, C. (2016). Gli effetti della Stampa 3D sulla competitivita aziendale. Il caso delle imprese orafe del distretto di Arezzo. *Piccola Impresa/Small Business, 2*(2), 80–100.

Michael D. Mattei, Ph.D.
Bellarmine University
Email: mmattei@bellarmine.edu

Stephen J. Hellebusch, Ph.D.
Hellebusch Research and Consulting
steve@hellrc.com

ABSTRACT

In Behavioral Economics, "arbitrary coherence" is when an arbitrary, randomly chosen number, influences the amount purchasers are willing to pay for a product. Arbitrary coherence is similar to anchoring which marketers sometimes use to help set optimal prices. This paper examines how the arbitrary coherence effect influences individual decision making.

KEYWORDS: Arbitrary coherence, Anchoring, Semantic differential, Monadic testing, Random numbers

INTRODUCTION

A landmark article in Science (Tversky and Kahneman, 1974) reported a research track concerning decision-making that essentially created the field of Behavioral Economics. In that article, the authors noted that humans normally use heuristics ("rules of thumb") to interpret information and make decisions. Humans are especially prone to errors concerning representivity (judging instances more or less representative than they are; ignoring sample size, etc.), availability (e.g., bias of imaginability), and adjustment and anchoring (e.g., bias due to conjunctive and disjunctive events).

In the book *Predictably Irrational* by Dan Ariely (2008), a study is described which reveals something very interesting about the effect of context on how we make decisions. Knowledgeable and very intelligent people who were fully aware that certain numbers were chosen in a completely arbitrary, random fashion were still influenced by those numbers when making decisions about what they would pay for certain items. This phenomenon is described as "arbitrary coherence", a term also found in the behavioral economics literature. One of the authors was working for a large national marketing research firm at the time and thought that the results of the study could provide a valuable tool for marketing research professionals especially in the area of setting new product prices.

A first attempt at replicating the Ariely study using four different products confirmed the phenomenon. A second study, with different items, completely failed to replicate the effect. Thus, began a nearly ten-year journey into answering the question, "When and under what conditions does the arbitrary coherence phenomenon manifest itself?" This paper details five studies beginning with the one that replicated Ariely's findings and the second study that did not demonstrate the effect. Three follow-up studies with students and national samples are described in an effort to identify the underlying factors when arbitrary coherence is observed and when it is not.

LITERATURE REVIEW

Beyond the landmark article in *Science* (Tversky and Kahneman, 1974), the book *Thinking, Fast and Slow* (Kahneman, 2011) elaborated on forty years of research in the area by describing two mental “systems.” Kahneman appropriately cautions that thinking is not as simple as the two-system concept suggests, but the theory helps understand and describe the findings. System 1 is the heuristic-using operator that relies on bits of available data and draws conclusions; it is very good at perceiving patterns, even where there are none. System 2 is the thinking component – focus, concentration, work. System 1 only examines and operates on what is present. Kahneman summarizes its method with WYSIATI – an admittedly awkward acronym for “What You See Is All There Is”. System 1 does not consider unseen possibilities. It makes quick, easy decisions, often accurate in everyday life.

System 2 is the analytical function. It can explore possibilities and elements that are not present, but this requires thinking, which is hard work. So, if System 1 makes a judgment and nothing present (WYSIATI) conflicts with it, System 2 will likely “rubber stamp” it, taking the easy way out. Given this, decisions are subject to systematic biases.

The specific instance of bias of interest here is “arbitrary coherence.” In a summary of six experiments, Ariely, Loewenstein, and Prelec (2003) describe an anchoring phenomenon wherein a random, arbitrary number that respondents know is irrelevant to the decision influenced their judgments of amounts willing to pay and reactions to hearing painful sounds. The authors’ own words are the best succinct description of the phenomenon.

> *In six experiments, we show that the initial valuations of familiar products and some simple hedonic experiences are strongly influenced by arbitrary “anchors” (sometimes derived from a person’s social security number). Because subsequent valuations are also coherent with respect to salient differences in perceived quality or quantity of these products and experiences, the entire pattern of valuations can easily create an illusion of order, as if it is being generated by stable underlying preferences. The experiments show that this combination of coherent arbitrariness (1) cannot be interpreted as a rational response to information, (2) does not decrease as a result of experience with a good, (3) is not necessarily reduced by market forces, and (4) is not unique to cash prices. The results imply that demand curves estimated from market data need not reveal true consumer preferences, in any normatively significant sense of the term.*

The authors say this in a less precise but perhaps clearer way in their introduction.

> *In this paper, we show that consumers’ absolute valuation of experience goods is surprisingly arbitrary, even under “full information” conditions. However, we also show that consumers’ relative valuations of different amounts of the good appear orderly, as if supported by demand curves derived from fundamental preferences. Valuations therefore display a combination of arbitrariness and coherence that we refer to as “coherent arbitrariness.”*

Arbitrary coherence is also referred to in *Predictably Irrational* (Ariely, 2008), and is put in the broader context of anchoring in *Priceless* (Poundstone,2010). The latter also points to numerous examples of anchoring and arbitrary coherence in setting market prices and determining price changes. One example is lowering the volume of peanut butter by indenting the bottom of the jar, since consumers will, by and large, fixate on the familiar price, not the volume. Once an anchor price is established, all other pricing in that particular arena can be coherent with that price in some way. Luxury goods can especially benefit from a high anchor price point.

In a literature search aimed specifically at finding work on “arbitrary coherence” only two articles specific to the topic appear. In a pilot project with 41 fourth year medical students, Pouthier (2009) found that having them write the last two digits of their mobile telephone numbers influenced their subsequent responses to a series of questions concerning surgical training. The questions required estimating numbers, and the mean results by student correlated ($r = 0.36$, $p \leq .05$) with the irrelevant telephone number digits. The author warns that steps should be taken to minimize this unexpected potential bias.

A second article with students at Cambridge (Scott and Lizieri, 2011) tested two hypotheses. Paraphrasing, they were (1) an arbitrary anchor will reliably influence judgment of a property’s value, even when incentives are provided for accuracy; (2) subsequent judgments of property values will be influenced by the value immediately prior, so a single subject’s valuation set will be coherent. After significant data manipulation (“eliminating noise”) primarily by eliminating outliers, they more or less confirmed both hypotheses. Their research clearly has implications for pricing in real estate marketing.

As noted, arbitrary coherence is put into a broader context of anchoring in *Priceless* (Poundstone,2010), pointing to numerous examples of anchoring and arbitrary coherence in setting market prices and determining price changes. Another example besides the peanut butter jar is placing a very expensive item on display, so that other items of less expense look reasonable – the $9,000 Gucci purse displayed with many $2,500 to $3,000 purses available for purchase.

Fudenberg, Levine and Maniadis (2010), in “Reexamining Coherent Arbitrariness for the Evaluation of Common Goods and Simple Lotteries” found that the effects of the arbitrary coherence manipulation on the valuation of common market goods produced very weak anchoring effects, and no effects at all on the valuation of binary lotteries.

> *The assumption that people make decisions based on a constant set of preferences, so that choices should not depend on context-specific cues (anchors), is one of the cornerstones of economic theory. We reexamined the effects of an anchoring manipulation on the valuation of common market goods that was introduced in Ariely, Lowenstein and Prelec (2003). We found much weaker anchoring effects. We performed the same manipulation on the evaluation of binary lotteries, and we found no anchoring effects. This suggests limits on the robustness of strong anchoring effects. Hence, the evidence that people have "arbitrary preferences" may not be conclusive, and economic theory may still be valid in many cases of interest.*

The experiments conducted by Fudenberg et. al. took place in the California Social Science Laboratory (CASSEL) at UCLA, in August of 2009, and the subjects were UCLA students. The first of four experiments were paper-and-pencil as was the case in Ariely et. al. and reexamined the effects of random anchors on subjects' valuation for common market goods of interest to students.

One other consistency in the reviewed research and in the initial research by the authors should be noted. Nearly all were convenience samples, primarily college students, but our earliest studies included marketing researchers who were at hand and willing. The authors are not familiar with any research that attempted to show arbitrary coherence in a national, random sample. In this paper, two national samples of 400 each are detailed.

In an interesting precursor of arbitrary coherence, Wilson et. al. (1996) examined anchoring effects to determine conditions under which anchoring appears. The authors supported a number of hypotheses that relate to arbitrary coherence. Each of their five studies required numerical responses. They found numerical basic anchoring effects occur when uninformative numerical anchors influence judgment, even when people are not asked to compare that anchor to the target value.

Based on the finding from their second study, Wilson et. al. eliminated any subject claiming to be knowledgeable from the next three, so their conclusions must be qualified accordingly. Wilson et. al. focused on aspects of the conditions where anchoring impacts decisions. Following Wilson et. al., questions abound about the limitations of both anchoring and the arbitrary coherence effect.

The fourth study explores the differences in perceptions which might account for when the arbitrary coherence effect is observed by using semantic differentials, specifically, do differences in the perceptions of each item tested correlate in any way to the existence or lack of arbitrary coherence? There is a wealth of literature on semantic differentials and the research is beyond the scope of this paper. Briefly, one of the earliest publications was by Charles Osgood (1957). Mindak (1961) applied the semantic differential to marketing problems.

FIVE STUDIES EXAMINING THE IMPACT OF ARBITRARY COHERENCE ON DECISION MAKING

Study #1 (AC1) – The Initial Attempt to Replicate Ariely's "Coherent Arbitrariness"

In the Ariely study (2003), six items with an average retail price of $70 were shown to 55 MBA students during the first marketing research class meeting at the Sloan School. Students were asked if they would buy each item for the dollar figure equal to the last two digits of their Social Security Number. They were then asked if an auction was held, what is the highest price they would be willing to pay (WTP) for each item.

The goal of the first study we conducted was to see if the arbitrary coherence (AC) effect would be observed in a less "naïve" group of respondents. Four items were selected at random from advertisements in a Sunday paper, specifically: Teleflora Spring Pitcher Floral Display, Dr. Scholl's Hand-Stitched Leather Loafers, Irish Spring 12-bar Value Pack and a bottle of Bailey's Irish Crème (Appendix 1). There are a few differences from Ariely's original research, specifically, this study was conducted with 69 marketing research professionals and four items were tested using photos only, not the actual items. To further test the effect, two items with generally unknown values were tested with two items with generally known values below $30. The results from Ariely are shown in Table 1 as the reference for our first study (AC1).

The survey contained the following three questions (Appendix 2):

1. Enter the last two digits of your Social Security number (SS#) as if it were a price in dollars.
2. Indicate if you would be willing to pay that amount for the item by writing in a Y for “yes” or an N for “no”.
3. Pretend there is an auction for the item. In whole dollars, write in the maximum amount you would be willing to pay for it (WTP).

Table 1. Results from the Ariely (2003)

Products	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Correlation
Cordless Trackball	$ 8.64	$11.82	$13.45	$21.18	$26.18	.415 p=.0015
Cordless Keyboard	$16.09	$26.82	$29.27	$34.55	$55.64	.516 p<.0001
Average Wine	$8.64	$14.45	$12.55	$15.45	$27.91	.328 p=.014
Rare Wine	$11.73	$22.45	$18.09	$24.55	$37.55	.328 p=.0153
Design Book	$12.82	$16.18	$15.82	$19.27	$30.00	.319 p=.0172
Belgian Chocolates	$9.55	$10.64	$12.45	$13.27	$20.64	.419 p=.0013

As Ariely points out,

> … the impact of the social security number on stated WTP was significant in every product category… the top quintile subjects were willing to pay $56 on average for the cordless computer keyboard, compared with $16 on average for the subjects with the bottom quartile numbers. Alongside this volatility of absolute preference, we also observed a marked stability of relative preference. For example, the vast majority of subjects (>95 percent) valued a cordless keyboard more than a trackball, and the highly rated wine more than the lower-rated wine.

The results of AC1 are shown in Table 2. The arbitrary coherence effect was found to exist for the two higher-priced items where the actual price was uncertain. Even though most respondents would have a good idea of the actual retail price for the Bailey’s Irish Crème there was a positive correlation. As expected, there was no correlation for the Irish Spring, the one item most subjects knew was inexpensive.

Table 2. Results of Arbitrary Coherence (AC1) Study

	Last 2 Digits Social Security Number as $				
Products	Under $25	$26 - $50	$51 - $75	$76+	Correlation
Teleflora	$22.13	$43.29	$38.70	$40.79	+.33*

Dr. Scholl's	$34.13	$34.07	$49.45	$51.53	+.28*
Bailey's Irish Crème	$15.50	$16.93	$16.75	$20.42	+.16
Irish Spring 12 Pack	$5.63	$9.50	$6.10	$6.21	-.05

* = the correlation is statistically significantly different than zero, $p \leq 0.05$.

If a group of professionals were "impacted" by the random number anchoring, the next question was, "Is there something in the questions that produces the results?" A second study was conducted about six months later with the same, but larger group of 84 professionals.

Study #2 (AC2) – Different Item Set, Two Questions versus Three

For the second study, four different items, all of reasonable and unknown value, were used: Crocodile Wallet, Copper Bowl, Necklace of Pearls from Tahiti, The Complete Works of Lewis Carroll (Appendix 3). These were again selected from newspaper ads, this time to have prices that would likely be unfamiliar to subjects, since uncertainty is a precondition.

In this study (AC2) half the group were not exposed to the middle question "indicate if you would be willing to pay that amount for the item" (Appendix 4). The other half of the respondents completed all three of the original questions. Subjects were randomly assigned, so that half were asked the original three questions, and half were not asked the middle question (would they buy it for the specified amount).

Table 3. Results of AC 2 Study with Purchase Decision Question (Three Questions)

	Last 2 Digits Social Security Number as $				
Products	Under $40	$40 to $59	$60 to $79	$80 to $99	Correlation
Crocodile Wallet	$38.22	$42.50	$18.60	$22.69	-.24
Copper Bowl	$90.56	$112.90	$158.40	$151.54	+.10
Necklace from Tahiti	$26.33	$45.80	$25.50	$47.85	+.12
Complete L. Carroll	$65.56	$34.20	$25.50	$75.38	-.03

Table 4. Results of AC2 with No Purchase Decision Question (Two Questions)

	Last 2 Digits Social Security Number as $				
Products	Under $30	$30 - $59	$60 - $79	$80 - $99	Correlation
Crocodile Wallet	$13.10	$20.22	$23.17	$20.64	+.08
Copper Bowl	$161.50	$189.89	$117.58	$135.45	-.05
Necklace from Tahiti	$14.10	$36.00	$36.58	$15.91	+.04
Complete L. Carroll	$137.10	$97.44	$95.00	$64.36	-.18

In the second study, there was no arbitrary coherence effect in either subgroup. This lack of effect meant that it was impossible to shed light on the effect of the second question. It also raised new questions about the limitations of arbitrary coherence.

Did the lack of effect in the test cell that received exactly the same questions occur because the same subjects were used? But prior research, reported in *Priceless* (Poundstone, 2010), found that experience did not impact arbitrary coherence. In this case, though, one wonders if these informed researchers were “conditioned?” Alternatively, did the lack occur because of the items used? Were the items of little interest, limiting the effect? Or, was it just the random assignment of subjects presenting an effect of its own?

Study #3 (AC3) – Two Items Sets and Two versus Three Questions

The third study was conducted among 84 university students to attempt to resolve the questions from AC2; specifically:

1. Is the arbitrary coherence effect present when each set of items is tested with naïve subjects?
2. Is making a purchase decision (“the middle question”) causing a different effect?
3. Is there an interaction, such that making or not making a purchase decision yields different results for each set of items?

The age range for respondents was 18 to 47 years old and all had Social Security numbers. The mean age is 22 and the median is 21 with 71% between 18 and 22 years old. Thirty-two (38%) are female. There are 53 undergraduates (63%) and 31 (37%) graduate students.

Since there are two types of questionnaires, one with two questions and the other with three questions and two sets of items there were four test groups (Table 5). Four color coded surveys were created and distributed to four independent groups with between 19 and 22 different students per group. Group 3 (16) and Group 4 (14) contained almost all of the graduate students.

Table 5. Table of Groupings for the Third AC Study (AC3)

	Item Set #1:	Item Set #2:	Totals
Purchase Decision (3Qs)	Group 1: 22 students AC3, Set1, 3Qs	Group 3: 19 students AC3, Set2, 3Qs	41
No Purchase Decision (2Qs)	Group 2: 21 students AC3, Set1, 2Qs	Group 4: 22 students AC3, Set2, 2Qs	43
Totals	43	41	

Item Set #1 consists of three of the four original items from the first study. In deference to the fact that some subjects were underage for alcohol consumption, the original Bailey’s Irish Crème was changed to “Sandy Flats Pure Maple Syrup – 250ml” (Appendix 1). Item Set #2 is the same four items used in the second study (Appendix 3). To conduct the survey a one-page paper questionnaire, color-coded by cell, was handed out for subjects to complete. A picture of each item was projected on a screen all at one time. Half the students in the room received the two-question survey and the other half received the three-question survey.

For each item, the amount subjects were Willing to Pay (WTP) was examined in two different ways: by quartile grouping and by looking at the correlation between the SS$ and the WTP. If arbitrary coherence holds, the WTP should increase with increasing SS$.

Group 1 – Item Set 1 and 3Qs plus Comparison to AC1

Group 1 shows a clear arbitrary coherence effect. The pattern parallels the original AC1 study. The Irish Spring 12-pack of bar soap shows stronger tendency to arbitrary coherence than it did in the first study, but the correlation is still not statistically significant. Teleflora, which had the strongest correlation with the older researchers is somewhat weaker with these younger students.

Table 6. Group 1 - Item Set 1 with Purchase Decision Question plus Comparison to AC2

	Last 2 Digits Social Security Number as $					
	Under $25	$26 - $59	$60 - $80	$80+	AC3-3Qs Correl.	AC1-3Qs Correl.
Teleflora	$15.83	$25.50	$22.00	$25.00	+.28	+.33*
Dr. Scholl's Loafers	$39.83	$43.83	$67.00	$56.80	+.36**	+.28*
Maple Syrup-250ml	$5.17	$5.50	$10.00	$23.60	+.39**	
Irish Spring 12 pack	$8.00	$10.00	$11.80	$10.80	+.24	-.05
Bailey's Irish Cream						+.16

* = the correlation is statistically significantly different than zero, $p \leq 0.05$.
** = the correlation is statistically significantly different than zero, $p \leq 0.10$.

One question this research addressed was whether results would be replicated with new subjects. They were. The respondents in the AC1 study were experienced marketing researchers, which may explain why the results with the more naïve students are overall stronger. The researchers not only were older, but also had more experience with questionnaire studies like this, which may have made them slightly less susceptible to the AC effect.

Group 2 - Item Set 1 with No Purchase Decision Question (No Prior Study for Comparison)

While there was a clear effect with all four products when questioning with the three questions on Item Set 1 (Table 6), the arbitrary coherence effect vanishes in all but one item when the Purchase Decision Question is removed (Table 7). Teleflora has a statistically significant correlation while the effect is almost nonexistent for the other three items.

Table 7. Group 2, Item Set 1 with No Purchase Decision Question
(no comparable prior study)

	Last 2 Digits Social Security Number as $ (SS$)				
	Under $18	$19 to $41	$42 to $79	$80 to $99	Correlation
Teleflora	$6.67	$23.00	$8.60	$26.40	+.41*
Dr. Scholl's loafers	$31.67	$45.00	$39.40	$40.80	+.07

Maple Syrup-250ml	$6.67	$9.20	$3.80	$7.00	+.02
Irish Spring 12-pack	$5.83	$9.00	$5.60	$6.20	-.04

* = the correlation is statistically significantly different than zero, $p \leq 0.05$.

Group 3 – Item Set 2 with 3Qs plus Comparison to AC2

The student survey results for the second set of products and three questions are similar to the AC2 results; there was no arbitrary coherence effect. All the correlations are weak and three of four are negative. There is little consistency in WTP across the quartiles.

Table 8. Group 3, Item Set 2 with Purchase Decision Question plus Comparison to AC2

	Last 2 Digits Social Security Number as $					
Products	Under $23	$23 to $40	$41 to 70	$71 to $99	AC3-3Qs Correl.	AC2-3Qs Correl.
Crocodile Wallet	$48.00	$20.83	$19.50	$50.00	+0.13	-0.24
Copper Bowl	$60.60	$25.83	$24.17	$34.00	-0.14	+0.10
Necklace from Tahiti	$80.00	$249.17	$53.67	$86.00	-0.12	+0.12
Complete L. Carroll	$18.20	$126.67	$27.50	$46.00	-0.04	-0.03

Group 4 – Item Set 2 and 2Qs plus Comparison to AC1

Given that there was no effect with these items when a SS$ purchase decision is required, it is not surprising that there is none when there is no purchase decision. With correlations hovering near zero, the pattern is consistent with that found with Group 2, Item Set 1 with No Purchase Decision.

Table 9. Group 4, Item Set 2 with No Purchase Decision Question plus Comparison to AC2

	Last 2 Digits Social Security Number as $					
Products	$00-$25	$26-$60	$61-$70	$71-$99	Correl.	AC2-2Qs Correl.
Crocodile Wallet	$40.00	$26.00	$50.00	$33.00	+0.05	+0.08
Copper Bowl	$27.00	$20.00	$16.25	$24.00	-0.05	-0.05
Necklace from Tahiti	$128.00	$71.00	$128.75	$153.00	+0.04	+0.04
Complete L. Carroll	$22.00	$15.00	$22.50	$51.00	+0.17	-0.18

Observations and Summary of Results for AC3

In the Purchase Decision groups, an indicator of the willingness to buy is how many said “yes” to the second question, “In column B, indicate if you would be willing to pay that amount for the item.” How do the proportions saying “yes” compare between this study and the earlier studies?

In Item Set 1, the overall interest in either Bailey’s Irish Crème or Maple Syrup is low, as is interest in the Irish Spring soap. The interest among the AC3 students in Teleflora is not much higher than for the Irish Spring. About the only consistency is that the Dr. Scholl's loafers received more interest among the students than the researchers, and the correlation with SS$ is stronger with students than with researchers.

Table 10. Item Set 1 - Percent Who Would Buy Product for SS$

Items	AC2 researchers	AC3 students
Bases	69	22
Teleflora	43.5%	18.2%
Dr. Scholl	49.3%	63.6%
Bailey's/Maple Syrup	18.8%	9.1%
Irish Spring	4.3%	13.6%

Interestingly, the percentages who would buy for the SS$ prices are very similar for Item Set 2 (Table 11). In neither study with Item Set 2 was there an arbitrary coherence effect.

Table 11. Item Set 2 - Percent Who Would Buy Product for SS$

Items	AC2 researchers	AC3 students
Bases	42	22
Wallet	21.4%	22.7%
Bowl	38.1%	45.5%
Pearl Necklace	85.7%	86.4%
Carroll Book	38.1%	31.8%

Overall, there does not appear to be a strong relationship between interest and the presence of arbitrary coherence.

Research Questions Answered:

1. Is the Arbitrary Coherence effect present when each set of items is tested with naïve subjects?

    No. It is present in Item Set 1 in three of the items with the researchers and, to some extent, with all four among the students. But it is not present anywhere with the products in Item Set 2, even though the actual cost of these is likely uncertain and, as nearly as can be assessed, levels of interest are similar.

2. Does making a purchase decision causes a different effect?

    The similarities in the AC2 study and groups 2 and 4 in the AC3 study (where no purchase decision was made) certainly suggest that eliminating the purchase decision eliminates arbitrary coherence.

3. Is there an interaction, such that the making or not making of a purchase decision yields different results for each set of items?

    This seems to be the case. Making a purchase decision is in some way responsible for creating arbitrary coherence in Item Set 1, as the comparison between Group 1 (Item Set 1 with Purchase Decision) and Group 2 (Item Set 1 with No Purchase Decision) demonstrates. But the situation is more complex, since Groups 3 and 4 show the same “no arbitrary coherence” results regardless of the inclusion of the purchase decision question.

One of the original intentions of AC3 was to see if the AC effect existed with Group 3 (Set 2, with purchase decision question), which would imply that the prior failure to replicate was because knowledgeable subjects were used in that research. The AC2 group’s greater experience was obviously not the reason for the lack of effect. It also appears unlikely that lack of interest in Item Set 2 products is responsible. If it is not simply greater variability in interest, what it is about Item Set 2 that obliterates arbitrary coherence?

Study #4 (AC4) – National Sample, Two Items Sets and Two versus Three Questions

This study attempts to identify the underlying factors that explain why arbitrary coherence is found in some of the previously mentioned situations but not in others. Factors tested in this study were, would a national sample, rather than a local convenience sample, make a difference? Also, what would happen if each item was presented alone, not in any set? Could respondents’ perception of the test items help explain when arbitrary coherence is observed? A semantic differential measurement was included in the survey to quantify factors that might identify the casual factors that produce the inconsistency of arbitrary coherence reported in the literature.

This study was conducted with a nationally representative sample of 402 adults with the objective of determining if a national sample provides more consistent evidence of arbitrary coherence than found in the literature. The eight items from the prior studies were used for this study.

The arbitrary coherence part of the questionnaire for this research was like that used in prior research. However, experience had determined that asking consumers in an online panel anything about their Social Security number reduced response rates. To compensate for this, the first question, which established the arbitrary anchor, was changed to write the last two digits of your primary telephone number as if it were a price in dollars. The first four items were in the original Item Set 1 and the last 4 items were in Item Set 2 as shown in Table 12.

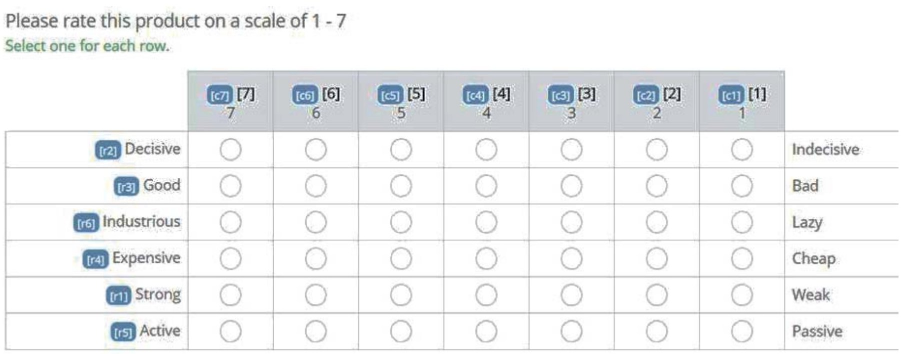
Table 12. Base Sizes of the Arbitrary Coherence Test Cells

The Eight Item Cells	Base -- Total per Cell
Teleflora	50
Dr. Scholl’s Loafers	51
Maple Syrup-250ml	51
Irish Spring 12-pack	50
Crocodile Wallet	50
Copper Bowl	50
Necklace from Tahiti	50

Complete Works of L. Carroll	50

The Semantic Differential attributes were added after the arbitrary coherence questions for each item. They were taken from a University of California Davis teaching website (Unknown 2017). The respondent saw the semantic differential scale shown in Figure 1. After examining the connotative meaning of thousands of concepts, Charles Osgood and his associates identified three major dimensions of meaning: *strength, value*, and *activity*. The first two examples below fit the theme of *strength.* The second two represent *value*, and the last two illustrate *activity*.

Figure 1. Semantic Differential Attribute Questions


Please rate this product on a scale of 1 - 7
Select one for each row.

	[c7] [7] 7	[c6] [6] 6	[c5] [5] 5	[c4] [4] 4	[c3] [3] 3	[c2] [2] 2	[c1] [1] 1	
[r2] Decisive	○	○	○	○	○	○	○	Indecisive
[r3] Good	○	○	○	○	○	○	○	Bad
[r6] Industrious	○	○	○	○	○	○	○	Lazy
[r4] Expensive	○	○	○	○	○	○	○	Cheap
[r1] Strong	○	○	○	○	○	○	○	Weak
[r5] Active	○	○	○	○	○	○	○	Passive

Discussion of the Arbitrary Coherence Effect

When these eight items are tested individually online with a sample national in scope, arbitrary coherence is greatly lessened. For only two of the eight items is there a clear statistically significant correlation ($p \leq 0.05$) between the last two-digit telephone dollars ($Phone#) and the amount respondents were willing to pay for the item, namely the Teleflora Flowers and the Crocodile Wallet. Note that arbitrary coherence was not seen with the Wallet in the second and third studies. The Pearl Necklace from Tahiti shows a directionally large correlation, but the correlation is weak for four of the other five items: Dr. Scholl’s Loafers, Sandy Flats Maple Syrup, Irish Spring Soap, and the Copper Bowl. It is modest for the Lewis Carroll Book.

For five of the items, the highest $Phone# group has a WTP that is greater than the lowest – Teleflora, Soap, Wallet, Necklace, and Book. For the other three (Scholl’s, Syrup, Bowl) the highest $Phone# subgroups will pay less than the lowest $Phone# would be willing to pay.

Table 13. Arbitrary Coherence: WTP by Quartile, Correlations

	$Phone#				
	$0 to $25	$26 to $50	$51 to $75	$76 to $99	$Phone# WTP Correlation

Teleflora	$11.16	$22.47	$27.75	$39.12	+0.36*
Scholl's	$41.67	$50.33	$44.58	$38.38	-0.01
Syrup	$20.82	$14.86	$11.42	$19.07	-0.11
Soap	$6.33	$9.83	$9.84	$7.69	+0.03
Wallet	$8.00	$26.19	$21.53	$40.00	+0.29*
Bowl	$18.00	$17.21	$27.22	$16.73	-0.01
Necklace	$31.00	$59.09	$56.83	$78.80	+0.26**
Carroll Bk	$9.43	$10.63	$32.67	$16.15	+0.19

* = the correlation is statistically significantly different than zero, $p \leq 0.05$.
** = the correlation is statistically significantly different than zero, $p \leq 0.10$.

A summary of the correlations from the four arbitrary coherence studies is shown below. Up to this point, items were always tested in sets. This study (AC4) is the only study in which the context was the single item. In all, there are 24 instances where arbitrary coherence might manifest; it is clearly evident in four of the 24, and possibly present in a total of seven of 24 instances.

Table 14. Summary of Correlations from Four Arbitrary Coherence Studies

	AC1	AC2	AC3	AC4
Item Set 1				
Teleflora	0.33*	na	0.28	0.36*
Scholl's	0.28*	na	0.36**	-0.01
Bailey's/Syrup	0.16	na	0.39**	-0.11
Soap	0.05	na	0.24	0.03
Item Set 2				
Wallet	na	-0.24	0.13	0.29*
Bowl	na	0.1	-0.14	-0.01
Necklace	na	0.12	-0.12	0.26**
Carroll Bk	na	-0.03	-0.04	0.19

na = Not Available.
* = the correlation is statistically significantly different than zero, $p \leq 0.05$.
** = the correlation is statistically significantly different than zero, $p \leq 0.10$.

Since arbitrary coherence is a sub-category of the well-known anchoring effect, data were examined across all four studies to see if the highest quartile gave a higher WTP than the lowest. This did occur in 19 of 24 instances. Three that did not are in the current study (AC4). One caution is that base sizes per quartile are small. In many of the earlier studies (and in a few instances here, in AC4) they are single-digits.

Even with one of the three items (Scholl's, Syrup, Bowl) that show a negative shift in WTP from the highest to the lowest quartile in AC4, two show evidence of some anchoring effect. With Scholl's, the lowest shift, from Quartile 1: $0-$25 to Quartile 2: $26-$50, shows an anchoring effect, going from $41.67 to $50.33, but the higher levels cancel any coherence effect. There is no anchoring effect at all, even at lowest-to-second lowest quartile for Syrup or the Bowl.

In the third study (AC3), there was directional evidence of arbitrary coherence with two of the eight items. Observationally, there is at least a modest anchoring effect evident in all four Set 1 items. In Set 2, with the Wallet and the Necklace, anchoring is again very modest, but at least possibly present. With the Bowl, there is no anchoring and no arbitrary coherence in AC3.

Only Set 2 items were used in the second study and no arbitrary coherence effect was found anywhere. There does appear to be some anchoring effect; even with the Wallet the WTP goes from $38.22 to $42.50 in going from the Quartile 1 to Quartile 2.

Table 15. Difference in WTP, Quartile 4 (Highest) Minus Quartile 1 (Lowest)

	AC1	AC2	AC3	AC4
Item Set 1				
Teleflora	$18.66*	na	$9.17	$27.96*
Scholl's	$17.40*	na	$16.97**	-$3.29
Bailey's/Syrup	$4.92	na	$18.43**	-$1.75
Soap	$0.58	na	$2.80	$1.36
Item Set 2				
Wallet	na	-$15.53	$2.00	$32.00*
Bowl	na	$60.98	-26.60	-$1.27
Necklace	na	$21.52	$6.00	$47.80**
Carroll Bk	na	$9.82	$27.80	$6.72

na = Not Available.
* = the correlation is statistically significantly different than zero, $p \leq 0.05$.
** = the correlation is statistically significantly different than zero, $p \leq 0.10$.

Semantic Differential Measures

As mentioned, there are three Semantic Differential dimensions: Strong-Weak, Good-Bad, Active-Passive. For this exercise, those adjectives were used plus three other pairs that echoed them; respectively; they are Decisive-Indecisive; Expensive-Cheap; Industrious-Lazy. The eight items were ranked on the mean ratings shown in Table 17. Essentially, the items for which arbitrary coherence appears all tend to be lower in rank on each of these attribute pairs, so they are viewed, relatively, as weak/indecisive, bad/cheap, and passive/lazy. That generalization is not perfect, since the Teleflora flowers are third on Decisive and fourth on Good, while the Wallet is rated Expensive and the Necklace is Active. Perhaps though, being relatively weak/indecisive, bad/cheap, and passive/lazy assists in generating the effect.

Table 16. Ranking of Items on Mean Seven-point Scale Values

Rank	Strong-Weak	Decisive-Indecisive	Good-Bad	Expensive-Cheap	Active-Passive	Industrious-Lazy
1	Soap	Bowl	Soap	Syrup	Soap	Soap
2	Bowl	Soap	Bowl	Bowl	**[Necklace]**	Bowl
3	Dr. Scholl's	**[Teleflora]**	L.Carroll	**[Wallet]**	Dr. Scholl's	Syrup
4	Syrup	Dr. Scholl's	**[Teleflora]**	**[Necklace]**	Bowl	Dr. Scholl's
5	**[Necklace]**	L.Carroll	Dr. Scholl's	Dr. Scholl's	L.Carroll	**[Wallet]**
6	L.Carroll	**[Necklace]**	**[Necklace]**	L.Carroll	Syrup	**[Necklace]**
7	**[Teleflora]**	Syrup	Syrup	**[Teleflora]**	**[Wallet]**	L.Carroll
8	**[Wallet]**	**[Wallet]**	**[Wallet]**	Soap	**[Teleflora]**	**[Teleflora]**

Note: items in bold and bracket indicate the Arbitrary Coherence effect is present

Table 17. Mean Values for Each Item on Seven-Point Scale

	Strong(7) Weak (1)	Decisive(7) Indecisive(1)	Good(7) Bad (1)	Expensive(7) Cheap (1)	Active(7) Passive(1)	Industrious(7) Lazy (1)
Teleflora	4.72	4.98	5.44	4.44	4.24	4.42
Dr. Scholl's	5.20	4.84	5.33	4.61	4.59	4.82
Syrup	5.10	4.67	4.92	5.29	4.53	4.88
Soap	5.72	5.08	5.80	4.22	5.34	5.42
Wallet	4.64	4.44	4.56	4.76	4.26	4.78
Bowl	5.64	5.10	5.54	5.26	4.58	5.34
Necklace	4.92	4.72	5.24	4.66	4.66	4.76
L.Carroll	4.92	4.80	5.52	4.46	4.56	4.60

AC 4 Research Questions Answered:

Is the arbitrary coherence phenomenon is found in a national sample of adults questioned about each of the eight items used in prior research?

> The best answer is “no.” Arbitrary coherence is clearly present in only two of eight items, weakly in a third. But over the quartiles, the average price willing to pay appears to randomly shift for all but the Teleflora, which rises consistently.

What happens if the eight items in the Sets are broken up and each article is tested individually?

> Apparently, the context of the sets matters in some nonspecific way. The reasonably strong effects seen with Set 1 items in prior research were lessened when the items were shown alone, and two of the items in Set 2, Necklace and the Wallet, were strengthened. Arbitrary coherence was not seen in Set 2 in two prior studies.

Does a semantic differential measurement of how each item is perceived shed light on the differences between items for which arbitrary coherence reliably appears and those for which it does not?

> The items for which arbitrary coherence appears generally tend to be lower in rank on each of these attributes, so they are viewed, relatively, as weak/indecisive, bad/cheap, and passive/lazy. The only item for which arbitrary coherence reliably appears is Teleflora, though not statistically significant in AC3, likely due to small base sizes. It is “middling” on the Semantic Differential and relatively weak, cheap, passive, and lazy. Other items were as well, so no conclusion about items for which arbitrary coherence reliably appears are possible.

Study #5 (AC5) – Arbitrary Coherence versus Monadic Testing

In addition to measuring the arbitrary coherence effect on an item from the original Ariely study, the objective of this final study was to examine the extent a price-demand curve based on AC results parallels one derived from more tradition monadic testing. Marketing researchers sometimes use monadic testing to determine price sensitivity and create a demand curve. If AC creates a similar demand curve, it might be a useful tool to more efficiently (i.e., fewer subjects, less expensive) create that curve.

This study is comprised of two groups which are compared to address the objectives. Group 1 completes a standard three question AC task, followed by a monadic purchase intent (PI) question including price, with the group split by four price points. Group 2 is the reverse, specifically the purchase intent question is first followed by the standard 3 question arbitrary coherence task. For this study, an item was needed with an ambiguous price and could easily be used by anyone, men or by women, young or old, etc. The Belgian chocolates used in the original Ariely article was selected.

The sample consists of 401 subjects (not the same sample as AC4) from across the United States. All were prescreened as having enjoyed chocolate in some form “recently.” Statistical testing shows that the two groups do not differ in terms of age, sex, race, education, or geographic distributions ($p > .05$). The highest percentages are older, between 50 and 69. Most, about seven in 10, are female. The majority, more than eight in 10, are White. Overall, 5% indicated they were Hispanic. The highest percent, just over a third, had some college, with about a fourth having a college degree. The respondents were fairly evenly split between Northeast, Midwest, South and West.

A Word version of the complete online questionnaire can be found in Appendix 5 and 6. After determining demographics, a question about whether or not the subject has enjoyed chocolate recently was asked in a blinded fashion to make certain only those who possibly would be interested in the Belgian Chocolates were questioned.

In Group 1, two hundred (200) were asked about a box of 24 Fine Belgian Chocolates, using the three arbitrary coherence questions with the last two digits of the telephone number used to determine the arbitrary anchor. Then, each respondent was randomly assigned to one of four price subgroups ($12.50, $37.50, $62.50, $87.50) and asked purchase intent (PI) for the same item. Price points were selected by taking the midpoints of the quartiles between zero and 100, after ensuring that some Fine Belgian Chocolates actually do sell for those amounts.

In Group 2, two hundred one (201) were randomly assigned to the four price points first. Each saw the same picture and description of the Belgian Chocolates used for Group 1, but were asked purchase intent (PI) using the traditional scale (Definitely Would Buy, Probably Would Buy, Might or Might Not Buy, Probably Would Not Buy, Definitely Would Not Buy). They were then asked the same arbitrary coherence questions as asked in Group 1.

Summary of Arbitrary Coherence Results for AC5

Given that the arbitrary coherence effect is often not found, it seemed reasonable to first learn if it could be found in either of these groups.

The effect was found in both groups. In Group 1 (AC task first) the correlation between the arbitrary number derived from the telephone number and the Willingness to Pay (WTP) amount is 0.31 (statistically significantly above zero) versus Ariely's .491. There is only one WTP reversal as one goes up the arbitrary telephone dollar quintiles, and that is at the highest level.

Table 27. Groups 1 Belgian Chocolates, Results When AC Task is First

	$Phone# for AC5 and $SSN for Ariely					
Products	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Correlation
AC5 Belgian Chocolates Average WTP (n=200)	$12.82	$18.15	$21.41	$30.08	$28.03	.310 $p<.05$
Ariely Belgian Chocolates Average WTP (n=55)	$9.55	$10.64	$12.45	$13.27	$20.64	.419 $p=.0013$

Similarly, the effect is found in Group 2, when the AC questions are asked after presenting one of the four prices and asking purchase intent. The effect is lessened, however, with a correlation of .20 and two reversals at the high end of the arbitrary price quintiles.

Table 28. Groups 2 Belgian Chocolates, Results When AC Task is Second

	Last Two Digits or Telephone Number Corresponding to the Monadic Prices					
	0 to $12	$13 to $37	$38 to $62	$63 to $87	$88+	Correl.
Average WTP	$14.13	$21.92	$33.92	$29.64	$29.44	0.20*

* = the correlation was statistically significantly different than zero, $p \leq 0.05$.

It appears that the lower stated prices in the purchase intent question had little effect on the lower quintiles, but restricted reactions to the higher quintiles. Seeing $87.50 in a purchase intent question and then being asked for a maximum bid may have biased those with arbitrary scores ≥$88.

Summary of Purchase Intent Results

Table 29 shows the results of the purchase intent (PI) questions for each group. An examination of the sum of those who said they definitely or probably would buy, ("top two buy" scores, or T2B scores), shows the order made little difference. In three of the four instances, the T2B scores are statistically the same in Groups 1 and 2. For the highest price, it is directionally different, suggesting completing the AC task first pulled interest down at the highest price.

Table 29. Purchase Intent Question Results by Price Points and Groups

	$12.50		$37.50		$62.50		$87.50	
	Grp 1	Grp 2	Grp 1	Grp 2	Grp 1	Grp 2	Grp 1	Grp 2
Base	*50*	*50*	*50*	*50*	*50*	*51*	*50*	*50*
Definitely buy	26%	22%	8%	18%	8%	6%	2%	14%
Probably buy	30%	20%	14%	16%	6%	12%	4%	4%
Might or might not	16%	32%	16%	22%	18%	20%	6%	20%
Probably not buy	10%	14%	32%	24%	16%	18%	20%	14%
Definitely not buy	18%	2%	30%	20%	52%	45%	68%	48%
Total	100%	100%	100%	100%	100%	100%	100%	100%
Definitely+Probably	56%	52%	22%	34%	14%	18%	6%	18%

One of the objectives of this final study was to examine the extent a demand curve based on AC results parallels one derived from more tradition monadic testing. Based on the results shown in Figures 2 and 3, it looks quite promising that an AC task can be employed to create a price-demand curve. Figure 2 is a graph of the percentages of respondents willing to pay the price range listed on the x-axis. In Figure 3, the same curve is shown, but with only the four price points (plus ≥$88) presented to Group 2 respondents. The curves are quite similar. It appears that if a product being tested is susceptible to arbitrary coherence, then the results would be reliable for creating a price demand curve.

The regression lines can be interpreted in the following manner. In Figure 2, for each $1 increase in price, the percentage of those willing to buy at that price drops by .284 times the log of the price. The regression is only valid over the data range of $0.00 to $99.00 (technically cannot be extrapolated beyond that range). In Figure 3, for each $1 increase in price, the percentage of those willing to buy at that price drops by .175 times the log of the price. From visual inspection of both graphs, the AC test shown in Figure 2 appears to provide a more reliable price-demand curve than the monadic price test shown in Figure 3.

Figure 2. Price Demand Curve from Arbitrary Coherence Data (Group 1)

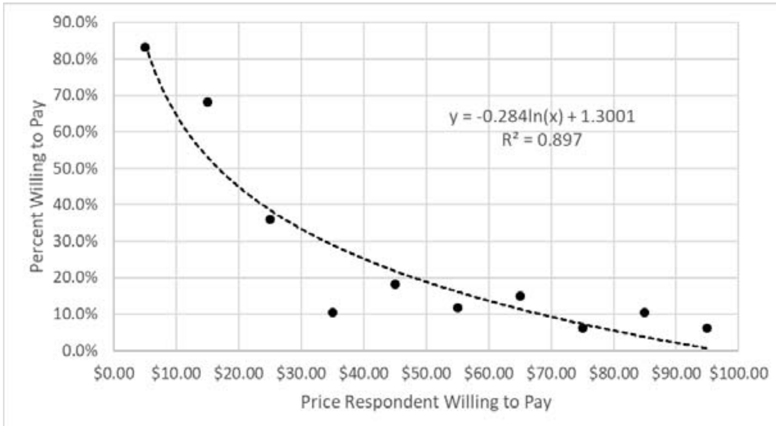
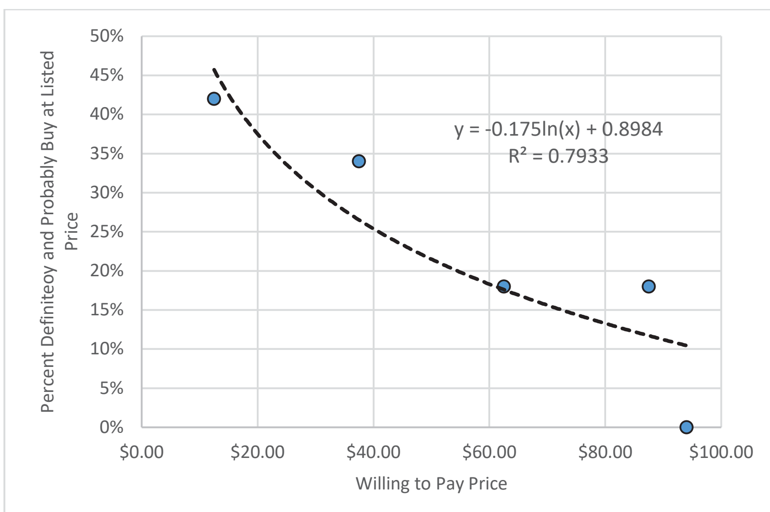


Figure 3. Price Demand Curve from Purchase Intent Data (Group 2)



Summary of AC5 Results

Under the carefully specified conditions created for this study the AC phenomenon is clearly present. Also, of interest is that whether the AC task was first or second had little effect on the results of the PI task, which were very similar regardless. The exception was at the highest price level asked in PI ($87.50), in which the AC task depressed the percent willing to buy. This may have been an anchoring effect, since most of the arbitrary numbers were lower than that price point.

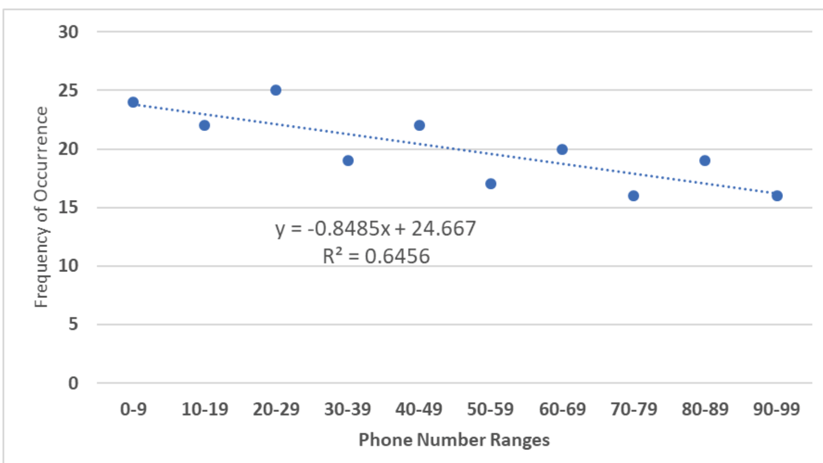
In the reverse order, half saw PI prices first that were lower than $63, so there appears to be an anchoring effect that way as well. The AC correlation was lessened, with the middle price point higher than when the AC task was first and those getting the AC task second indicating lower prices in the higher arbitrary number quintiles.

Observation on a Random Number Distribution

The original work by Ariely used the last two digits of the social security number as the random number which seems quite reasonable. With the more recent privacy concerns around social security numbers, it seemed reasonable to use the last two digits of a person's telephone number as the random number for testing purposes. While creating the tables for AC5 it was observed that the random numbers were not quite so randomly distributed. As shown in Figure 4, as the decile range increased the frequency of occurrence decreased.

This phenomenon does not adversely affect the results of AC4 or AC5 since any random number is sufficient for the study of arbitrary coherence. However, if this phenomenon is not an aberration, then research requiring a uniform distribution of random numbers using a telephone number may need to sample in such a way as to obtain a more uniform distribution of respondents.

Figure 4. Random Number Distribution from AC5 Study



It would be interesting to determine if the last two digits of one's telephone number is not as random as one would expect or if respondents are also reluctant to give their actual numbers. If respondents are making up a number, it seems remarkable that the distribution is so linearly predicable. It is unlikely however, that any more than a small percentage of people willing to cooperate in an interview of this nature would deliberately mislead. Even more amazing is that numbers beginning with 2, 4, 6 and 8 consistently deviate positively from the regression line and numbers beginning with 1, 3, 5 and 7 consistently deviate negatively from the regression line.

CONCLUSIONS

There are two aspects of arbitrary coherence. The first is that the anchor, which everyone knows is arbitrary, influences the judgment of subjects in the selection of a starting point when asked about pricing in an ambiguous situation. As noted, in our three-question work, we found (along with Wilson) that some processing of the arbitrary number must occur to see the influence. Hence, no effect when “question two,” the purchase decision, is not asked.

There are two hypotheses for why the arbitrary number has this effect. One is the “reach back” hypothesis proposed by Kahnemann. Subjects need something to think of when asked the third question (the auction question) and so they reach back for a handy recent number. The other is the “imprinting” hypothesis of Ariely et. al., similar to the “carry forward” alternative Wilson discusses. The idea is simply that since the arbitrary number is still in mind, it has an influence.

The second aspect of arbitrary coherence is sequentiality, or coherence. The first WTP influences all subsequent ones. However, the three-question demonstration does not show sequential or coherent impact. In it, each subject is exposed to a single arbitrary number and creates a single judgment. If the anchoring effect is not present, then there is no “coherence” using this method. Ariely et. al. did find evidence for sequential impact in the other five “annoying sound” experiments they conducted, since in those the subjects experienced repeated exposures to sounds of varying durations. However, the “arbitrary coherence” shown in their Experiment 1 and (where it appears) in our work only demonstrates an anchoring effect. Coherence (when it is found) is an artifact produced by anchoring across a group of subjects.

While the “coherence” in arbitrary coherence is weak or nonexistent for some items tested, the “arbitrary” reaction, i.e., anchoring, is present. As noted previously, data were examined across the first four studies to see if the highest quartile gave a higher WTP than the lowest. This did occur in 19 of 24 instances. The fact that 19 of 24 instances show an impact of the arbitrary value in raising WTP, most by more than five dollars, suggest the anchor influenced the respondent’s willing to pay price.

The Semantic Differential results are challenging to interpret. While being the strongest, best, and most active is not conducive to arbitrary coherence or consistency in anchoring, the conclusion is based on relative ranks within these eight items, and suffers that limitation. Finally, despite its limitations, it appears from AC5 that the three arbitrary coherence questions can be quite useful in exploring the price-demand relationship, one that is often challenging in the development of new products.

REFERENCES

Ariely D, Loewenstein G & Prelac D. (2003). Coherent Arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, February.

Ariel, D. (2008). *Predictably Irrational*. Harper Collins, NY, NY.

Fudenberg D, Levine DK, and Maniadis Z, Reexamining Coherent Arbitrariness for the Evaluation of Common Goods and Simple Lotteries, published online at ResearchGate. Uploaded 16 June 2015.

Kahneman, D. (2010). T*hinking, Fast and Slow*, Farrar, Straus and Giroux, New York, NY.

Pothier DD, Mangal T, Loizou P, Mason KA. (2009). Arbitrary coherence in theoretical decision making about surgical training: the effect of irrelevant subliminal anchoring. *Journal of Surgical Education,* 66 (3), 129-131.

Poundstone, W. (2010). *Priceless*, Hill and Wang, a division of Farrar, Straus, and Giroux, NY, NY.

Scott P.J. & Lazier C. (2012). Consumer house price judgements: new evidence of anchoring and arbitrary coherence. *Journal of Property Research,* 29 (1), 49-68.

Tversky A. & Kahneman D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, New Series, Vol. 185, No. 4157.

Unknown (2017). Scaling: Semantic Differential. Retrieved October 17, 2017 from http://psc.dss.ucdavis.edu/sommerb/sommerdemo/scaling/semdiff.htm.

APPENDICES

Appendix 1 - Images for Item Set 1 (AC1 and AC3)

Note: Maple syrup replaced Bailey's Irish Crème for AC3.

<table>
<tr>
<td>


Teleflora Spring Pitcher Floral Display
</td>
<td>


Dr. Scholl's Hand-Stitched Leather Loafers
</td>
</tr>
<tr>
<td>


Sandy Flats Pure Maple Syrup – 250ml
</td>
<td>


Irish Spring 12-Bar Value Pack
</td>
</tr>
</table>

Appendix 2 - 3 Question Survey

Thank you in advance for your help with this project. There are three questions. When the project is complete, we will be happy to share the results with you.

A. In column A, next to each of the four items on the list below, please enter the last two digits of you Social Security number (SS#) as if it were a price in dollars.

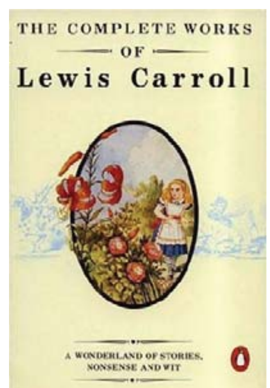
   For example, if your SS# is 678-54-9876, you would put $76. If it is 123-45-0001, you would enter $01.

B. In column B, indicate if you would be willing to pay that amount for the item by writing in a Y for “yes” or an N for “no”.

C. In column C, pretend there is an auction for the item. In whole dollars, write in the maximum amount you would be willing to pay for it.

Item	A. SS# as $	B. Willing to Purchase for that price? (Y or N)	C. Max willing to pay:
Item 1			
Item 2			
Item 3			
Item 4			

Appendix 3 - Images for Item Set 2 (AC 2 and AC 3)

 **Crocodile Wallet**	 **Copper Cooking Bowl**
 **Necklace of Pearls from Tahiti**	 **The Complete Works of Lewis Carroll**

Appendix 4 - 2 Question Survey

Thank you in advance for your help with this project. There are two questions. When the project is complete, we will be happy to share the results with you.

A. In column A, next to each of the four items on the list below, please enter the last two digits of you Social Security number (SS#) as if it were a price in dollars.

B. For example, if your SS# is 678-54-9876, you would put $76. If it is 123-45-0001, you would enter $01.

C. In column C, pretend there is an auction for the item. In whole dollars, write in the maximum amount you would be willing to pay for it.

Item	A. SS# as $	B. Max willing to pay:
Item 1		
Item 2		
Item 3		
Item 4		

Appendix 5 – Part 1 of Questionnaire for AC5

1. How old were you on your last birthday?

2. Please select your gender:
   Female
   Male

3. What is your primary U.S. state of residence?
   [Drop down list.]

4. To ensure representation, please check the box best representing your race.
   Black/African American
   Mixed racial background
   White/Caucasian
   Prefer Not to Answer

5.Are you Hispanic?
   Yes
   No

6. Please select your level of education:
   Some High School or Less
   High School Graduate
   Some College
   4 Year College/University Degree
   Some Graduate School
   Master's Degree or Above
   Prefer Not to Answer

Randomize responses to Q.7. All must say Chocolate to continue.
7. Please select which of these foods you have enjoyed recently:
   Chocolate
   Jalapenos
   Hamburgers
   Ice Cream
   Pasta

[Illogical responses: if respondent said would pay $x at AC Q.2 and max bid <$x at AC Q3, OR would not pay $x at AC Q.2 and max bid > $x at AC Q3, point out and re-ask AC Q.2 and AC Q.3.]

Appendix 6 – Part 2 of Questionnaire for AC5

Fine Belgian Chocolate

This is a rich, quality Belgian chocolate. The ingredients have been regulated by law since 1894 when a minimum level of 35 percent pure cocoa was imposed, making Belgian chocolates renowned throughout the world. They are made according to a voluntary quality standard to ensure excellence.

Each question appeared on a separate page. The product and description stayed in view the entire time.

Ask of Group 1 first, Group 2 second

AC Q1. Please write the last two digits of your primary telephone number as if it were a price in dollars. For example, if your primary telephone number is 555-721-1701 , you would put $01. If it is 555-721-0171, you would enter $71.

AC Q2. Please indicate if you would be willing to pay the dollar amount you just listed for the item in the picture.

AC Q3. Pretend there is an auction for the pictured item. In whole dollars, write in the maximum amount you would be willing to pay for it.

Ask of Group 1 second, Group 2 first

PI Q1.How likely are you to buy the box of Belgian chocolates as shown. That is, would you say you... *[N = 50 for each of 4 prices for a box of 24 --$12.50; $37.50; $62.50; 87.50. Random assignment.]*

Definitely would buy it
Probably would buy it
Might or might not buy it
Probably would not buy it
Definitely would not buy it

DECISION SCIENCES INSTITUTE

Product Pricing and Allocation Strategies in a Prioritized Multi-Segment Market: An Extended Newsvendor Model

Hani I. Mesak
Louisiana Tech University
Email: mesak@latech.edu

Hongkai Zhang
East Central University
Email: hzhang@ecok.edu

Jiehui Ma
University of Wisconsin-Whitewater
Email: maj@uww.edu

ABSTRACT

In this paper, we tackle in the newsvendor context a problem of optimal product allocation and pricing under uncertain demand in a prioritized multi-segment market for a monopolistic retailer. A mathematical programming model is developed and then solved via dynamic programming to find the optimal allocation scheme and determine the prices correspondingly. A numerical experiment is designed and performed to solve the model via the dynamic programming approach and examine the effects of the key model parameters upon the retailer's optimal expected total profit and allocation schemes.

KEYWORDS: Product allocation, Pricing, Prioritized segmentation, Optimization, Dynamic programming

INTRODUCTION

The classical news-vendor problem aims at finding a product's order quantity that maximizes the expected profit in a single-period probabilistic demand framework. Khouja (1999) classifies two approaches to solve this problem. The first approach intends to minimize the expected costs of overestimating and underestimating demand. The second approach is to maximize the expected profit. A general extension of this classical newsvendor problem in conjunction with the second approach in a retail setting is how to allocate a fixed quantity of a product for sale and set the price in a multi-segment market to maximize the expected profit. We focus on a profit-maximizing monopolistic retailer facing uncertain demand at the aggregate level. In this paper, we address three research questions stated as follows: (i) What is the optimal scheme for a monopolistic retailer to allocate and price K identical units of a product in a prioritized n-segment market under uniformly distributed uncertain demand during a single selling season so that her expected total profit is maximized? (ii) What is the sensitivity of the optimal expected profit to simultaneous variations in the key parameters of the mathematical programming model? (iii) What is the relative importance of each of these parameters upon the optimal allocation scheme?

We make the following basic assumptions while addressing the three strategic issues mentioned above:

(i) The price charged by the retailer in each segment is the only independent variable that affects the aggregate demand in that segment.
(ii) Consumers in each segment are well-informed of the price charged in that segment, which is set at the beginning of a selling season.
(iii) The demand in each segment is heterogeneous and independently follows a uniform probability distribution conditioned by the price charged in that segment.
(iv) The allocation decisions are made following a predetermined prioritized sequence.

Priority service, a form of product differentiation, is widely used in practice. Priority-based capacity allocation allows the market to be segmented into a spectrum of priority classes. For example, those customers willing to pay higher prices are assigned higher priority in receiving the product or service (Chao & Wilson, 1987). We analyze a model of *n* prioritized customer segments, in which higher priority segments are served before any lower priority segment. If capacity is exhausted by higher priority segments, the lower priority segment goes unfilled. We are not deriving optimal prioritization; rather we assume that priorities are set. Factors to consider in setting priorities may include willingness to pay higher prices, expected volume of business and expected customer satisfaction, among others.

The addressed problem is envisioned to belong to the class of “inverse newsvendor problems.” The classical newsvendor problem is one of optimally choosing a level of capacity to respond to a known demand distribution. The inverse newsvendor problem is one of optimally choosing a demand distribution with fixed capacity. In many firms, capacity is relatively more difficult to adjust than demand because of the time required to install new capacity (Carr & Lovejoy, 2000).

This paper is organized as follows. In the next section, we review the relevant literature and position our study afterwards within the reviewed literature to highlight its contributions. In the third section, a mathematical programming model is developed. The fourth section presents a dynamic programming (DP) formulation developed to solve the mathematical programming model. A numerical study is conducted for the model in the fifth section. Finally, the paper concludes in the six section with a summary of its contributions, managerial implications, and directions for future research.

LITERATURE REVIEW

The newsvendor problem of setting quantity to maximize expected profit has been extensively addressed for several decades. Khouja (1995) and Qin et al. (2011) provide lucid reviews of the newsvendor literature. In a classical newsvendor environment, a retailer has little control over the exogenously determined selling price. In practice, however, a retailer may adjust the current selling price in order to shift demand (Qin et al., 2011). There is a substantial body of newsvendor literature exploring the impacts of the price-dependent demand on the seller’s pricing and ordering decisions (e.g., Whitin, 1955; Mills, 1959; Karlin & Carr, 1962; Lau & Lau, 1988; Chen et al., 2004; Arcelus et al., 2005; Wang et al., 2017; Berenguer et al., 2017).

Wang et al. (2014) argue that the newsvendor literature has focused on the quantity decision, whereas the revenue management literature has focused on the pricing decision. In the newsvendor literature, three notable studies are more relevant to our current study. Ng and Lee

(2008), applying a game theoretic approach, analyze the optimal capacity allocation in a duopolistic two-segment market where demand is deterministic. A closed-form solution is obtained in their study. Carr and Lovejoy (2000) study an inverse newsvendor problem where demand distributions are chosen from an opportunity set, which represents the set of market opportunities for the firm. The firm's profit is formulated as a function of those demand distributions. The objective is to maximize the firm's profit by finding the optimal demand distribution from the opportunity set. The results of the solution are extended to cases of multiple prioritized market segments sharing the firm's capacity, where the priorities are taken as given. In the work of Deng et al. (2008), demand in each segment of the market is assumed to independently follow a Poisson distribution. A marginal-revenue-based capacity management model is formulated in their study to manage stochastic demand in a three-segment market and develop the policies for the firm to allocate capacity to those segments of higher revenue.

In the revenue management literature, two recent studies are related to our current paper. Zhang (2012) investigates a multiple-segment price discrimination problem with different segments' demand elasticity and supply constraints, where all segments are subject to the limited common capacity. In addition, each segment is subject to a limited segment-specific capacity, while the firm incurs segment-specific production and selling costs in each segment. In his model, the demand function of each segment is deterministic. The objective is to maximize the firm's profit by finding the optimal price for each segment. Zeng (2013) proposes multiple pricing strategies for a retailer who operates in a two-segment market composed of experienced and inexperienced consumers.

In recent years, there is a growing research stream that focuses on both quantity and price decisions. Three papers of this research stream are noteworthy. Gallego et al. (2011) address the problem of capacity allocation and pricing in a two-segment market where an entrant sells at a lower price to compete against an established incumbent. Their model is designed to study the competitive interactions between the two rivals for different assumptions about market structure, demand and available capacity. Wang et al. (2014) tackle a price-setting problem with partial information. Their study focused on robust capacity and pricing decisions for an innovative product with limited consumer valuation and market size information. The objective is to minimize the maximum loss in expected profit due to incomplete information on the probability distribution of demand. Closed-form expressions are derived for the optimal quantity and price. Duan et al. (2016) present a model as a three-stage dynamic game for the allocation of a limited spectrum capacity and the pricing decision in a wireless market with two segments for femtocell and macro-cell services.

Our present study possesses the following five characteristics that make it distinct from the articles cited above: First, a general n-segment market in a prioritized setting is taken into account in our modeling framework. Second, the uncertain demand in each segment is modeled as a continuous random variable following a uniform probability distribution. Third, a mathematical programming model is formulated and then solved via dynamic programing (DP) to find the optimal allocation scheme, which then determines the pricing strategy. Fourth, a numerical experiment is designed and implemented to solve the mathematical programming model with the DP approach. Fifth, we perform a global sensitivity analysis of the mathematical programming model using the approach proposed by Wagner (1995) and subsequently examine the relative importance of each key parameter on the model's optimal output as in Druehl et al. (2009).

MODEL DEVELOPMENT

Let us consider a monopolistic retailer who has K identical units of a product to be allocated in a general n-segment market during a single selling season. The n segments, denoted as segments i (i = 1, 2, …, n), are prioritized to receive their allocations of the product in such a way that the allocation decision pertaining to segment i is made prior to the allocation decision related to segment i + 1. Without loss of generality, we assume that the K units of the product are successively allocated to segments i (i = 1, 2, …, n).

The following notations are used to formulate the retailer's profit and price functions:

K	the quantity of the product available for sale in the n-segment market;
K_i	the quantity of the product available for sale in segment i;
y_i	the quantity of the product to be allocated to segment i (a decision variable);
P_i	the price per unit of the product charged in segment i ($P_i > 0$);
C_i	the cost per unit of the product allocated to segment i ($C_i > 0$);
S_i	the salvage value per unit of the product disposed of after the selling season in segment i ($S_i > 0$);
d_i	the aggregate demand of consumers in segment i;
ω_i	the demand parameter in segment i ($\omega_i > 0$);
$f(x\|\omega_i)$	the probability density function (p.d.f.) of d_i being a continuous random variable;
$E(d_i)$	the expected value of d_i;
π_i	the retailer's profit yielded from segment i;
$E(\pi_i)$	the expected value of π_i;
π	the retailer's total profit yielded from the entire n-segment market;
$E(\pi)$	the expected value of π.

Consumers' price sensitivity may substantially vary across segments, a thorough understanding of the dynamic interaction between price and demand is required to find the optimal allocation scheme and set the price accordingly in each segment of the market. Since consumers in segment i are assumed to be well-informed of the price charged in that segment, they would take the price into consideration while making their purchases. Thus, the uncertain demand of segment i, d_i (i = 1, 2, …, n) can be modeled as a random variable following a probability distribution conditioned by P_i. Following Azoury (1985), we model the aggregate demand, d_i, as a continuous random variable following a uniform probability distribution. The p.d.f. of d_i is given by

$$f(x|\omega_i) = \begin{cases} 1/\omega_i, & if \quad 0 < x < \omega_i, \\ 0, & if \quad x \geq \omega_i, \end{cases} \tag{1}$$

where $\omega_i > 0$. A price-dependent model proposed by Karlin and Carr (1962) is used in this study to define the functional relationship between the demand parameter ω_i and selling price P_i:

$$\omega_i = \alpha_i P_i^{-\beta_i} \tag{2}$$

where, $\alpha_i > 0$ and $\beta_i > 1$ (i = 1, 2, …, n).

In expression (2), the parameter, α_i, reflects the size of segment *i*. A larger value of α_i shows stronger demand in segment *i*. The power parameter β_i measures the responsiveness of the aggregate demand of segment *i* to the price charged in that segment. According to Huang et al. (2013), one of the advantages of this model is that it does not require a finite upper limit on price. The model is also called the constant elasticity model because the demand function exhibits a constant elasticity (i.e., $-\beta_i$) (see Huang et al., 2013).

The expected demand of segment *i*, based on expression (1), is given by

$$E(d_i) = \int_0^{\omega_i} \frac{x}{\omega_i} dx = \frac{\omega_i}{2}\,. \tag{3}$$

Expression (3) reveals that the demand parameter, ω_i, equals twice the expected demand, $E(d_i)$, and thus serves as an indicator of the aggregate demand in segment *i*.

From expressions (2) and (3), the expected demand of segment *i*, can be rewritten as follows:

$$E(d_i) = \frac{1}{2}\alpha_i P_i^{-\beta_i} \tag{4}$$

Cost-plus pricing is the practice of "adding a standard markup to the cost of the product" (see Kotler & Armstrong, 2004, p.357). Zeithaml et al. (1985) report that 63% of the studied firms base their prices primarily on costs. Also, there are studies in the literature that incorporate price markup in their modeling frameworks (e.g., Shah & Jha, 1991; Hanson, 1992). In this paper, the unit price charged in segment *i*, P_i (i = 1, 2, …, n), is set by a markup over the unit cost; that is, $P_i = \theta_i C_i$ where $\theta_i > 1$.

As the risk of larger orders is mainly carried by buyers, sellers are often more than willing to offer quantity discounts when inventory costs are high. Quantity discounts help sellers identify price-insensitive buyers from the price-sensitive ones to practice price discrimination (Qin et al., 2011). While implementing price discrimination, a retailer usually charges a higher price when the quantity of the product available for sale is relatively scarce and a lower price when it is relatively abundant. We may therefore assume that the price markup factor, θ_i, is a decreasing function of K_i :

$$\theta_i = \phi(K_i), \tag{5}$$

where, $d\phi(K_i)/dK_i < 0$. Accordingly, the price charged in segment *i* is a function of K_i :

$$P_i = C_i\phi(K_i) \tag{6}$$

In segment *i* (i = 1, 2, …, n), if the demand (d_i) exceeds the quantity of the product allocated by the retailer (y_i), the profit (π_i) will equal the profit per unit multiplied by the number of units sold. On the other hand, if d_i is smaller than y_i, a portion of the allocated quantity, $y_i - d_i$, will be unsold and disposed of after the selling season at the unit salvage value, S_i. Hence, the retailer's profit yielded from segment *i* (i = 1, 2, …, n) is stated as

$$\pi_i = \begin{cases} P_i d_i - C_i y_i + S_i (y_i - d_i), & d_i \le y_i, \\ (P_i - C_i) y_i, & d_i > y_i. \end{cases} \tag{7}$$

Since d_i is modeled as a continuous random variable, we obtain from expression (7) the expected value of π_i:

$$E(\pi_i) = (P_i - S_i)\int_0^{y_i} x f_i(x|\omega_i)dx - (C_i - S_i) y_i \int_0^{y_i} f_i(x|\omega_i)dx + (P_i - C_i) y_i \int_{y_i}^{\infty} f_i(x|\omega_i)dx. \tag{8}$$

Substituting (1) into (8) and carrying out the integrations, we obtain

$$E(\pi_i) = (P_i - C_i) y_i - \frac{(P_i - S_i)}{2\omega_i} y_i^2, \quad if \quad y_i < \omega_i; \tag{9}$$

$$E(\pi_i) = \frac{(P_i - S_i)}{2}\omega_i - (C_i - S_i) y_i, \quad if \quad y_i \ge \omega_i. \tag{10}$$

The derivations of expressions (9) and (10) are found in the Appendix.

The retailer's expected total profit from the entire *n*-segment market can be expressed as

$$E(\pi) = \sum_{i=1}^{n} E(\pi_i). \tag{11}$$

Given *K* identical units of a product, which is to be allocated by the retailer during a single selling season for sale in the *n*-segment market, we aim at finding the optimal quantity to be allocated to segment *i* (i = 1, 2, …, n), y_i^*, to maximize the expected total profit. The problem may therefore be formulated as the following mathematical programming model:

$$\underset{y_1, y_2, \ldots, y_n}{Max} \sum_{i=1}^{n} E(\pi_i)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} y_i \le K,$$

$$y_i \ge 0 \text{ for } i = 1, 2, \ldots, n. \tag{12}$$

The model (12) can be cast into a dynamic programming (DP) formulation that is solved by applying the principle of decomposition. The DP formulation is presented next.

DYNAMIC PROGRAMMING FORMULATION

There are six key components in the DP formulation: (i) the sequence of decision stages, (ii) the input state variable of each stage, (iii) the decision variable of each stage, (iv) the transition function linking the input and output state variables of each stage, (v) the return at each stage, and (vi) the recursive relationship. These components, shown in Figure 1, are identified and discussed below.

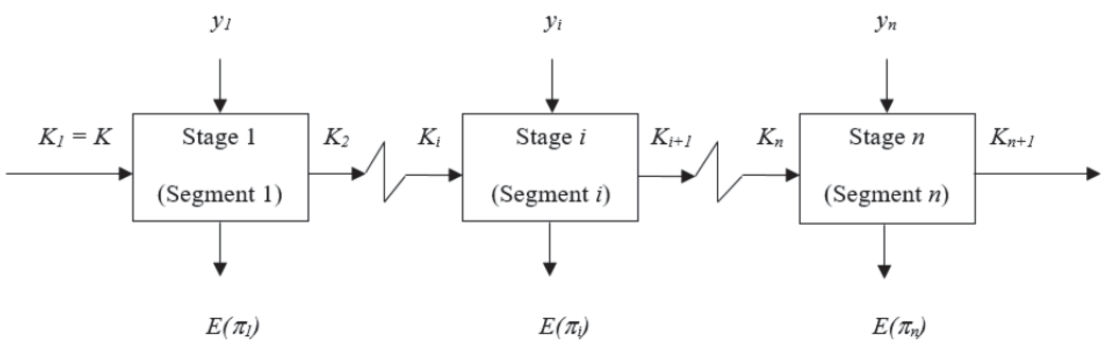


Fig. 1. A dynamic programming representation for the mathematical programming model (12)

Each segment of the market stands for a decision stage to which a certain quantity of the product is allocated. These consecutive stages are indexed corresponding to the indices of the segments defined in the previous section. The mathematical programming model (12) is decomposed into a sequence of smaller and computationally simpler sub-problems, as discussed in the remainder of this section.

The input state variable of stage i is the quantity of the product available for allocation at the beginning of the stage, K_i (i = 1, 2, …, n). As shown in Figure 1, each stage's output state variable K_{i+1}, serves as the input state variable of the next stage. In particular, K_{n+1} is the output state variable of the last stage, which is the quantity of the product unallocated for the purpose of maximizing the expected total profit. The quantity allocated to stage i, y_i (i = 1, 2, …, n), is a decision variable. As a whole, there are n decision variables in the DP formulation for the n-segment market. The transition function defines the linkage between the input and output state variables of a stage and the decision made for the stage, and may be expressed as

$$K_{i+1} = t_i\,(K_i, y_i),$$

where,
$K_1 = K$ is given;
$K_{i+1} = K_i - y_i$, $i = 1, 2, \ldots, n-1$;
$t_i(\cdot)$ stands for 'transition function of.'

Expressions (9) and (10) both show that the expected profit yielded from stage i (i.e., the return at stage i), $E(\pi_i)$, is determined by the decision variable y_i, where $y_i \le K_i$ and the price P_i. It is shown in expression (6) that P_i is a function of K_i. Hence, $E(\pi_i)$ is a function of K_i and y_i and can be expressed as $E(\pi_i\,(K_i, y_i))$.

A backward induction process is employed to formulate the recursive relationship, which links the optimal decision for a stage to the optimal decisions made in previously considered stages. Starting from stage n, the recursive relationship is given by expressions (13) and (14).

For stage n,

$$F_n^*(K_n) = \max_{\forall y_n \le K_n} E(\pi_n(K_n, y_n)). \tag{13}$$

For stage i = 1, 2, …, n-1,

$$F_i^*(K_i) = \underset{\forall y_i \le K_i}{Max}\{E(\pi_i(K_i, y_i)) + F_{i+1}^*(K_{i+1})\}, \tag{14}$$

where $K_{i+1} = K_i - y_i$.

The optimal solution to the DP model formulated above, y_i^* (i = 1, 2, …, n), is a function of the input state variable K_i and hence can be expressed as $y_i^*(K_i)$. The recursive optimization is carried out in backward fashion until the first stage is reached. At stage 1, the maximum total return (i.e., the expected total profit), $F_1^*(K_1)$, and the corresponding optimal quantity to be allocated to stage 1, $y_1^* = y_1^*(K_1)$, are both determined. $y_1^*(K_1)$ is a unique value because $K_1 = K$ is a given constant. Then, it is possible to backtrack from the first stage through the succeeding stages to obtain the optimal quantity to be allocated to each of the other stages in the following manner: At stage 2, we compute the optimal input value $K_2^* = K_1 - y_1^*$ and then determine the optimal quantity to be allocated to stage 2 through the function $y_2^* = y_2^*(K_2^*)$. Afterwards, for stage i (i = 3, 4, …, n), we compute the optimal input state value $K_i^* = K_{i-1}^* - y_{i-1}^*$ and then determine the optimal quantity to be allocated to stage i through the function $y_i^* = y_i^*(K_i^*)$. Using the optimal input state value K_i^*, the optimal price P_i^* (i = 1, 2…, n) is determined through the price function (6).

NUMERICAL EXPERIMENT

In this section, we design and conduct a numerical experiment to (i) find the optimal solution to the mathematical programming model (12) with the DP formulation presented in the fourth section, (ii) perform a global sensitivity analysis of the model, and (iii) examine the effects of eight key parameters on the optimal allocation scheme. In the numerical study, the price markup factor (5) in segment i, θ_i, is specified to take the following functional form:

$$\theta_i = \phi(K_i) = \frac{b_1 + K_i}{b_2 + K_i}, \tag{15}$$

where, $b_1 > b_2 > 0$. It can be easily shown that $d\phi(K_i)/dK_i < 0$ and $d^2\phi(K_i)/dK_i^2 > 0$, implying that $\phi(\cdot)$ is a convex decreasing function of K_i. We set b_1 = 2000, b_2 = 1 and, K = 10000, 15000, 20000, and 25000 units, respectively. Due to the RAM limit of the computer used to run the DP formulation, we consider $C_i = C$ and $S_i = S$ (i = 1, 2, 3) in our numerical study. The results of the numerical study for K = 15000 units are reported in this paper. All the unreported results for K = 10000, 20000, and 25000 units are available from the authors upon request.

Experimental Design and Implementation

Table 1 displays six marketing parameters, α_i and β_i (i = 1, 2, 3) and two operations parameters, C and S, encompassed in our factorial design. Three values (factor levels) are chosen for every single parameter. This factorial structure has a total of 3^8 = 6561 possible combinations (treatments), which makes our numerical experiment extensive and robust. Given K, b_1 and b_2, each treatment leads to a DP optimal solution.

Table 1. Factor levels of eight key parameters

Parameter	Factor levels
α_1	0.8×10^7 , 1.0×10^7, 1.2×10^7
α_2	0.8×10^7 , 1.0×10^7, 1.2×10^7
α_3	0.8×10^7 , 1.0×10^7, 1.2×10^7
β_1	1.8, 2.0, 2.2
β_2	1.8, 2.0, 2.2
β_3	1.8, 2.0, 2.2
C	20, 25, 30
S	0, 5, 10

The factor levels of α_i and β_i are chosen based on their domains defined by Karlin and Carr (1962). The levels of the parameter, C, are chosen to ensure that the cost per unit is smaller than the corresponding selling price. The choice of the factor levels for S reflects the assumption that the salvage value per unit is smaller than the cost per unit.

Profit Optimization

To improve exposition, we confine the decision variables y_i to take on nonnegative integer values. For each of the four levels of K mentioned above, the domain of the decision variable y_i (i = 1, 2, 3) is uniformly discretized as $\{500q \mid q = 0, 1, \ldots, K/500\}$. The mathematical programming model (12) is solved through the DP formulation for every single treatment of the factorial design. A computer program is developed by coding in C++ the recursive relationship characterized by expressions (13) and (14) and the backtracking procedure described in Section 4. Given the factorial design, we obtained 6561 optimal solutions of the model (12) for each of the four values of K along with the values of b_1 and b_2.

Table 2 displays the optimal solution for four treatments, respectively. For instance, given the combination of the parameters' values in Example 1, the retailer should allocate 8500, 5500 and 1000 units to segments1, 2 and 3, respectively. As a result, her expected total profit would be $55306.86, and the unallocated quantity of the product is zero unit. It is noted in Table 2 that the retailer could earn a higher expected total profit by allocating larger quantities of the product to satisfy the stronger demand of those consumers who are less price-sensitive. It is also found that the retailer may keep certain quantities of the product unallocated to maximize her profit if consumers are more price-sensitive.

Table 2. Examples of the DP optimal solutions
(K = 15000 units, C = 25, S = 5, $\alpha_i = 1.0\times10^7$, i = 1, 2, 3)

	β_1	β_2	β_3	y_1^*	y_2^*	y_3^*	$K-\Sigma y_i^*$
Example 1	1.8	1.8	1.8	8500	5500	1000	0
Example 2	1.8	2.0	2.2	5500	2500	1500	5500
Example 3	2.2	2.0	1.8	1500	2500	4000	7000
Example 4	2.2	2.2	2.2	1000	1000	1000	12000
	P_1^*	P_2^*	P_3^*	ω_1^*	ω_2^*	ω_3^*	$E^*(\pi)$
Example 1	28.33	32.69	74.93	24317.4	18798.3	4223.5	55306.86
Example 2	28.33	30.26	32.14	24317.4	10921.0	4836.7	14128.22
Example 3	28.33	28.70	29.54	6382.8	12139.2	22552.2	13502.21
Example 4	28.33	28.57	28.84	6382.8	6266.4	6135.9	5093.48

Parameter Effects Upon Profit

We follow the approach of global sensitivity analysis introduced by Wagner (1995) to investigate the effects of simultaneous variations of the parameters listed in Table 1 upon the retailer's expected total profit. A multiple linear regression is performed after all the 6561 treatments are completed with the eight parameters as independent variables in the regression and the expected total profit as the dependent variable. In order to circumvent potential multicollinearity, all the independent variables are centered around their respective means before the regression is run. The magnitude of the effect of each parameter on the expected total profit is gauged by the absolute value of the parameter's t-statistic generated by the regression. Table 3 summarizes the regression results with the R^2 value = 0.8414. As shown in Table 3, the expected total profit is positively impacted by α_i (i = 1, 2, 3) and S but negatively impacted by β_i (i = 1, 2, 3) and C.

Table 3. Regression results for the expected total profit as the dependent variable (K = 15000 units)

Parameter (Independent variable)	t-statistic
α_1	26.67229367*
α_2	22.09137115*
α_3	17.00855894*
β_1	-81.52676609*
β_2	-71.64766765*
β_3	-64.59967864*
C	-112.2403325*
S	68.72406362*

"*": p-value < 0.001.

Table 4. Regression results with interactions for the expected total profit (K = 15000 units)

Parameter (Independent variable)	t-statistic
α_1	44.195*
α_2	37.201*
α_3	28.641*
β_1	-137.287*
β_2	-120.651*
β_3	-108.782*
C	-189.006*
S	115.727*
$\alpha_1 \times \alpha_2$	0.898
$\alpha_1 \times \alpha_3$	1.513
$\alpha_1 \times \beta_1$	-24.410*
$\alpha_1 \times \beta_2$	-4.867*
$\alpha_1 \times \beta_3$	-8.550*
$\alpha_1 \times C$	-15.401*
$\alpha_1 \times S$	7.235*
$\alpha_2 \times \alpha_3$	1.344
$\alpha_2 \times \beta_1$	-3.697 *
$\alpha_2 \times \beta_2$	-16.049*
$\alpha_2 \times \beta_3$	-7.275*
$\alpha_2 \times C$	-10.575*
$\alpha_2 \times S$	5.399*
$\alpha_3 \times \beta_1$	-4.701*
$\alpha_3 \times \beta_2$	-4.475*
$\alpha_3 \times \beta_3$	-3.945*
$\alpha_3 \times C$	-7.068*
$\alpha_3 \times S$	4.900*
$\beta_1 \times \beta_2$	17.311*
$\beta_1 \times \beta_3$	26.071*
$\beta_1 \times C$	51.699*
$\beta_1 \times S$	-25.998*
$\beta_2 \times \beta_3$	23.837*
$\beta_2 \times C$	37.375*
$\beta_2 \times S$	-20.260*
$\beta_3 \times C$	28.417*
$\beta_3 \times S$	-20.220*
$C \times S$	-50.512*

"*": p-value < 0.001; otherwise, p-value > 0.1.

We also investigate the effects of two-way interactions among the eight parameters mentioned earlier. Table 4 shows that the three interaction terms among the segment-size parameters are insignificant, whereas all the other interaction terms turn out to be significant. Given the absolute values of their t –statistics in Tables 3 and 4, the eight parameters are ranked as $C > \beta_1 > \beta_2 >$

$S > \beta_3 > \alpha_1 > \alpha_2 > \alpha_3$ with respect to the importance of their effects upon the expected total profit. Apparently, the cost per unit exerts a leading influence on the expected total profit. The three price-sensitivity parameters, as a whole, are more influential than those segment-size parameters. The price-sensitivity and segment-size parameters of the segment of a higher priority have larger effects than their counterparts of a lower-priority segment, respectively.

Parameter Effects Upon the Allocation Scheme

Based on the factorial design, each level of a parameter is exclusively associated with 6561/3 = 2187 treatments. Out of the 2187 treatments associated with each level of a parameter, we now focus on the proportion of those treatments where the quantity K is not fully allocated based on the optimal allocation scheme. If a parameter's influence on the outcome of "partial allocation" is insignificant, then we should expect that the proportion of "partial allocation" is roughly equal across its three levels, classified as L (low), M (moderate), and H (high). This hypothesis is statistically tested. As shown in Table 5, the hypothesis of the equality of the three proportions is rejected for each of the eight parameters.

Table 5. Results of the 4084 treatments where $y_1^* + y_2^* + y_3^* < 15000$

Parameter	α_1	α_2	α_3	β_1	β_2	β_3	C	S
Frequency observed at the low level	1460	1449	1515	1014	1045	949	581	1614
Frequency observed at the medium level	1357	1360	1337	1408	1397	1270	1481	1359
Frequency observed at the high level	1267	1275	1232	1662	1642	1865	2022	1111
χ^2 – value †	36.3	29.5	79.6	414.9	350.5	840.6	2061.9	246.2

† For each parameter, the proportion of $y_1^* + y_2^* + y_3^* < 15000$ is significantly different across its three levels at p-value < 0.001.

For every single parameter, the Marascuilo procedure (Anderson et al., 2016) is used to make pairwise comparisons of the three proportions. Table 6 indicates that the proportions in each pair are significantly different from each other. This finding implies that each parameter plays a significant part in the outcome of "partial allocation." Furthermore, Tables 5 and 6 reveal that given K = 15000 units, the optimal partial allocation of the product by the retailer is more likely to occur at the lower levels of α_i (i = 1, 2, 3) and S and the higher levels of β_i (i = 1, 2, 3) and C.

The general results were found consistent with respect to the remaining three values of *K* not reported herein.

Table 6. Results of the Marascuilo procedure: The absolute differences in proportion pairs and the corresponding critical ranges (i.e., the numbers in parentheses)

Parameter	(L, M)	(L, H)	(M, H)
α_1	0.0471** (0.0439)	0.0882** (0.0443)	0.0412* (0.0449)
α_2	0.0407* (0.0440)	0.0796** (0.0443)	0.0389* (0.0449)
α_3	0.0814** (0.0436)	0.1294** (0.0440)	0.0480** (0.0451)
β_1	0.1802** (0.0449)	0.2963** (0.0426)	0.1161** (0.0416)
β_2	0.1610** (0.0450)	0.2730** (0.0429)	0.1120** (0.0419)
β_3	0.1468** (0.0454)	0.4188** (0.0395)	0.2721** (0.0394)
C	0.4115** (0.0417)	0.6589** (0.0334)	0.2474** (0.0348)
S	0.1166** (0.0425)	0.2300** (0.0432)	0.1134** (0.0452)

"**": p-value < 0.001; "*": 0.001 < p-value < 0.05.

SUMMARY AND CONCLUSIONS

In this paper, we tackle a general newsvendor problem of optimal product allocation and pricing for a profit-maximizing monopolistic retailer in a prioritized multi-segment market. A price-dependent demand model proposed by Karlin and Carr (1962) is employed to describe the functional relationship between the selling price and the aggregate consumer demand. A mathematical programming model is formulated and then numerically solved via dynamic programming for an exogenously given quantity of the product available for allocation over a three-segment market to determine the optimal allocation scheme and the corresponding pricing strategy for all the 6561 treatments of a factorial design.

The presented model is not only a valuable decision tool for making pricing and allocation decisions but also provides valuable insights on the roles of the six marketing parameters and two operations parameters. The effects of these parameters on the retailer's expected total profit and optimal allocation scheme are statistically investigated. The findings of the numerical study offer practical managerial implication for a retailer. Given a certain quantity of the product to be allocated to a multi-segment product, the optimal allocation scheme should be tilted towards those segments exhibiting low price-sensitivity. Reducing the cost per unit is the most important task to carry out due to its significant impact on profit. While enhancing the values of the marketing parameters of the same type to improve profitability, the retailer should focus her efforts on those segments of higher priorities. The price-sensitivity parameters as a whole deserve more enhancing efforts than the segment-size parameters. In addition, a partial-allocation scheme may maximize the retailer's profit.

This exploratory study suggests some possibilities for future research. First, we have addressed a problem of optimal product allocation and pricing for a monopolistic retailer in the newsvendor context. An interesting research direction would be to solve the problem in a competitive environment. Second, although the findings reported in this paper are appealing, their generalizability should be considered with care. Third, the mathematical programming model developed in this study is based on the uniformly distributed uncertain demand. Consumer demand following other probability distributions may be incorporated in the modeling framework in future research. Fourth, we employ a price-dependent demand model in this research. In future extensions, different demand models such as the price and lead time-dependent models could be considered for the profit-optimization framework.

APPENDIX

Derivation of Expression (9)

Given $y_i < \omega_i$, substituting (1) into (8) and carrying out the three integrations yield:

$$\begin{aligned} E(\pi_i) &= (P_i - S_i)\int_0^{y_i} \frac{x}{\omega_i}\,dx - (C_i - S_i)y_i\int_0^{y_i}\frac{1}{\omega_i}\,dx + (P_i - C_i)y_i\int_{y_i}^{\omega_i}\frac{1}{\omega_i}\,dx \\ &= \frac{(P_i - S_i)}{2\omega_i}y_i^2 - \frac{(C_i - S_i)}{\omega_i}y_i^2 + (P_i - C_i)y_i - \frac{(P_i - C_i)}{\omega_i}y_i^2 \\ &= (P_i - C_i)y_i - \frac{(P_i - S_i)}{2\omega_i}y_i^2. \end{aligned} \tag{A.1}$$

Derivation of expression (10)

Given $y_i \geq \omega_i$, substituting (1) into (8) and carrying out the three integrations yield:

$$E(\pi_i) = (P_i - S_i)\int_0^{\omega_i}\frac{x}{\omega_i}\,dx - (C_i - S_i)y_i\int_0^{\omega_i}\frac{1}{\omega_i}\,dx = \frac{(P_i - S_i)}{2}\omega_i - (C_i - S_i)y_i. \tag{A.2}$$

REFERENCES

Anderson, D.R., Sweeney, D.J., Williams, T.A., Camm, J.D.& Cochran, J.J. (2016). Essentials of modern business statistics (7th Edition), Cengage Learning.

Arcelus, F.J., Kumar, S., & Srinivasan, G. (2005). Retailer's response to alternate manufacturer's incentives under a single-period, price-dependent, stochastic-demand framework. *Decision Sciences*, 36 (4), 599–626.

Azoury, K.S. (1985). Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science*, 31 (9), 1150–1160.

Berenguer. G., Feng, Q., Shanthikumar, J. G., & Xu, L. (2017). The effects of subsidies on increasing consumption through for-profit and not-for-profit newsvendors, *Production and Operations Management,* 26 (6), 1191–1206.

Carr, S., & Lovejoy, W. (2000). The inverse newsvendor problem: Choosing an optimal demand portfolio for capacitated resources. *Management Science,* 46 (7), 912–927.

Chao, H., & Wilson, R. (1987). Priority service: pricing, investment, and market organization. *The American Economic Review*, 77 (5), 899–916.

Chen, F. Y., Yan, H. , & Yao, L. (2004). A newsvendor pricing game. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(4), 450–456.

Deng, H., Wang, Q., Leong, G.K., & Sun, S.X. (2008). The usage of opportunity cost to maximize performance in revenue management. *Decision Sciences*, 39 (4), 737–758.

Druehl, C., Schmidt, G., & Souza, G. (2009). The optimal pace of product updates. *European Journal of Operational Research*, 192, 621–633.

Duan, L., Shou, B., & Huang, J. (2016). Capacity allocation and pricing strategies for new wireless services. *Production and Operations Management*, 25 (5), 866–882.

Gallego, G., Krishnamoorthy, S., & Phillips, R. (2011). Competitive revenue management with forward and spot markets. *Journal of Revenue and Pricing Management*, 10 (2), 132–160.

Hanson, W. (1992). The dynamics of cost-plus pricing. *Managerial and Decision Economics*, 13 (2), March/April, 149–161.

Huang, J., Leng M., & Parlar, M. (2013). Demand functions in decision modeling: A comprehensive survey and research directions. *Decision Sciences*, 44 (3), June, 557-609.

Karlin, S., & Carr, C. R. (1962). Prices and optimal inventory policy. In H. Scarf K. Arrow, S. Karlin (Eds.), Studies in Applied Probability and Management Science. Stanford University Press, Standford, CA, 159–172.

Khouja, M. (1999). The single-period (news-vendor) problem: literature review and suggestions for future research. *Omega, Int. J. Mgmt. Sci*., 27, 537–553.

Kotler, P., & Armstrong, G. (2004). Principles of Marketing (10th Edition). Prentice Hall.

Lau, A. H. L. , & Lau, H. S. (1988). The newsboy problem with price-dependent demand distribution. *IIE Transactions*, 20(2), 168–175.

Mills, E. S. (1959). Uncertainty and price theory. *The Quarterly Journal of Economics*, 73(1), 116–130.

Ng., I. C. L., & Lee, K. S. (2008). Competition in the advanced sale of service capacity. *International Journal of Revenue Management.* 2 (4), 307–326.

Qin, Y., Wang, R., Vakharia, A. J., Chen, Y., & Seref, M. M. H. (2011). The newsvendor problem: Review and directions for future research, *European Journal of Operational Research* 213 (2), 361–374.

Shah, Y. K., & Jha, P. J. (1991). A single-period stochastic inventory model under the influence of marketing policies. *Journal of the Operational Research Society*, 42 (2), February, 173–176.

Wagner, H. M. (1995). Global sensitivity analysis. *Operations Research*, 43(6), 948–969.

Wang, C. X., Webster, S., & Zhang, S. (2014). Robust price-setting newsvendor model with interval market size and consumer willingness-to-pay. *International Journal of Production Economics*, 154, 100–112.

Wang, Y. Y., Tao, F., & Wang, J. C. (2017). Designing optimal and practical volume discounting contracts for the dominant retailer with information asymmetry. *International Transactions in Operational Research*, 24, 1405–1433.

Whitin, T.M. (1955). Inventory control and price theory. *Management Science*, 2(1), 61–68.

Zeithaml, V.A., Parasuraman, A., & Berry, L.L. (1985). Problems and strategies in services marketing. *Journal of Marketing*, 49 (2), Spring, 33–46.

Zeng, C. (2013). Optimal advance selling strategy under price commitment. *Pacific Economic Review*, 18 (2), Spring, 233–258.

Zhang, B. (2012). Capacity-constrained multiple-market price discrimination. *Computers & Operations Research*, 39, 105–111.

DECISION SCIENCES INSTITUTE

Engaging Millennials in Gig Economy: An Exploratory Study Using Direct Selling Association's National Survey on Independent Sales Representatives

Yulong Li
Simmons University
Email: yulong.li@simmons.edu

Ying Liao
East Carolina University
Email: liaoy17@ecu.edu

ABSTRACT

Based on data from DSA's 2018 National Salesforce survey, this study revealed a three sub-dimensional structure of the construct of engagement, different from findings in previous research. Additional cluster and discriminant analyses investigated factors related to gig workers' perceived work experience and their likelihood to stay with the same company.

KEYWORDS: Workforce engagement, Gig economy, Direct selling, Factor analysis, Discriminant analysis

INTRODUCTION

As organizations respond to the forces of globalization and technology innovation, they see the rise of "gig economy" and the career preference change of the Millennials, the generation that is "notoriously" known for their zest for entrepreneurship and more frequent job-hopping than any other generations before (Deal and Levenson, 2016). According to a recent employer survey by Ernst & Young, the global workforce has been transformed dramatically: one in two organizations in the U.S. indicated an increase in the use of contingent workforce in the last five years (EY Global, 2018). Intuit further predicted that 40 percent of American workers would be independent contractors by 2020 (Intuit, 2018).

The gig economy is defined as "a collection of markets that match providers to consumers on a gig (or job) basis, where gig workers enter into formal agreements with on-demand companies to provide service to the companies' clients" (Donovan, Bradley and Shimabukuru, 2016). For organizations, maintaining a highly flexible workforce gives them the agility to scale and descale resources rapidly depending on dynamic market demands. For contingent workers, also known as consultants, freelancers, independent contractors, independent professionals, or temps in various occasions, the gig economy offers them freedom, instant gratification, bridge employment opportunities and work-life balance that many Millennials rank highly on their career aspiration list (PWC, 2017). For researchers on the other hand, such drastic paradigm shift in employment context and structure challenges many fundamental beliefs and principles of workforce planning and management, thus calls for a re-examination of the applicability of many traditional theories and frameworks on employee management among Millennial in the new economy. Among them, employee engagement, a concept regularly used to attribute job performance to employee morale, motivation, and behavior, deserves exceptional research attention. This is because of three reasons: firstly, employee engagement has been documented in numerous previous studies in the literature as one of the most important enablers for potentially high organizational performances. Secondly, the alternative job arrangements and the temporary nature of contingent work significantly alter workers' sense of job security, work scheduling, and compensation structure. The bonds that customarily ties workers to the organization and to his or her bosses, colleagues and sometimes clients are thus believed to become much weaker than in a typical full-time employment environment, in which traditional employee engagement theories were developed (Donovan, Bradley and Shimabukuru, 2016). Organizations, those utilize a large number of contingent employees in particular, must develop new understandings on the challenges and possible strategies of engaging the new labor force population of Millennials in the new business landscape of the gig economy. Thirdly, the sheer size of the Millennial population, which will account for up to 75 percent of the entire labor force in the U.S by 2025 (Fry, 2018) and the impact and nationwide reach of gig work have even attracted attention from members of the U.S. Congress (Donovan, Bradley and Shimabukuru, 2016). All these three reasons combined make this current study interesting to explore, relevant to business practices, and impactful from a national policy perspective.

In the rest of this paper, we will first start with a literature review on employee engagement as a construct in section 2. In the next section on research methodology, we will conduct exploratory factor analysis using the data from the 2018 Direct Selling Association's National Salesforce Study Questionnaire to examine the conceptual dimensions of engagement among freelance contractors (referred as direct selling independent sales representatives or ISR in the direct selling industry). After that, we will conduct K-means cluster analysis and a series of discriminant analyses to evaluate various factors that are believed to have impact on ISR's work experiences and their intention to continue with their current sponsoring direct selling companies. The last section (#4) will discuss the implications of our findings and draw conclusions for this study.

LITERATURE REVIEW ON EMPLOYEE ENGAGEMENT

Built upon job characteristics theory (Hackman and Oldham, 1980) and growth theory (Alderfer, 1972), scholars in academic circles conceptually define *Employee Engagement* as the "harnessing of the organizational members' selves in their work roles that enables them to devote discretionary effort in achieving organizational intent*" (*Kahn, 1990, p 694). Engagement is a motivational state that employees develop a sense of meaningfulness (i.e. feeling work is valuable and worth doing), availability (i.e. having the physical and mental energy for work), and psychological safety (i.e. having the freedom of being oneself at work). Despite its preciseness and ability to delineate engagement from other similar but distinctive constructs (e.g. employee participation and empowerment), the academic definition lacks the level of clarity for measurement and practical application (Podskoff, Mackenzie, and Podskoff, 2016). Most researchers typically chose to define engagement by the characteristics or outcomes of engaged employees. For example, Sugirtha and Snehe (2017) argue that engagement makes employees feel cognitively and emotionally connected to their jobs, managers, co-workers and organization's vision, mission and objectives. Mathews (2010) describes engaged employees as those who "take pride in their organization and work; take ownership of projects; talk positively about themselves, their employer and the goods and services they help deliver; and view working for their organization as a career, not just a job" (p29).

Regardless of which way it is defined, engagement is universally perceived as a key enabler for potentially high-performing organizations, because engaged employees are more likely to go extra miles in their assignments, contribute vigorously and thus achieve greater productivity, higher customer satisfaction, and increase profitability of their employer. Numerous studies have demonstrated strong relationships between engagement and employee turnover (e.g. Shuck et al, 2014; *Smith & Macko, 2014*). More recently, additional studies have empirically observed the positive effects of engagement on reducing turnover intentions in specific industry sectors, such as among federal employees (Byrne, Hayes, and Holcombe, 2017), in public accounting firms (Carrillo et al., 2017), IT sector (Pandita and Singhal, 2017), and higher education (Chong and Lee, 2017).

A variety of approaches have been proposed in the literature on measuring engagement as a complex and holistic construct. The Gallup's Q^{12} Survey is one of the earliest and also commonly considered as one of the most influential. It consists of 12 simple one-sentence questions (Table 1) and has been used on more than 25 million employees in 3 million workgroups from 195 countries since 1996 (Gallup, 2018). Forbringer (2002) consequently categorized the Gallup Q^{12} into 4 discrete levels/dimensions: 1) basic job needs, 2) self-image and worth, 3) feeling of belongingness, and 4) opportunities for learning and growth. More recently, SurveyMonkey Business Solutions (2016) designed a different version of 12 questions (Table 1). As compared with the Gallup Q^{12}, which primarily focus on employees' perception about themselves, the SurveyMonkey questionnaire measures engagement at the organizational level. Besides the 6 questions about how engaged one feel about him or herself, it includes another 6 questions on the respondent's perception on co-workers' level of engagement. The assumption is that an engaged organization is one where all employees should feel engaged (rather than just felt by the individual who responded to the questionnaire). Following the same two-level approach (i.e. individual plus organization), the Society of Human Resource Management (SHRM) measures engagement in its Employee Job Satisfaction and Engagement survey (2016) that are administered to 600 U.S. employees annually. The engagement section contains 37 questions (Table 1), categorized into 3 separate dimensions: 1) *Conditions for Engagement* dimension captures the employee's satisfaction level to a wide array of job characteristics, which are considered as the capacity and reason for the employee to feel engaged. 2) The *Engagement Opinion* dimension depicts the personal reflection of employee's relationship with their own work. 3) The *Engagement Behaviors* dimension describes the "look" of engaged employees and workgroups and their "actions" that can positively contribute to the welfare of the organization. Despite the greater details and more straightforward verbiage of the questions, the foundation of the SHRM survey is similar to the 4 levels of Gallup Q^{12} as proposed by Forbringer (2002).

The applicability of traditional concept of employee engagement among Millennials in the gig economy has been questioned by many practitioners and researchers. Typical concerns are around: 1) the confusing definition of gig workers being employees vs. independent contractors and the associated legal rights, responsibilities and other ramifications (Donovan, Bradley and Shimabukuru, 2016); 2) the weaker bond and loyalty from contingent workers to employers (Klazema, 2018); 3) the typical part-time and temporary status of gig workers and their and spatial detachment (i.e. the lack of face-to-face interactions) with fellow employees and the geographic location of the company (ibid); 4) the shrinking differences between self-employment and external work in the gig economy so that companies have to compete for talents not only against other employers, but also with self-employment (McGuire, 2017); and 5) company's lack of understanding to the importance of gig worker engagement and the appropriate strategies to engage them (Thompson, 2017). We would thus like to use the secondary data from the 2018 Direct Selling Association's National Salesforce Study Questionnaire to help us develop a snapshot understanding to the current status of ISR engagement and compare its internal sub-dimensions with the traditionally defined construct of employee engagement.

Table 1. Literature review on the construct of Employee Engagement

Forbringer's Engagement levels	Gallup's Q[12] Survey	Survey Monkey's Professional Engagement Survey		SHRM's Employee Job Satisfaction and Engagement Study	
Level 1: Basic Needs (What do I get?) • Gallup Q1 • Gallup Q2	Q1. Do you know what is expected of you at work?	Self-perception	Q1. I am inspired to meet my goals at work.	Conditions for Engagement	• Relationship with co-workers • Opportunities to use skills/abilities • Meaningfulness of the job • The work itself • Relationship with immediate supervisor • Organization's financial stability • Contribution of work to organization's business goals • Autonomy and independence • Variety of work • Overall corporate culture • Communication between employee and senior management • Organization's commitment to corporate social responsibility • Management's recognition of employee performance • Job-specific training • Organization's commitment to professional development • Network opportunities • Career development opportunities • Career advancement opportunities within the organization
	Q2. Do you have the materials and equipment to do your work right?		Q2. I feel completely involved in my work		
Level 2: Self-image and worth (What do I give?) • Gallup Q3 • Gallup Q4 • Gallup Q5 • Gallup Q6	Q3. At work, do you have the opportunity to do what you do best every day?		Q3. I get excited about going to work.		
	Q4. In the last seven days, have you received recognition or praise for doing good work?		Q4. I am often so involved in my work that they day goes by very quickly.		
	Q5. Does your supervisor, or someone at work, seem to care about you as a person?		Q5. I am determined to give my best effort at work each day.		
	Q6. Is there someone at work who encourages your development?		Q6. When at work, I am completely focused on my job duties.		
Level 3: Feeling of belongingness (Do I belong here?) • Gallup Q7 • Gallup Q8 • Gallup Q9 • Gallup Q10	Q7. At work, do your opinions seem to count?	Perception of the organization / other employees	Q7. In my organization, employees adapt quickly to difficult situations.	Engagement Opinion	• I am confident to meet work goals • I am determined to accomplish work goals • I understand the organization's vision/mission • I am highly motivated by my work goals • While at work, I always completely focus on work projects • I am provided the resources to do job well. • I frequently feel that I am putting all efforts to my work • I have passion and excitement about my work • I am often so wrapped up in work that hours go by fast • I enjoy volunteering for activities beyond my job requirements • I feel completely plugged in at work, like I am always on full power.
	Q8. Does the mission/purpose of your company make you feel your job is important?		Q8. Employees here always keep going when the going gets tough.		
	Q9. Are your associates (fellow employees) committed to doing quality work?		Q9. Employees proactively identify future challenges and opportunities.		
	Q10. Do you have a best friend at work?		Q10. Employees in my organization take the initiative to help other employees when the need arises.	En ga ge	• Employees are encouraged to take actions when problem/opportunity arises
Level 4: Learning, growth and innovation (How can we all grow?) • Gallup Q11 • Gallup Q12	Q11. In the last six months, has someone at work talked to you about your progress?		Q11. Employees here are willing to take on new tasks as needed.		• My colleagues quickly adopt to challenging or crisis situations. • My workgroup never gives up despite difficulties • Employee deal very week with unpredictable or changing work situations. • My workgroup is constantly looking out for future challenges • People in my workgroup are always flexible • Others in my organization view unexpected responsibilities as opportunities • Other people in my organization often volunteer for new projects
	Q12. In the last year, have you had opportunities to learn and grow?		Q12. Employees in my organization willingly accept change.		

RESEARCH METHODOLOGY AND FINDINGS

Research Sample and Data

Direct Selling Association (DSA) is the national trade association for companies using direct selling business model. Serving nearly 200 member companies, DSA provides educational opportunities and support to advance the understanding and practice of direct selling. Every 3-5 years, DSA conducts a study on current direct selling salesforce to learn about their attitudes and opinions toward direct selling and outline demographic profile of workforce in the industry. The study may be completed in English or Spanish. As faculty fellows of the Direct Selling Education Foundation (DSEF), a 501(c)3 organization affiliated with DSA to advance direct selling-related research and teaching, we were granted access to the variable descriptions of the questionnaire and approximately 8,700 responses DSA had received from this study conducted in May 2018.

We first removed records 1) with incomplete data, 2) answered “Don’t know/Not Applicable” option in measurement items of ISR engagement, and 3) indicating company and upline “does not provide” training and support. In screening the data for outliers, we detected extreme values using Mahalanobis distance. As wide range of commitment to direct selling was represented in the data set, we further identified those have no other paid jobs and spend at least 10 hours per week on direct selling business. In addition, it has been recognized that engagement for success is different across cultures in arena of management and marketing (Akutsu, Haga, Fujikawa and Ono, 2013; Huang, Ma and Meng, 2018). To control the effect of ethnic culture in this study, we chose respondents whose primary language spoken at home is English. The resulted final research sample contains 267 responses.

Measurement and Analysis

To investigate the number of sub-constructs and structure of the measurement of engagement, an exploratory factor analysis (EFA) was conducted using SPSS on the 33 items about ISR's opinions on the direct selling company they represent and the upline agents who sponsor the ISR. Using principal component analysis, varimax rotation and extraction criterion of eigenvalue greater than 1.00, the analysis results recommended a five-factor solution that explained 67.78% of the variance. All measurement items with their factor loadings and reliability are reported in Table 2. Cross-loaded items were examined using the threshold factor loading value of 0.50, underlying structure in response to proposed construct and the impact on measurement item reliability (Cronbach's α). For example, if item # 10 (The sponsor and/or upline is honest and accurate about the effort required in direct selling and the rewards you can expect to receive) is cross-loaded on the engagement by upline and the basic engagement condition at corporate level, with factor loading of 0.62 and 0.47, respectively. The cross-loading could be because being "accurate" about effort required and potential reward makes ISRs feel that they know what to expect at work (the condition for engagement); while being "honest" manifests upline's willingness to enhance sense of ISRs' well-being and help them succeed. As the item reflects two levels of engagement and the Cronbach's α is not improved if the item is added, it was removed. Factors 4 and 5 revealed in the EFA were also removed because they obviously manifest inventory-related operations not engagement.

We thus eventually identified three sub-dimensions of ISR engagement as shown in table 2: 1) UPLINE: Emotional attachment with sponsor/upline (items 1-9); 2) OPR: Operational needs from company (items 13-20); and 3) BELONG: Feeling belongingness to the company (items 21-26). The cumulative variances for these three factors are observed to be 56.47%.

Cluster Analysis

To study the ISR's perception towards direct selling experience, we performed a K-means cluster analysis choosing experience comparing to initial expectation and the rating of actual experience in direct selling as clustering variables. It needs to be noted that the responses for initial expectation (on horizontal axis) was in a 1-3 scale; while the actual experience (on vertical axis) was in a 1-5 scale as shown in Figure 2.

The results of the cluster analysis revealed two groups with 82.5% and 17.5% share of the sample, respectively. As shown in Table 3, ISRs in Cluster 1 tend to have very good experience that above their expectations when they started direct selling. ISRs in Cluster 2 have fair to good experience but the experience doesn't meet their expectations in general.

Table 2. Construct items (Factor loadings and reliability)

Item	Factor Loading 1	2	3	4	5	Cronbach's α
How much do you agree that your sponsor and/or upline…?						0.94
1. Is interested in your personal development	**0.87**	0.20	0.20	0.03	0.04	
2. Has recognized your accomplishments appropriately	**0.86**	0.17	0.22	-0.01	0.02	
3. Values you	**0.85**	0.19	0.14	0.03	0.05	
4. Helps you to set and achieve challenging goals	**0.84**	0.18	0.22	0.05	0.03	
5. Provides an appropriate amount of training tools and information for you to succeed	**0.83**	0.15	0.08	0.09	0.01	
6. Is good at anticipating the needs of independent direct sellers	**0.81**	0.15	0.29	0.10	0.06	
7. Has provided the means for you to develop more friendships	**0.80**	0.13	0.28	0.06	0.15	
8. Educates you about avoiding sales practices that are deceptive, unlawful or unethical	**0.79**	0.20	0.21	0.14	0.14	
9. Has provided you with the chance to be more involved in your community	**0.76**	0.09	0.26	0.19	0.15	
*10. Is honest and accurate about the effort required in direct selling and the rewards you can expect to receive	0.62	0.47	0.25	0.11	0.21	
* 11. Allows you to make your own business decisions freely	0.58	0.46	0.03	0.06	0.11	
* 12. Is very clear about the full costs of your orders, including any shipping and service charges	0.56	0.48	0.18	0.08	0.29	
How much do you agree that the company…?						0.81
13. Is very clear about the full costs of your orders, including any shipping and service charges	0.17	**0.73**	0.16	0.00	0.08	
14. Is honest and accurate about the effort required in direct selling and the rewards you can expect to receive	0.24	**0.73**	0.33	0.18	0.01	
15. Gives you a way to earn a fair return on your time and effort	0.16	**0.71**	0.26	0.14	0.06	
16. Educates you about avoiding sales practices that are deceptive, unlawful or unethical	0.19	**0.67**	0.26	0.21	0.12	
17. Provides an easy-to-use ordering system	0.10	**0.57**	0.24	0.05	0.20	
18. Allows you to make your own business decisions freely	0.21	**0.51**	0.31	-0.03	0.06	
19. Is good at anticipating the needs of independent direct sellers	0.24	**0.52**	0.52	0.20	0.04	
20. Has competent corporate personnel that work with you	0.26	**0.52**	0.50	0.00	0.05	
						0.89
21. Is interested in your personal development	0.25	0.26	**0.78**	0.13	0.02	
22. Values you	0.23	0.39	**0.73**	0.06	-0.03	
23. Has recognized your accomplishments appropriately	0.29	0.27	**0.68**	0.05	0.07	
24. Has provided you with the chance to be more involved in your community	0.19	0.22	**0.61**	0.32	0.16	
25. Has provided the means for you to develop more friendships	0.19	0.18	**0.61**	0.18	0.39	

26. Helps you to set and achieve challenging goals	0.30	0.48	**0.55**	0.17	0.04
*27. Makes you welcome in a new country or location	0.24	0.48	0.46	0.26	0.14
*28. Gives you a way to earn a fair return on your time and effort	0.27	0.42	0.57	0.10	0.15
*29. Has materials that are truthful in describing product quality, value and performance	0.16	0.44	0.49	0.11	0.21
*30. Will buy back your unexpired or undamaged inventory if you request it	0.14	0.20	0.15	0.88	0.12
*31. Will buy back product inventory, sales aids and materials, including the starter kit, if you decide to leave the business	0.12	0.12	0.29	0.86	0.08
*32. Requires or encourages you to purchase only a reasonable amount of inventory	0.05	0.20	0.14	0.12	0.83
*33. Provides an appropriate amount of training tools and information for you to succeed	0.49	0.17	0.13	0.09	0.64

Retained items are denoted in bold font.
* Item removed

Figure 1. Conceptual framework of the cluster analysis

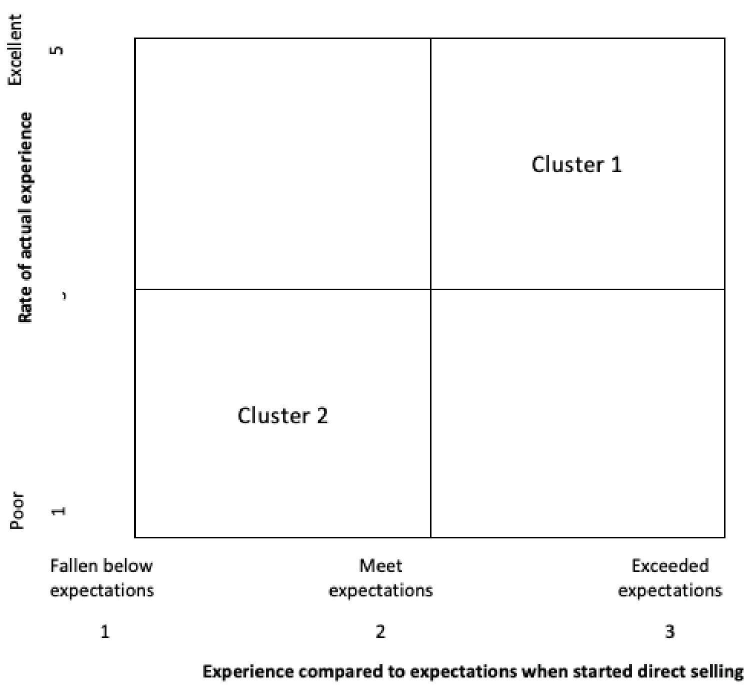


Table 3. Results of cluster analysis

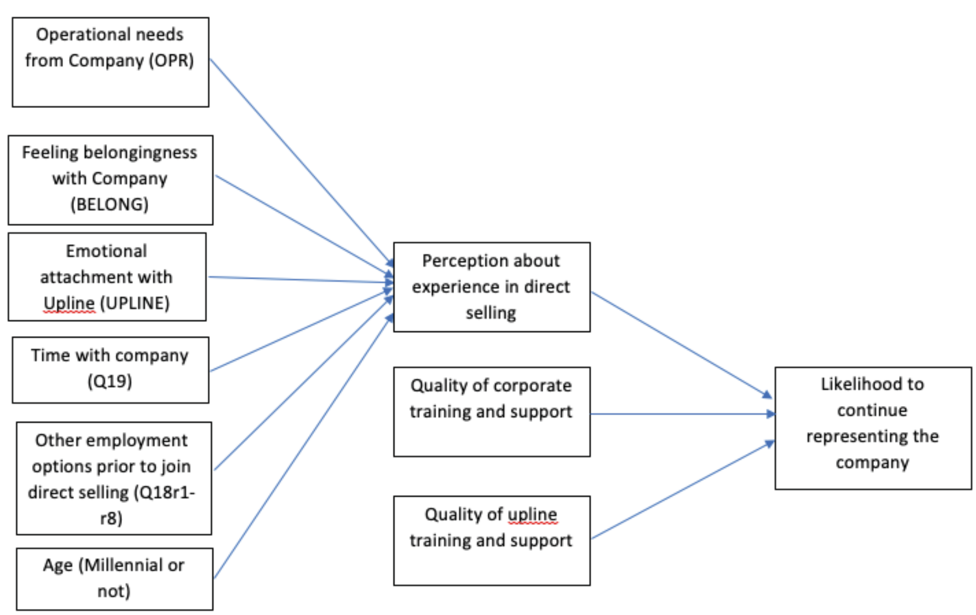
		Cluster 1	Cluster 2	F
Experience comparing to initial expectation	Mean	2.79	1.82 (out of 3)	163.94***
	S.D.	0.41	0.66	
Actual experience in direct selling	Mean	4.64	2.95 (out of 5	407.6***
	S.D.	0.48	0.61	

*** p < 0.001

<u>Discriminant Analysis</u>

Based upon cluster analysis results, we conducted discriminant analysis to further investigate the effectiveness of engagement sub-dimensions and a host of demographic descriptors on ISR's perceptions about their experiences in direct selling. These descriptors include a) time with the company, b) ISR age (millennial vs. baby boomer) and c) a series of 8 employment options respondents had explored when deciding to join direct selling (labeled as Q18_n in table 4). The conceptual model was illustrated in Figure 2. As the analysis involved two groups of ISR, one function was derived and was significant at 0.01 level ($\lambda=0.868$, $\chi^2(df=13) = 34.43$, $p = 0.001$). The hit ratio was 84.9%, which was higher than 1.25 times that achieved by chance, suggested acceptable classification accuracy (Hair et al., 2010. Press's Q is another measure of predictive accuracy of a discriminant function. The value was calculated at 122.92 and significant at 0.01 level. All above indicates that the discriminant function was statistically significant.

Figure 2. Conceptual model



As shown in Table 4, the Univariate ANOVA assesses the significance between the means of the two groups, variables of engagement sub-dimension of feeling belongingness to the company (BELONG), time with company, age and employment option before deciding to become an ISR showed significant univariate differences between the two groups at 0.05 or 0.01 level. The other aspect of engagement (Operational needs from company or OPR) was significant at the 0.1 level. It is noteworthy that Emotional attachment with upline (UPLINE) was not significant (not reported in Table 4). In other words, ISRs who had high levels of engagement with their upline sponsors rated their perception of direct selling experience very similarly to those with low upline engagement. Hair et al. (2010) suggested discriminant loading as a valid parameter to interpret the discriminating power of an independent variable. To be consistent with previous research, we consider those with loadings absolute value of 0.3 or above as substantive in this study. The results thus indicated that those ISR who had considered other full-time employment options are most likely to have a better experience with direct selling, while those who had not made no other career plans would likely to have a bad experience in direct selling (a strong negative discriminant loading factor). Feeling belongingness to the company (BELONG) also appeared to have strong relationship with ISR experiences, so did their time with that company and being the Millennials generation.

Finally, we conducted an additional discriminant analysis to determine whether three variables - perception about experience in direct selling, quality of training and support provided by the company and upline, could anyhow predict an ISR's likelihood to stay with the company in direct selling (Figure 2). Although the data consists of four level of likelihood to continue representing the company for the next year. The sample size of those who were very unlikely or somewhat unlikely to continue were only 4 and 1, respectively. Therefore, the discriminant analysis was only performed on two groups of ISRs – those were somewhat likely or very likely to continue. All three variables - perception on experience, quality of training and support the company provides, quality of training and support the upline or sponsor provides, are primary source of differentiating the two groups (Table 6). The ISRs who obtained high quality training and support from upline and company were more likely to continue with the company (Table 5). The groups of ISRs with relatively good and poor perception about experience were automatically coded in reverse order as 1 and 2 in cluster analysis. The results demonstrated that ISRs who had better perception about their direct selling experience were also more likely to continue.

DISCUSSIONS AND CONCLUSION

This study has made the following contributions to practitioners and future researchers.

Firstly, our empirical analysis confirmed the uniqueness of millennial ISRs engagement in the gig economy. The factor analysis revealed three distinctive sub-dimensions of engagement that are very different from findings of earlier studies on this construct. The dimension of emotional attachment with one's sponsor or upline (UPLINE) recognizes the importance of the people element of the construct. In order to feel engaged with one's work, people need to develop connections and emotional involvement with colleagues at their workplace (Kapoor & Meachem, 2012). With the absence of typical coworkers and managers in the direct selling context, the upline sponsors become the "human figurehead" of the work that ISRs can relate to. Direct selling companies are thus advised to pay close attention to the role of these upline sponsors in developing ISR engagement. The other two company related engagement dimensions (OPR and BELONG) align relatively well with Maslow's need hierarchy. At the minimum level, the direct selling company must meet basic operational needs of ISRs so that they can function properly for their contracted work (OPR). For many gig workers who use their own equipment and facilities for their work (e.g. Uber drivers use their private vehicles and direct selling ISRs

Table 4. Discriminant Analysis on Experience in Direct Selling

	Discriminant loading	Standardized Function coefficient	F
Full time employment option before deciding to join direct selling (Q18_1)	0.639	0.425	15.491***
Nothing before deciding to join direct selling (Q18_8)	-0.549	-0.385	11.441***
Feeling belongingness with company (BELONG)	0.424	0.546	6.811**
Time with company	0.323	0.431	3.956**
(Age) Millennials	0.320	0.396	3.887**
Operational needs from company (OPR)	0.274	0.348	2.859*
Emotional attachment with upline sponsors	0.209	0.263	ns
Part time employment option (8-29 hours per week) before deciding to join direct selling (Q18-2)	0.006	0.052	ns
Part time employment option (under 8 hours per week) before deciding to join direct selling (Q18_3)	-0.065	-0.198	ns
Other type of self-employment before deciding to join direct selling (Q18_4)	0.245	0.117	ns
Further education (Q18_5)	0.140	0.009	ns
Collaborative economy earnings opportunities (e.g. Uber, Airbnb, etc.) before deciding to join direct selling (Q18_6)	0.083	0.183	ns
Selling products through an online marketplace (e.g. eBay, Craigslist etc.) before deciding to join direct selling (Q18_7)	0.032	0.114	ns

*p<0.1, **p<0.05, ***p<0.01

set up sales parties at their own home), good quality products, clearly defined work flow, and easy-to-work-with supportive representatives from company are particular important, because these equipment and facilities are less standardized and more likely to cause confusion in the gig work environment. Once the basic needs have been met, gig companies should also look at the dimension of feeling belongingness to the company (BELONG) to engage their workforce. This dimension allows the company to move up the Maslow's hierarchy and address the needs of gig workers for personal development opportunities and the sense of ownership and pride. Meeting these higher-level needs helps gig workers align personal goals with the organizational goals, policies and practices, and is believed to be able to sustain a high level of engagement.

Table 5. Descriptive statistics for group statistics

Likelihood to continue		Mean	S.D.
Somewhat likely	Perception on experience	1.64	0.505
	Quality of training and support the company provides	4.36	0.674
	Quality of training and support the upline or sponsor provides	3.91	1.136
Very likely	Perception on experience	1.15	0.356
	Quality of training and support the company provides	4.89	0.352
	Quality of training and support the upline or sponsor provides	4.58	0.834

Table 6. Discriminant analysis on likelihood to continue representing the company

	Correlation Coefficient	Standardized Function coefficient	F
Quality of training and support the company provides	0.759	0.640	21.168***
Perception on experience	-0.718	-0.632	18.957***
Quality of training and support the upline or sponsor provides	0.426	0.140	6.671**

*p<0.1, **p<0.05, ***p<0.01

Secondly, our cluster analysis and discriminant analysis aim at understanding the factors related to gig workers' job experiences and their longevity in the gig professions. It is interesting to notice that being Millennial is indeed strongly related to having a good work experience in the gig economy. This is a confirmation to the findings from many generational studies in the literature. It is fortunate for the gig companies that the nature of gig work does seem to be perceived favorably among Millennials. They should continue their expansion in the Millennial generation.

Thirdly, another interesting finding from table 4 indicates that those people who have considered other full-time job opportunities and those who have considered no other career options (and jump straight into direct selling) are the ones more likely to have better perceived work experience at a later time. While others who have considered part-time employment opportunities, any type of self-employment, or those only consider direct selling as an "educational opportunity" prior to joining to direct selling are not found to have statistically significant relationship with high work experience. One possible interpretation to this result is that gig companies may want to rethink the strategy of billing themselves as a sector that offers people bridge employment or supplement income when attracting new contractors, because people with these expectations may find the actual gig work not meeting their expectations and drop out quickly. Only those who are serious with the gig work and would like to regard it as a long-term career path are more likely to feel engaged, enjoy their work experience, and deliver quality work results to the customers. When studying ISR's likelihood to stay with the company in direct selling, our results in table 6 indicated the importance of gig worker's perceived experience at work and the training and support they receive from both upline sponsors and the direct selling company. It thus highlighted another strategy for companies to sustain their gig workers: proper on-boarding at the commencement of contract work and ongoing training should never take lightly. Gig workers are still workers of your organization after all- they should have the same level of training and support as serious as you provide to any other employees.

Finally, we would like to acknowledge the challenges related to working with secondary data. The questionnaire was not designed specifically for this study nor even created/administered by the authors. The wording of some items was not as clear, the conceptual thoroughness of some questions less than ideal, and scales of different questions were not always constant. All these resulted in a lot of missing data and some difficulties for our model fitting. Despite these challenges, we appreciate the opportunity given by the DSEF. We hope our findings have stimulated more interest in studying Millennials in the burgeoning gig economy.

REFERENCES

References available upon request

DECISION SCIENCES INSTITUTE

Using Auditing and Commitment for Engaging the Supplier for Supply Chain Sustainability

Hossein Rikhtehgar Berenji
College of Business, Pacific University
hossein@pacificu.edu

Nagesh N. Murthy
Lundquist College of Business, University of Oregon
nmurthy@uoregon.edu

Zhibin (Ben) Yang
Lundquist College of Business, University of Oregon
zyang@uoregon.edu

ABSTRACT

We model a supply chain in which the buyer audits the supplier's compliance with the code of conduct. We investigate the effect of buyer's upfront commitment to price and quantity on supplier's compliance. We also analyze implications of raising the standard of code of conduct on sustainability and financial performance.

KEYWORDS: Supplier's Social and Environmental Compliance, Sustainability Code of Conduct, Commitment to Contract Terms, and Game Theory

INTRODUCTION

Major brands in many industries outsource their production and hence lose direct control in managing sustainability in their supplier base. An overwhelming majority (i.e., as much as 70%) of sustainability issues (i.e., social and environmental violations) come from suppliers (UNGC 2015). Brands strive to reduce the likelihood of social and environmental violations in their supplier base. Many brands such as Nike, GAP, Wal-Mart and Apple have developed a code of conduct that sets expectations for their suppliers' social and environmental practices. These brands audit their suppliers' compliance (GIIRS 2014).

Yet suppliers' likelihood of compliance varies significantly across industries with significant consequences for the buyer. For example, a supplier of multiple fashion brands (i.e., GAP, Old Navy, and Banana Republic), released untreated wastewater into local streams and rivers in China that resulted in significant brand damage (Price 2017, Backhaus and Fischer 2017). In another incident, a supplier of Nike employed child labor and paid low wages. Consequently, the customers organized boycotts, which resulted in loss of revenue for Nike (Nisen 2013).

Brands face the challenge of engaging the supplier for improving social and environmental sustainability. Previous researches have investigated the efficacy of using various tools such as offering price premium (see Karaer et al. 2017), auditing (see Plambeck and Taylor 2016), and certification (see Chen and Lee 2016) to motivate suppliers to be more socially and environmentally responsible. We investigate the efficacy of a brands' use of commitment to contract terms in combination with auditing (i.e., committing to price and/or quantity prior to auditing) to enhance supplier's compliance to the code of conduct. We address the following research questions:

1) How does commitment affect the compliance to the code of conduct?
2) How does commitment affect the financial performance of buyer and supplier?
3) What is the relative value of different modes of commitment (i.e., price only, quantity only, or both, or neither)?
4) What are the implications of raising the standard of code of conduct on sustainability and financial performance?

MODEL

We develop and analyze multistage game-theoretic models. We consider a supply chain with a buyer and a supplier. The buyer sources a product from the supplier with a wholesale price contract and sells the product to the market. The supplier may misbehave socially or environmentally and cause the buyer to lose market revenue. The buyer audits the supplier with respect to the code of conduct. The supplier exerts an effort to comply with the code of conduct. To induce higher level of effort of the supplier, the buyer may commit to some or all the contract terms (i.e., price, quantity) before auditing.

RESULTS AND IMPLICATIONS:

We solve for the subgame-perfect equilibrium and characterize the contract terms, the buyer's auditing effort, and the supplier's compliance effort at the equilibrium. We also derive the overall compliance of the supply chain and the two parties' profits. To the first research question, our results confirm that increasing the level of commitment improves the supplier's compliance. Partial commitment to the price may decrease the overall compliance of the supply chain in comparison to the scenario of no commitment.

To the second research question, we find that increasing the level of commitment always increases the buyer's profit, but the supplier's profit may decrease. Interestingly, we find that when the buyer commits to price only, there exists a win-win-win outcome where both parties will be financially better off and overall compliance of the supply chain is improved, in comparison to the scenario of no commitment.

To the third research question, we find that there is a value for the buyer to commit to quantity only if the buyer has already committed to the price. Therefore, committing to the price and committing to the quantity are complementary to each other for the buyer. The supplier suffers a loss from buyer's commitment to quantity only if the buyer also commits to the price. Therefore, committing to price and committing to quantity are substitutes for the supplier. Also, we find that the change in the overall compliance of supply chain due to committing to quantity is positive only if the buyer also commits to the price.

To the last research question, our results show that when the buyer raises the standard of the code of conduct, interestingly, the supplier is more likely to be compliant with the code of conduct, even if the chance of passing the audit decreases. We show that when the auditing cost is very sensitive to increasing standard and the buyer commits to the price, an increases of the standard increases both the supplier and buyer's profit. We also find that as the standard of the code of conduct increases, the supply chain's total compliance and total profit may decrease.

REFERENCES

Backhaus M, Fischer M (2017) “Brand damage from product harm and corporate social irresponsibility: How deep and how long?” *Working paper*, University of Cologne.

Chen L, Lee HL (2016) “Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment” *Management Science* 63(9).

GIIRS (2014) “Resource guide: Creating a supplier code of conduct” ww.bcorporation.net, https://www.bcorporation.net/sites/default/files/documents/bestpractices/EM_Creating_a_Supplier_Code_of_Conduct.pdf.

Karaer O, Kraft T, Khawam J (2017) “Buyer and nonprofit levers to improve supplier environmental performance” *Production and Operations Management* 26(6).

Nisen M (2013) “How Nike solved its sweatshop problem” *Business Insider* http://www.Businessinsider.com/how-nike-solved-its’s-sweatshop-problem-2013-5.

Plambeck EL, Taylor TA (2016) “Supplier evasion of a buyer’s audit: Implications for motivating supplier social and environmental responsibility” *Manufacturing & Service Operations Management* 18(2).

Price D (2017) “Worse for wear: Indonesia’s textile boom” *Undark- Special Project*.

UNGC - United Nation Global Compact (2015), “Supply Chain Sustainability - A Practical Guide for Continuous Improvement, Second Edition”.

THE DECISION OF THE SCIENCE INSTITUTE

NLP analysis of incidents descriptions in case of IT services

Attila Sóti Széchenyi István University of Hungary-soti.attial@hallgato.sze.hu

ABSTRACT

This article is an empirical approach to the linguistic analysis of incident description using natural language processing (NLP). Considering the customer feedback as texts are mainly written in an unstructured way, there is a need to use text meaning techniques to gain insight or perceive the focusing area of the text. The case is more complicated when events or incidents description is associated with a large organization providing a wide range of IT services. This article describes how to use semantic analysis with other data meaning techniques which can help in finding focus, patterns and trends in texts connected to user feedback.

Keywords: Statistical Natural Language Processing, User feedback, IT services

INTRODUCTION

It is estimated that 80% of business-relevant information is in unstructured and semi-structured text data. In other words, without the use of textual analysis, to discover the 80 percent of data, all business information and customer's behaviour data will be lost (Halevy, Norvig, & Pereira, 2009).

Incident descriptions are usually written by non-professionals and many times they are not formulated adequate to English, but they contain a lot of information.(Pan, Liu, Xiang, & Yang, 2011) Using this information in a thorough and systematic way, it is increasingly necessary to understand customer's behaviour and attitudes (Song & He, 2010). For example, if the cause of most incidents is known, measures can be taken to decrease their occurrences. There are no standard rules for writing customer's feedback so that the computer can understand them. The language and the meaning for every piece of text vary depending on the purpose (Dan Jurafsky & Martin, 2009). The only way to include accurately unstructured data in a data-meaning project is to understand the language and the context within which the text is created (Briscoe, 2011).

LITERATURE REVIEW

Understanding human language is based on linguistics, commonly referred to as Natural Language Processing (NLP) (Nadkarni, Ohno-Machado, & Chapman, 2011). NLP is a way for computers to analyse, understand, and derive meaning from human language in a smart and useful way (Turney & Pantel, 2010). By utilizing NLP, developers can organize and structure the knowledge to perform tasks such as automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, and topic segmentation. (Briscoe & Carroll, 1993) NLP algorithms are typically based on machine learning algorithms (Reddy et al., 2015). Instead of hand-coding, large sets of rules, NLP can rely on machine learning to automatic learning these rules by analysing a set of (Chan & Franklin, 2011). A Statistical NLP approach seeks to solve problems by automatic learning lexical and structural preferences (Cohen & Dolbey, 2007) from corpora and if the existing dictionary is extended (rules, types, synonyms) to support this process we might be able to analyse our non-formal descriptions. (Majumder, Mitra, & Chaudhuri, 2002) (See Figures below 1, 2) as examples (Jordan & Mitchell, 2015). A system that incorporates NLP can intelligently extract terms, including compound phrases, and permit classification of terms into related groups (Le & David Jeong, 2017). Linguistic systems are knowledge-sensitive: the more information is contained in the linguistic resources (dictionaries), the higher is the quality of results (Yu, Li, Merigó, & Fang, 2016). Modification of the dictionary content, such as synonym definitions, can simplify the resulting information and focus attention on the most relevant concepts.

Figure 1 – Statistical NLP approach phase I

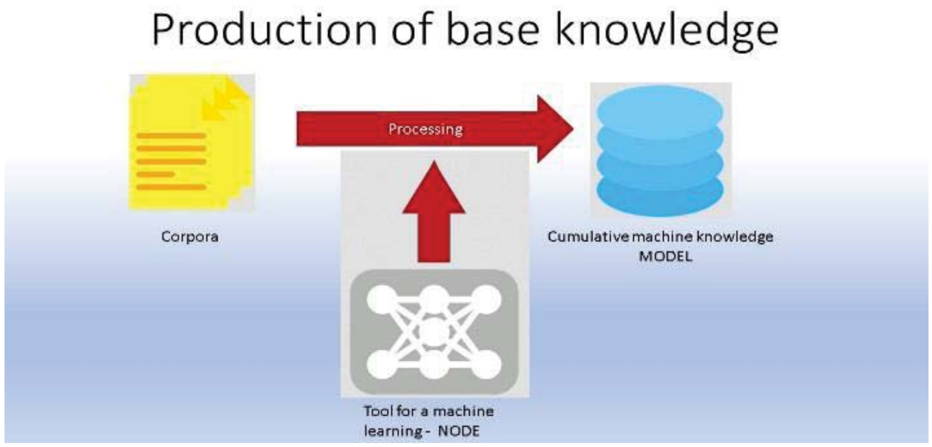
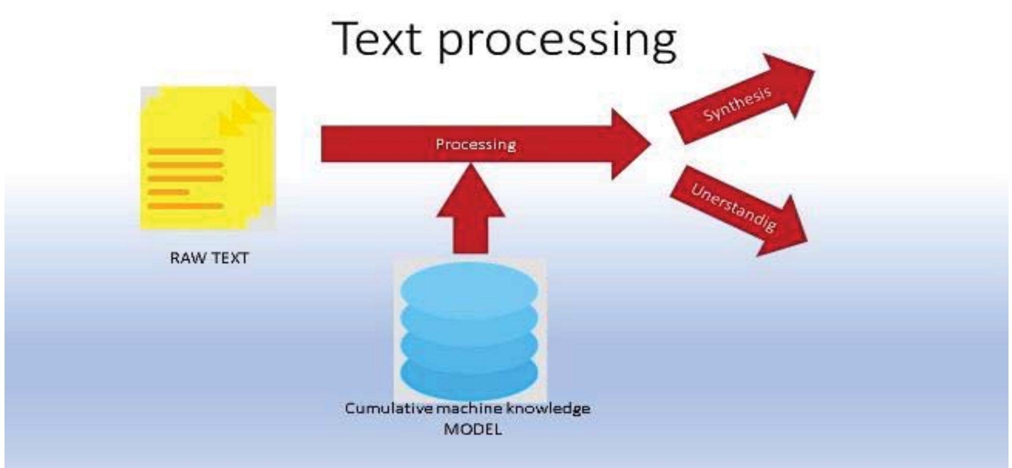


Figure 2 – Statistical NLP approach phase II



Text meaning must consider the universal fact that languages contain ambiguities (Kiss & Strunk, 2006). The same words can be different parts of speech (nouns, pronouns, verbs, adjectives, adverbs, etc.), and therefore play different roles in meaning (Goldberg, 2018). The same word, even when used as the same part of speech, can have different meanings depending on how it is used and the context within which it appears. The linguistic analysis involves the study of the elements, structure, and meaning of language: (Fraser & Hudson, 1992).

THEORETICAL DEVELOPMENT/MODEL

Text meaning is the process of extracting knowledge and information from natural language texts (Li, Zhou, Bruza, Xu, & Lau, 2008). There are several ways how NLP can help to gather insights out of any text description (Manning & Schütze, 2000). In this research are used segmentation of text description (by creating clusters) which can help to identify the problem area, where we can detect which parts of the service have the most impact (Alotaibi & Ahmad, 2019). This is the most difficult task for text meaning and involves in part the use of dictionaries, glossaries, lexicons, typologies, and so forth (Briscoe, Copestake, & Boguraev, 1990). The following steps were taken during text processing:(Yessenov et al., 2013).

Text preprocessing- Replacement of some special characters with spaces, and determination at the end of sentences and at the end of paragraphs.

The extraction step begins by attaching a label to each word and making the first attempt at the identification of candidate terms. Recall that terms are the building blocks of text analysis, corresponding to single or complex words that are considered relevant or interesting. During this step, the following actions are taken:

Text tokenization: A process that identifies character strings (tokens) from the input text, based on delimiters. Examples of delimiters are spaces, tabs, carriage returns, and punctuation marks.

Text normalization: Helps to manage poor punctuation in the text, such as improper use of a period, comma, semi-colon, colon, forward-slash, etc. The input text is "corrected" internally to place spaces around improper punctuation.

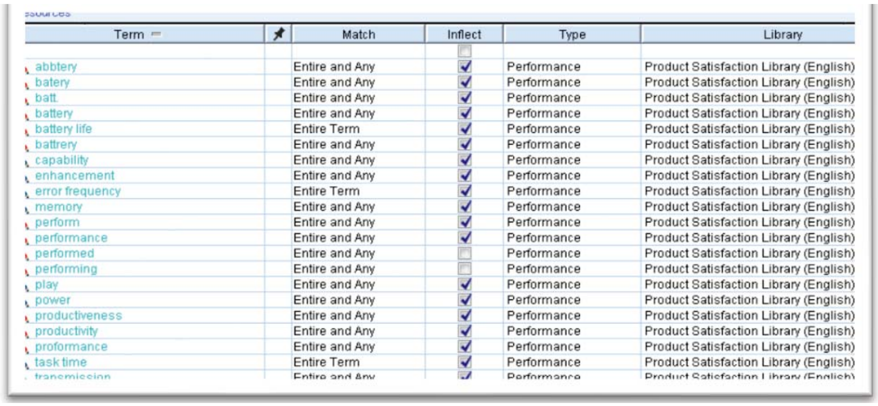
Candidate term extraction: Identifies relevant words and compound words from the input text.

Labelling the Part of the speech: Each token in the text stream is labelled by a Part of Speech tag, which comes from the base of the dictionary. Traditional English grammar classifies words based on eight parts of speech: verb, noun, pronoun, adjective, adverb, preposition, conjunction, and interjection. Each part of speech does not explain what kind of word is, but how the word is used.

Forcing / Excluding step: When we reviewed the text data after extraction, we should discover that some words or phrases were not extracted. While normally these words are verbs or adjectives that we are not interested in. These may be terms that are specific to an industry as product names, acronyms, etc. If we would like to have these words and phrases extracted, we can force a term into our dictionary, or we can exclude expressions that we no longer consider important.

Typing step: A type is a higher-level concept that contains one or more terms. There are default types for Organization, Product, Person. After extracting and modifying terms as described above, the extractor next assigns each concept to a type, wherever possible. See the example in figure 3.

Figure 3 -The type (Performance) and associated terms.



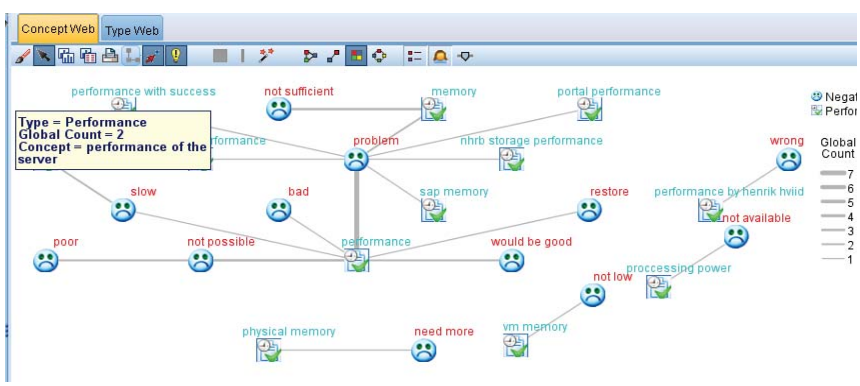
Term	Match	Inflect	Type	Library
abbtery	Entire and Any	✓	Performance	Product Satisfaction Library (English)
batery	Entire and Any	✓	Performance	Product Satisfaction Library (English)
batt.	Entire and Any	✓	Performance	Product Satisfaction Library (English)
battery	Entire and Any	✓	Performance	Product Satisfaction Library (English)
battery life	Entire Term	✓	Performance	Product Satisfaction Library (English)
battrery	Entire and Any	✓	Performance	Product Satisfaction Library (English)
capability	Entire and Any	✓	Performance	Product Satisfaction Library (English)
enhancement	Entire and Any	✓	Performance	Product Satisfaction Library (English)
error frequency	Entire Term	✓	Performance	Product Satisfaction Library (English)
memory	Entire and Any	✓	Performance	Product Satisfaction Library (English)
perform	Entire and Any	✓	Performance	Product Satisfaction Library (English)
performance	Entire and Any	✓	Performance	Product Satisfaction Library (English)
performed	Entire and Any		Performance	Product Satisfaction Library (English)
performing	Entire and Any		Performance	Product Satisfaction Library (English)
play	Entire and Any	✓	Performance	Product Satisfaction Library (English)
power	Entire and Any	✓	Performance	Product Satisfaction Library (English)
productiveness	Entire and Any	✓	Performance	Product Satisfaction Library (English)
productivity	Entire and Any	✓	Performance	Product Satisfaction Library (English)
proformance	Entire and Any	✓	Performance	Product Satisfaction Library (English)
task time	Entire Term	✓	Performance	Product Satisfaction Library (English)

Creation of categories step (grouping): The next step in text meaning is normally the creation of categories, which represent the information that we consider to be important in the responses. There are three categorization methods:

- concept inclusion,
- concept derivation
- semantic networks

The category refers to a group of closely related concept, opinion, or attitude used in text meaning. Each category is defined by one or more descriptors that are concepts, types, patterns, or conditional rules. Categories can only be created from a single concept or type, but it is common to combine multiple descriptors. The above-mentioned three language methods provide a natural language-based approach to categorization. For example, figure 3 presents the overlap for the categories to which the type (Performance) belongs according to the selection.

Figure 4 The overlaps for the categories



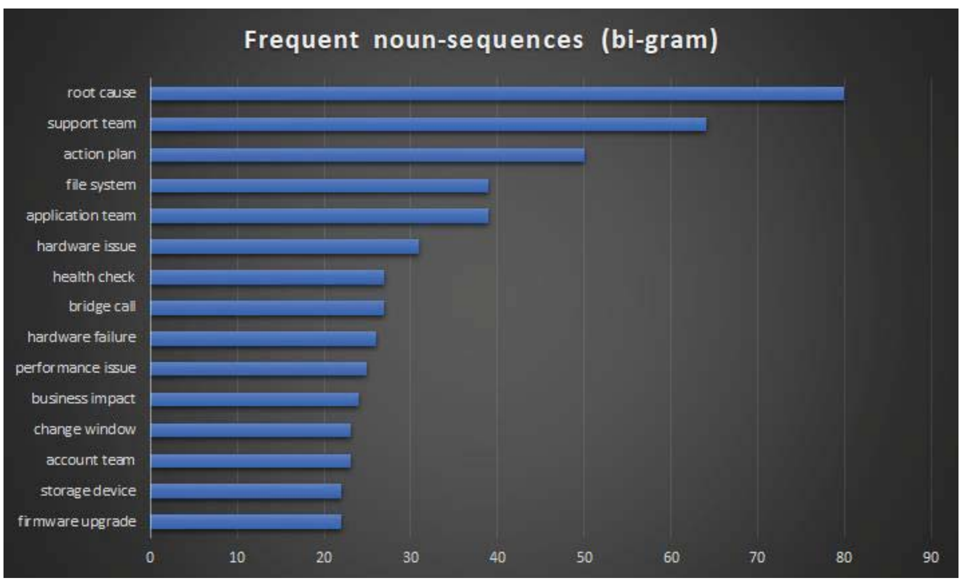
.

Indexing step: The types are indexed by establishing a pointer between a text position and the representative term for each equivalent class.
Matching patterns and events extraction step: NLP can discover not only types and concepts but also relationships among them.

RESULTS

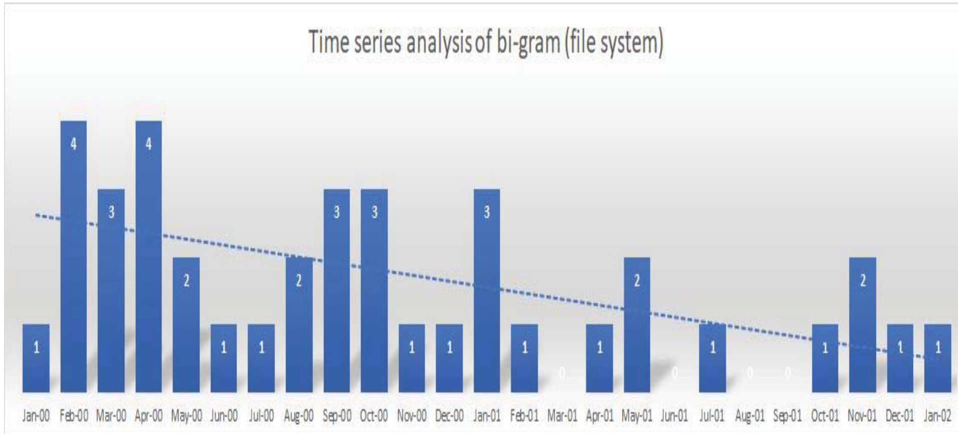
Semantic analysis has several advantages, it can detect connection among different tags, There are several combinations, which have powerful meaning like verb-noun connections (Huttenlocher & Lui, 1979). In the case of the noun, we can increase the significance if we identify not a single term, but a sequence. A typical noun sequence is a bi-gram, but that can be any number of nouns (n-grams) (Daniel Jurafsky & Martin, 2014; Majumder et al., 2002; Santos, Penya, Devesa, & Bringas, 2015). Advantage of the n-grams, that are comparing with the daily-speech environment occurrences of n-grams is relatively low (Gries & Mukherjee, 2010) while in the technical environments these occurrences are high (typically identifying technical terms or expressions) and have higher business meaning value, what we can use for analysis (Wang, McCallum, & Wei, 2007). One example, how bi-grams help to identify areas, what we can see here that several customer feedback refer to hardware problems or issues related to file systems or problems describing cases related to performance (See figure 5), In such cases, we need to look at the relationship between different IT environments where the problem occurs. We discovered, for example, that some types of databases could not work together with old operating systems (performance problem). In other cases, we have discovered that a poorly written application program in different IT environments has always caused a storage problem. The third case refers to the vulnerability of a hardware element.

Figure 5 – Frequent noun-sequence



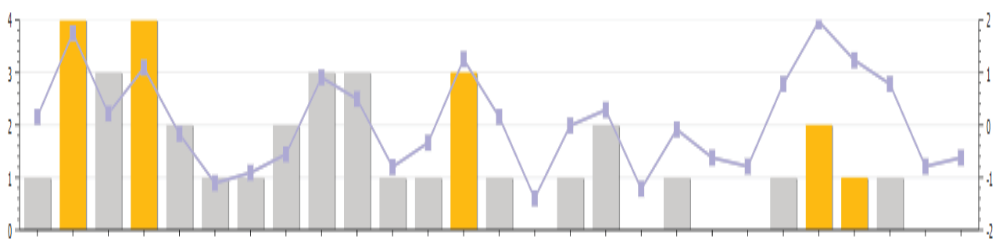
Since problems do not occur at the same time, so by enriching our data with time scales can easily help us to identify trends in some common grammar tags. For example, very often, the usage of some IT environments is not linear, but there are some frequent periods. In such cases, for example, periodical usage of the term "slow response time" can be observed. In another case, however, we can follow the elimination of the problem. For example, the mention of the "file system" bi-gram was reduced at the beginning of the year, because some applications started automatically to save the data. (See Figure 6)

Figure 6 – Time series analysis of bi-gram (file system)



At the same time series can be analysed from deviations points of view, which helps to identify those points in the time, which behave differently in time than what we would predict from previous data points (See figure 7). We could use this method as well in the case of problems with the smaller parts of the system. For example, a poorly formatted SQL query occasionally causes high memory usage. It is easy to found out which of the thousands of queries working badly with this method since only the running time of queries had to be compared with the time of deviation.

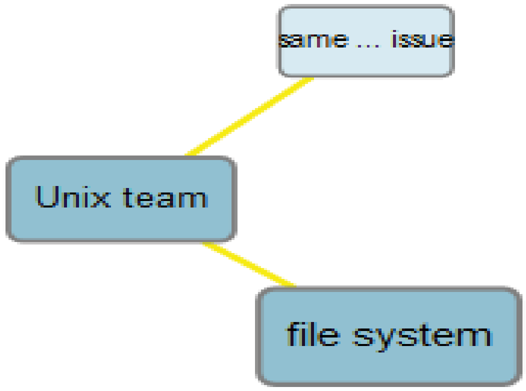
Figure 7 – Deviation detection



Semantic analysis can help to detect meaningful connections among different grammatical parts of the customer feedback.

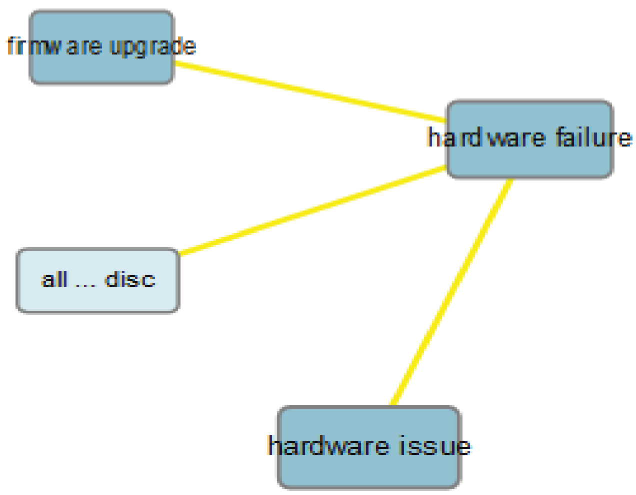
Let's see two examples of connection by detecting a correlation between noun sequences (bi-gram) and modified nouns (Chen, Seymore, & Rosenfeld, 1998). For example, we gain more insight by seeing that a larger part of file system issues is connected to Unix team (See figure 8). So we can identify that this problem occurs mainly around Unix environment. In such cases, the parts of a text ("same issue", "file system", "Unix team") are not necessarily mentioned sequentially. It may also happen that in different customer feedback only some synonyms of terms appear, nevertheless, we discover the relationships among the text elements.

Figure 8 - Connection network sample I



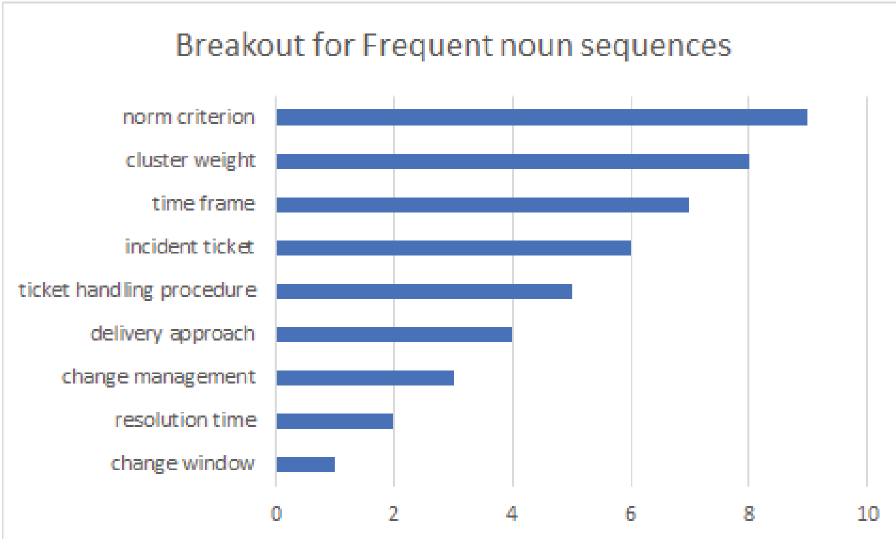
Another example shows that there is a correlation between hardware failure and firmware upgrade, (See figure 9) which can help us to identify common issues with firmware level. In this case, a correlation between the four elements of text occurs. The incorrect writing of "hardware" word does not cause trouble to recognize the connection among the text elements.

Figure 9 - Connection network sample II



Beyond the forms mentioned so far, we can identify structures such as noun sequences, modified nouns or nouns with predicates, which help us to identify more relevant text structures in describing our problem space (See figure 10).

Figure 10–Frequent noun sequences



The frequency of these terms and text structures may reflect the severity of the problematic area. In this case, we need to consider what solution shall we choose. For example, if many posts contain a "time frame" expression. That refers to a certain period when something planned is to happen in the IT system. In this case, we need to pay special attention to planning the introduction of changes in the future, because this is an important area for our customers.

DISCUSSION AND CONCLUSIONS

NLP is an efficient way to discover patterns on a large set of feedback descriptions, simply identifying grammatical structures supports to detect trends and deviations related to the provided service or the applied service components (Gao et al., 2007). There are grammatical structures, which are more valuable for this purpose, basically, nouns and verbs are the key language components from this point of view (Huttenlocher & Lui, 1979). Customers' feedback and incident descriptions individually should describe some parts of a customer' opinion but contains more information in a complex enterprise environment. These details are highly unstructured, and we need to apply techniques to extend information about them. By applying NLP procedures with other methods such as time series analysis, deviation detector connection, detecting, we can identify trends and deviations focusing on business-relevant expressions and terms.

REFERENCES

Alotaibi, F. M., & Ahmad, S. (2019). Keywords Extraction from the Text of Holy Quran Using Linguistic and Heuristic Rules. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY*.

Briscoe, T. (2011). Introduction to Linguistics for Natural Language Processing. *October*.

Briscoe, T., & Carroll, J. (1993). Generalized Probabilistic LR Parsing of Natural Language (Corpora) with Unification-Based Grammars. *Computational Linguistics*.

Briscoe, T., Copestake, A., & Boguraev, B. (1990). Enjoy the paper: lexical semantics via lexicology. *Proceedings of the 13th Conference on Computational Linguistics*.

Chan, S. W. K., & Franklin, J. (2011). A text-based decision support system for financial sequence prediction. *Decision Support Systems*. https://doi.org/10.1016/j.dss.2011.07.003

Chen, S. F., Seymore, K., & Rosenfeld, R. (1998). Topic adaptation for language modeling using unnormalized exponential models. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. https://doi.org/10.1109/ICASSP.1998.675356

Cohen, K. B., & Dolbey, A. (2007). Foundations of Statistical Natural Language Processing (review). *Language*. https://doi.org/10.1353/lan.2002.0150

Fraser, N., & Hudson, R. (1992). Inheritance in word grammar. *Computational Linguistics*, (October). Retrieved from http://dl.acm.org/citation.cfm?id=142239

Gao, J., Gao, J., Andrew, G., Andrew, G., Johnson, M., Johnson, M., … Toutanova, K. (2007). A Comparative Study of Parameter Estimation Methods for Statistical Natural Language Processing. *ACL*.

Goldberg, Y. (2018). Neural network methods for natural language processing. *Computational Linguistics*. https://doi.org/10.1162/COLI_r_00312

Gries, S. T., & Mukherjee, J. (2010). Lexical gravity across varieties of English: An ICE-based study of n -grams in Asian Englishes . *International Journal of Corpus Linguistics*. https://doi.org/10.1075/ijcl.15.4.04gri

Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*. https://doi.org/10.1109/mis.2009.36

Huttenlocher, J., & Lui, F. (1979). The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior*. https://doi.org/10.1016/S0022-5371(79)90091-4

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*. https://doi.org/10.1126/science.aaa8415

Jurafsky, Dan, & Martin, J. H. (2009). Speech and Language Processing 18 BT - An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. In *An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*.

Jurafsky, Daniel, & Martin, J. H. (2014). N-Grams. In *Speech and Language Processing*.

Kiss, T., & Strunk, J. (2006). Unsupervised multilingual sentence boundary detection. *Computational Linguistics*. https://doi.org/10.1162/coli.2006.32.4.485

Le, T., & David Jeong, H. (2017). NLP-Based Approach to Semantic Classification of Heterogeneous Transportation Asset Data Terminology. *Journal of Computing in Civil Engineering*. https://doi.org/10.1061/(asce)cp.1943-5487.0000701

Li, Y., Zhou, X., Bruza, P., Xu, Y., & Lau, R. Y. K. (2008). A two-stage text mining model for information filtering. *Proceeding of the 17th ACM Conference on Information and Knowledge Mining - CIKM '08*. https://doi.org/10.1145/1458082.1458218

Majumder, P., Mitra, M., & Chaudhuri, B. (2002). N-gram: a language independent approach to IR and NLP. *… Knowledge and Language*.

Manning, C. D., & Schütze, H. (2000). Foundations of Natural Language Processing. *Proceedings of the 20th International Conference on Computer Processing of Oriental Languages*.

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*. https://doi.org/10.1136/amiajnl-2011-000464

Pan, W., Liu, N. N., Xiang, E. W., & Yang, Q. (2011). Transfer learning to predict missing ratings via heterogeneous user feedback. *IJCAI International Joint Conference on Artificial Intelligence*. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-386

Reddy, S., Sharoff, S., Collobert, R., Weston, J., Bottou, L., Karlen, M., … Sangal, R. (2015). Deep Learning for NLP (without magic). *Arxiv*. https://doi.org/10.1017/CBO9781107415324.004

Santos, I., Penya, Y. K., Devesa, J., & Bringas, P. G. (2015). *N-GRAMS-BASED FILE SIGNATURES FOR MALWARE DETECTION*.

https://doi.org/10.5220/0001863603170320

Song, Y., & He, L. (2010). *Optimal rare query suggestion with implicit user feedback*. https://doi.org/10.1145/1772690.1772782

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*. https://doi.org/10.1613/jair.2934

Wang, X., McCallum, A., & Wei, X. (2007). Topical N-grams: Phrase and topic discovery, with an application to information retrieval. *Proceedings - IEEE International Conference on Data Mining, ICDM*. https://doi.org/10.1109/ICDM.2007.86

Yessenov, K., Tulsiani, S., Menon, A., Miller, R. C., Gulwani, S., Lampson, B., & Kalai, A. (2013). *A colorful approach to text processing by example*. https://doi.org/10.1145/2501988.2502040

Yu, D., Li, D. F., Merigó, J. M., & Fang, L. (2016). Mapping development of linguistic decision making studies. *Journal of Intelligent and Fuzzy Systems*. https://doi.org/10.3233/IFS-152026

DECISION SCIENCES INSTITUTE

Gamification in Entrepreneurial Education: Highlighting Major Concerns through a Systematic Mapping Study

Katalin Czakó
Széchenyi István University
Email: ckatalin@sze.hu

László Szendrői
Széchenyi István University
Email: szendroi.laszlo@sze.hu

ABSTRACT

Gamification is innovation in classroom instruction and consequence of following the global standards in university operations. In this study, through the application of systematic mapping, we screened empirical articles, where universities developed gamification in entrepreneurial education. Our goal is to present trends among the following questions: Who are the actors of applying gamification in entrepreneurial university atmosphere? How can we categorize the applied gamification tools? What research gaps can be defined? We observed that academics and institutional changes are less in focus of research. Categorization of gamification tools is constraint due to the dominance of serious games and online based (mobile) applications.

KEYWORDS: Games and Simulations, Classroom management, Gamification, Entrepreneurship

INTRODUCTION

Goal of this paper is to give an insight in present trends and research directions dealing with the efficient operations of higher education institutions and the application of gamification in entrepreneurial education.
Method applied in this paper is the mapping study, which is usually used as one of the first steps of comprehensive research in a given field and it is a useful method for defining the future research interest based on the findings of the systematic categorization of the existing research. Based on a systematic mapping study logic, represented in this paper, 20 papers were selected as suitable material for deeper data extraction.
We use the word “entrepreneurial education” once because of the logic of the analyses presented in this paper. Beside this under entrepreneurial education we understand all the courses, which are in field of business administration and economics. They are business or entrepreneurship related courses. We also focus on other courses, where developing the entrepreneurial mindset of students is the goal of the implemented gamification tools.
Literature of gamification in education provides several empirical analyses and can be categorized in sub topics. In order to be more specific, this analysis provides synthesis on two regularly researched topics – gamification and entrepreneurial university. General question of our interest is that is there any overlap among the two topics? If yes, what are the major elements and findings in this link. If no, what are the characteristics of the research gap? We know, that eventually we cannot separate two answers for the general question. In order to get closer insight to the extent of the assumed overlap, we highlight the following categories in the selected papers:

- Number of studies by disciplines or subjects.
- Course' types mentioned in the studies.
- Way of implementation of gamification tools.
- Examined variables in the studies.
- The mostly mentioned gamification tools in the studies
- General outcome or findings of the studies.

Recovering the common fields and elements in literature through the above listed categories may give more exact conclusions in the efficient operations of universities, which have the strategic focus to enforce the entrepreneurial spirit among students, faculty members and other shareholders of higher education operations. Since Clark's original definition (1998) there are changes in the role of entrepreneurial universities. In many cases, being an entrepreneurial university is not only a financial must, it is also a strategic and or differentiating goal, which is still in development phase in many cases (Czakó, 2017). Gamification is the tool to drive development processes and enforce different practices in operations of organizations in competitive sphere and also in education. Pelling (2011) was the first who coined the term of gamification when there were several authors used it before. He used it to describe those techniques which promote products and services. (Buckley and Doyle, 2017) The most common definition accepted by many researchers comes from Deterding et al., (2011) who defined gamification as a term for the use of video-game elements in the context of non-gaming systems to improve user engagement and experience. Since then, we can see several surveys, which general message is that applying gamification in education results better student performance, better experience, higher engagement and overall better teaching processes in educational institutions. Yilirdim (2017) focuses on students' attitude toward mathematic lessons. We would like to turn toward lessons of entrepreneurship and business and lessons in other topic, which improve the entrepreneurial skills of students. Basic assumption of us is that both of the topic of gamification in education and entrepreneurial university are upcoming areas among researchers and also in decision making processes of universities or other educational institutions. This gives the reasonability to make the synthesis, which arises several research questions.

In this paper we focus on the following questions: (Q1) Based on the categorization of this paper, what kind of tools are in the focus of the studies? (Q2) Who are the actors of gamification in entrepreneurial universities, how do they contribute to the usage of gamification tools? (Q3) Based on the synthesis of the paper, which research gaps can be identified?

This paper shows that the mapping study prospect is adequate to answer these questions. First part of the paper presents the major literature of gamification and entrepreneurial university with the goal of highlighting the link between the two topics. The theoretical model represents the main functions of mapping study and after this, the logic of mapping study process applied in this paper. Discussion part includes the extracted information from classification, which leads to the answering the above defined questions.

LITERATURE REVIEW

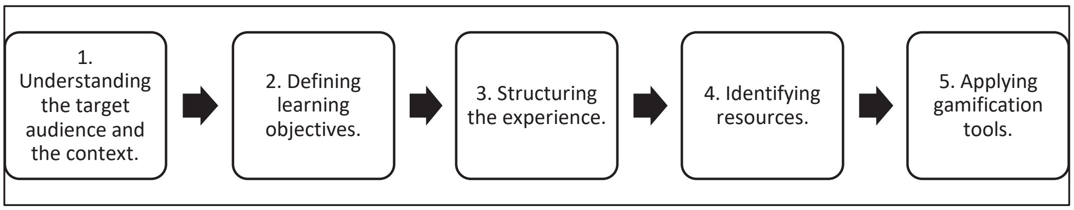
As it was mentioned before, this paper's theoretical framework has two pillars. The first important topic is gamification in education. The second is entrepreneurial university and more precisely, the integration of entrepreneurial practices in higher education.

Gamification in education

Huang and Soman (2013) elaborated a five-step process to apply gamification in education. Figure 1. gives shows those conditions, which are needed to the successful application

Understanding the target audience and defining the learning objectives are the very first steps before the application. Usually there are more versions of application. Structuring the experiences supports to collect needs or mistakes from previous applications. As it is sign in step 4. Application of gamification tools are strongly resource oriented actions. It is a basic requirement to provide equal availability to the members of the target audience.

Figure 1. The five-step process of applying gamification in education.



Source: Huang and Soman (2013)

Gamification can be found in the Oxford dictionary, which explains it in the following way: "The application of typical elements of game playing (e.g. point scoring, competition with others, rules of play) to other areas of activity, typically as an online marketing technique to encourage engagement with a product or service." (Oxforddictionaries, 2019).
Hamari et al. (2014) found the most common motivational terms related to gamification: points, leader boards, achievements/badges, levels, story/theme, clear goals, feedback, rewards, progress, and challenge. Dicheva et al. (2015) in an educational context found that the following gamification elements are the most commonly used: points, badges, leaderboards, levels, virtual currency, progress bars, and avatars. These elements have different motivational values; therefore, they have to be customized according to the environment and different types of individuals (Barata et al. 2015). According to Costa et al. (2017) it is difficult to define each gamification element as they are in many cases related to each other and sometimes researchers use them in different ways, meaning there are always overlaps between them.
There are both studies which examined the attitudes and knowledge of the teachers toward gamification (Adukaite et al., 2017; Fisher et al., 2014; Martí-Parreño et al., 2016) and studies investigating the student motivational effects and the learning outcome of using gamified classes (Antonio et al., 2015; Wintermeyer & Knautz, 2015).

Entrepreneurial universities – external and internal research perspectives

Clark (1998) gives the prototype criterions of entrepreneurial universities through a continuous operation analysis and case study building in the middle of 90's in chosen European Universities. There are internal criterions, which descript the changing mindset of students and faculty members toward entrepreneurial practices like taking part in common research activity or providing creative lessons and courses. Topic of entrepreneurial universities got great interest, it is showed by Mascarenhas et al. (2017), who gives a comprehensive literature review about the study of entrepreneurial universities from 1900 to nowadays. Considering external point of views Slaughter (1997) focuses on the effects of policy changes and centralized strategies on the corporatization process. Fekete (2015) collects economic development tools in governance of European cases and highlights the growing role of universities in industrial cities. (Abeles 2001) links external and internal point of views with the finding of growing number of university partnerships and other outsourcing processes can have also mainly cultural effects on internal university operations toward being more entrepreneurial.

Gamification and entrepreneurial universities

Basic assumption of us is that both of the topic of gamification in education and entrepreneurial university are upcoming areas among researchers and also in decision making processes of universities or other educational institutions. Gamification serves as supporting aspect in the transformation process to be entrepreneurial. This statement is supported by Troudt et al. (2017), who says that in order to turn entrepreneurship education into a semiautomated system of rewards, we must understand the actions of the entrepreneur. This statement can be strongly linked to the Criteria of Clark (1998) toward faculty members in an entrepreneurial transformation of a university.
Rippa and Scundo (2018) listing and summarize different types of technologic tools used in academic entrepreneurship, which has also clear focus in term of changing attitude of faculty members (e.g. thinking of Strike (2009); Normand (2016); Jessop (2017)). They refer to Galan (2013), whose case study illustrated that gamification is presented in these tools as improving factor of not only in the student learning experience but it is closely linked to research activity, which leads to effective community outreach.
Belotti et al. (2014) measures in European Universities, that serious games are welcomed by Students and the use of them and found them to be useful for introducing some difficult topics. Entrepreneurial motivation, business competence, and business acumen were identified as key entrepreneurial skills. There is a general question by the authors that what are the needed skills and competences by an educator to apply these games successfully. Are these competences closely linked to the entrepreneurial mindset? According to the above mentioned basic assumption, we think that applying gamification tools requires an entrepreneurial atmosphere. In order to strengthen this assumption, we provide the systematic mapping study to recover the literature was made in the topics of gamification and entrepreneurial university.

METHODS

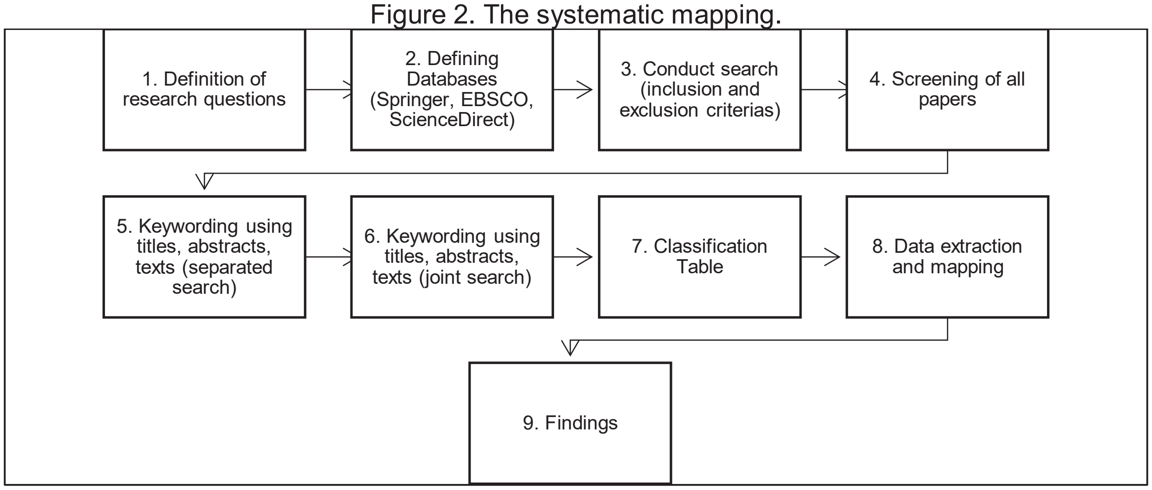
The mapping study methodology is usually used as a first step of the research and it is a useful method for defining the future research interest. It is similar to a systematic literature review, however, there are differences regarding the goals and other aspects between them. The mapping study focuses more on the classification of the collected studies and thematic analysis of literature and not the aggregation of the information from the comparative studies. The research questions are related more to research trends and not the outcomes of empirical studies. The scope of the research is broader, which means that all papers are selected to a topic area but only the classification data are collected from them and individual research outcomes are not extracted from each paper. The research strategy requirement is less stringent, because only trends are of interest, therefore authors may search only for journal papers or restrict themselves to only one or two search engines. So, it is not necessary to find all relevant studies like in a literature review (Kitcenham, 2011). Analysis of this paper is close to logic of Dicheva et al. (2015).
In this paper, we used more search engines and applied the definition of inclusion and exclusion criteria in the searching process. Another difference is that, we collected outcomes or findings from each paper after the screening process based on the inclusion and exclusion criteria.
A mapping study has several benefits for researchers. After a thorough mapping study, it is easier and less time-consuming to move forward in future research. As it provides a comprehensive overview of the literature it can also help in understanding the literature and to construct research questions. A theoretical contribution of a mapping study is usually the definition of the needed primary studies in sub-topics. The procedures, forms and experiences

can be also reused, and the results can provide a basis for comparison with the follow-on studies. Finally, the primary studies provided, can be used to validate further research and results (Kitcenham, 2011).

Mapping study of the topic of gamification in entrepreneurial universities

Analysis in the paper follows the logic mapping study process. Figure two shows the steps of the mapping study represented in the paper. Conducting the research, the process of the research was accomplished based to the systematic mapping steps, provided by Petersen et al. (2008).
The screening was conducted through the search in Springer, EBSCO and ScienceDirect. As it is visible on Figure 2, in the 4th step, there is a screening after defining inclusion and exclusion criteria. The used terms in screening is showed by Table 1.

Figure 2. The systematic mapping.



Own source based on: Petersen et al., (2008)

The inclusion criteria were the following:

- Where several papers reported the same study, only the most complete one is included.
- Where several studies were reported in the same paper, each relevant study was treated separately.
- Studies that answer at least one research question.

The exclusion criteria consisted the following:

- Studies that are not written in English.
- The study is not related to institutional education or learning.
- The study does not contain empirical research.

Screening of all papers

Firstly, the screening was conducted based on a separated mapping activity. In the first round, keywords “gamification” and “education” and key words of “entrepreneurial” and “university” or “universities” were screened in the databases (5th step in Figure 2). Our assumption was that entrepreneurial university is an older term. So, the keywords “gamification” and “education” were searched not only in titles of papers but also in abstracts or the whole text. Keywords “entrepreneurial” and “university” or “universities” were searched only in titles of papers. Through the separated search, we could extract data in the following fields:

- Number of selected papers by year,
- distribution by database,
- number of papers by journal,
- number of studies by education level, which is interesting input in case of “gamification”, but not in the focus of this paper and also recall low level of interest as this topic shows increasing number of studies year by year.

The second round of the searching activity was the joint search (6th step in Figure 2). Two ways were applied. In the first case „gamification” was searched in the title and „entrepreneurial” „university” or „universities” in the text - or „entrepreneurial university” or „universities” in the title and „gamification” in the whole text. In the second case, „gamification” or „entrepreneurial” or university” or „universities” words were searched in the whole text. Trough the joint search, we could reflect information in the following categories, represented in the introduction and later in Table 2. of this paper:

- Number of studies by disciplines or subjects.
- Course’ types mentioned in the studies.
- Way of implementation of gamification tools.
- Examined variables in the studies.
- The mostly mentioned gamification tools in the studies
- General outcome or findings of the studies.

RESULTS

In this section, we introduce the results from the step of the screening of all papers (4th step) and then the categorization phase (7th step in Figure 2) in order to represent the possibility of extracting information, which provide a general picture of the research trends. Based on the selection of articles. Table 1. Shows the numerical data of the screening:

Table 1. Number of the searched papers by database. Selection of articles

	FIRST ROUND – SEPARATED SEARCH		SECOND ROUND – JOINT SEARCH	
	„gamification” „education” in abstracts and the whole text	„entrepreneurial university” in titles	In the whole text: „gamification” and “entrepreneur” or „entrepreneurial” and „university”	Duplicates
Springer	269	35	3	1
EBSCO	246	167	11	9
ScienceDirect	194	60	14	10
SUM			28	20

Own source based on the screening process

After making the separated search of papers (5th step in Figure 2) it could be concluded that the number of articles are constantly growing in gamification and there is a continuous increasing trend that more and more empirical papers are written related to gamification in educational context. In case of entrepreneurial university, the first articles have been found from 1989. Number of articles are peaking for 2015. Dominating journals in case of separated search if the

topics can be defined. In case of entrepreneurial university Journal of Technology Transfer and Small Business Economics are the mostly screened Journals in the databases. In case of gamification, there are many Journals in field of IT.
We looked at the education level on this stage of the screening as well. Gamification related articles could be analyzed in term of education level. Most papers investigated the effects of gamification in higher education. In case of entrepreneurial university, this question is not relevant.

Gamification in entrepreneurial universities - specific categories

After making the joint search (6th step in Figure 2), the following information could be extracted to the Classification Table (Table 2) from 20 studies.

Table 2.: Classification Table

AUTHORS	FIELD	COURSE TYPE	WAY OF IMPLEMENTATION	VARIABLES/RESEARCH FOCUS	TOOLS	MAIN FINDINGS
Antonaci et al. (2015)	IT	Entrepreneurship	Stimulation	Students' perception	Serious games	Theoretical model supporting the choice of serious games.
Felker (2014)	IT	Library usage	Online platforms	Online and physical presence of users – Online usage	Information sharing (mobile) applications	Describing ways to gamify libraries and details the effect on students' attitudes to library usage
Holotescu et al. (2017)	Smart Cities	Entrepreneurship	Online platforms	Online usage	Open online courses	Effects of courses on smart city interventions.
Ciupe et al. (2018)	Career	Business and entrepreneurship - entrepreneurship	Simulation	Institutional elements	Virtual enterprise	Link between applying virtual enterprise and five practices in entrepreneurship and entrepreneurial competency model and responses in education.

Sanip & Rahman (2018)	Education	Medicine	Performance systems	Experiences of implementing the system – Professionals' experience	Integrated cumulative GPA	Needed competences and responses for the medical faculty members.
Bodnar et al. (2016)	Education	Engineering	Games	Student learning outcomes – Students' performance	Serious games	Undergraduate student learning was improved by game-based activities
Pianchi & Lió (2017)	Education	Medicine: genomics and bioinformatics - Medicine	Online platforms	Digital-genomic divine – Software performance	Community awareness platforms	Bottom-up development efforts in education. Role of bioinformaticians.
Naaji et al. (2015)	Education	Humanistic, economic, social, medical studies – Mixed courses	Online platforms	e-learning experience of Romanian educational system – exact case analysis	Flipped classrooms + social media	Listing related sciences. Standards required for evaluating the quality of online and blended courses.
Vaughan (2013)	IT	Library usage	Online platforms	Responses of members of the Association of Research Libraries – Library users' responses	Information sharing (mobile) applications	Suggestions to innovative solutions in libraries.
Dziob (2018)	Education	Physics	Serious games	Students' responses with control group's responses – Students' perception	Serious games	New method of student assessment.
Zaina & Alváro (2015)	IT	Entrepreneurship	Online platforms	End user requirements and needs – Users' requirements	Information sharing (mobile) applications	New ideas for software development.

Sousa et al. (2019)	Education	Business and ICT Entrepreneur trainings - Entrepreneurship	Online platforms and games – Mixed platforms	Perceptions of students and entrepreneurs about E-education methods. – Students' and Entrepreneurs' perceptions	Serious games	Gamification is more adequate in creating new ideas and market and product analysis. Skills, motivations, support from the HEI, barriers, difficulties and sociodemographic traits.
Mincha et al. (2018)	Education	Business – Entrepreneurship	Online platforms	Case based outcomes of implementing the program - Exact case analysis	GTP-global talent program	Effects of software usage: skills and competence development.
Seixas et al. (2016)	Education	Elementary school courses – Mixed courses	Online platforms and games – Mixed platforms	Students' engagement level – Students' perception	ClassBadges	Classification of student groups. Strong causality between more rewards and better performances.
Faghihi et al. (2014)	Education	Mathematics	Software	Difference in student performance – Students' performance	Serious games	Students, who used the gamified system reached greater scores Video game elements vs. AI tutoring system (MathDungeon).
Yen-Chun Jim Wu	Education	Entrepreneurship	Online platform, software –	Students' perception	Information sharing (mobile)	mobile-based CRS technology is a useful and

et al. (2017)			Online platforms		applications	effective tool for facilitating interaction among learners and content, enhancing students' engagement with entrepreneurial knowledge acquisition, and improving students' motivation toward increased entrepreneurial capability.
Severengiz (2018)	Education	Engineering education – Manufacturing - Engineering	Gamified class	Students' perception and performance	Gamified examination	Gamifying a classic MC test has no significant impact on the test results and does not influence the perceived level of difficulty of the test questions.
Reise et al. (2014)	Education	Manufacturing - Engineering	Games	Game elements	Serious games	Description of usage of the game
Belotti et al. (2014)	Education	Entrepreneurship	Games	Game elements	Serious games	Overview of relevant SGs available on the market and identifies, through an expert analysis, key benefits and issues concerning their adoption

						in teaching entrepreneurship for the target students.
Popoiu et al. (2012)	Education	Medical courses - Medicine	Social media – Online platforms	Medical educators' experiences – Professionals' experience	Blogs and microblogs	Best practices in teaching.

Own source

Number of studies by discipline or subject

The articles were categorized by the disciplines or fields of subjects in which context's researchers examined gamification tools in universities. Education (14 studies) Information Technology (4 studies), Career, Smart City (1-1 study). It is worth to note, that in empirical studies, in which students are the target, the findings are concluded to effective way of teaching and the positive effects of choosing a gamified process.

Number of studies by courses

In a wide aspect, seven courses could be identified. Obvious outcome is that entrepreneurship courses were represented in the highest number (8 studies) others are medical courses (3 studies), engineering courses (3 studies) and Library usage (2 studies). There were two studies, where more courses were mentioned and physics and math in one-one studies. The fact that entrepreneurship classes were overrepresented is not distorting the answers of the research questions of this paper.

Ways of implementation

It can be defined as novel result that in most of the studies the implementation of online platforms are presented. There are several offline ways of implementation gamification tools (competitions, offline badges, group games). In this focus of papers – in the topic of gamification in entrepreneurial teaching – online based games, social media platforms, other online communication channels are implemented. There are few cases, where gamified examination is presented or performance systems are implemented, but there is no exact information whether they are based on online platform or not. It is more surely that these are based on mixed – online and offline platforms.

Variables or quantifiable outputs

We tried to extract more precisely the variables or quantifiable research focuses in case of empirical researches. The following variables are presented in the selected papers:

- Students' perception.
- Student's performance.
- Students' and entrepreneurs' perception.

- Level of online usage.
- Institutional elements.
- Professionals experience.
- Software performance.
- Case analysis: exact case-based data.
- Library user's experience.
- User requirement: in case of a given program.
- Game elements.

Gamification tools mentioned in the papers

We can agree with the finding of Costa et al. (2017), that it is difficult to categorize gamification tools according to their function, because there are several games with licenced names, which have customized, for their own target. After this mapping, we focused on the followings when we categorized the tools: all the games, which had exact names or brands was categorized in the group of serious games. All the platforms, which had the only function of communication and interaction, were categorized in the group of information sharing "mobile" applications. All the others outside of these two major categories can be differ from serious games and information sharing mobile applications in their functions:

- Open online courses.
- Virtual enterprise.
- Integrated, cumulative GPA.
- Flipped classrooms (social media).
- Global talent program.
- Community awareness platforms.
- Blogs.
- Gamified exams.
- Class badges.

Logic behind this is that we did not want to list all the exact game names presented by several papers, but we did not want to be too general with creating big categories of tools. To do so, we focused on the function of the mentioned tools in the selected papers and categorized them based on it. As a result, we could conduct the list of gamification tools in entrepreneurial teaching based on their function.

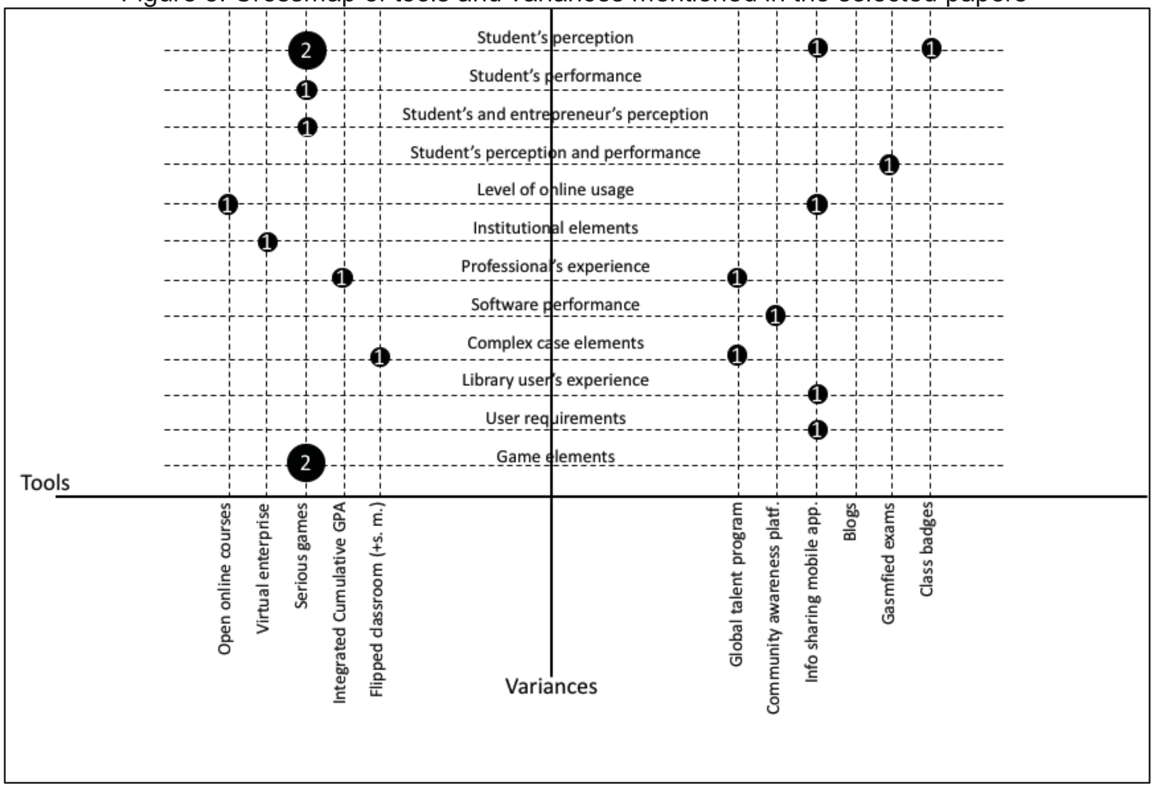
Outcomes or findings of the studies

As it is written in the methodological discussion of mapping study process, it is not needed to collect outcomes in case of each studies. In order to be more focused, thanks to the joint research, we could list the major findings, which can be valuable if they are in one table. Categorization in these ones was also not so direct, like in case of variables. Through the collection of outcomes and findings we were looking for whether gamification provided a positive effect on the different variables that were examined by authors in the articles. Authors represent most of the cases positive effects, which meant that gamification had an overall positive influence on each examined variable. The application of gamification tools in entrepreneurial teaching is presented as effect in the context of open innovation or technology-based entrepreneurship. Obvious, but worth to mention that in studies, where authors mentioned an exact tool (e.g.: exact program names in group of serious games), the applied methodology was case study and the findings were generalized less. There were studies, where the market of the given program was represented in the results (Belotti et al. 2014).

DISCUSSION

Figure 3 shows the mapping of gamification tools mentioned and variables analyzed in the studies in a crossmap. Beside answering the above-mentioned Q1 question of the paper, this figure makes us able to extract the actors of gamification applied in entrepreneurial education (Q2).

Figure 3: Crossmap of tools and variances mentioned in the selected papers



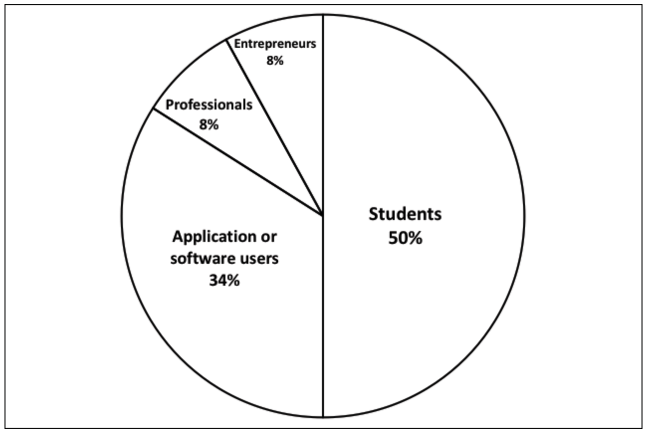
Own source based on the selected papers

Major group of actors, based on the in the selected papers:

- Students
- Professionals
- External users of apps and software applied by the university: entrepreneurs, customers

Main and dominant group of actors based on the analyzed studies are students on the courses, where a given gamification tool is applied. Their perception, performance, requirements, level of usage and soft variables like experience are measured in the studies, mostly in case-based dimension. To be keen on the papers in the final selection, we named customers as users and we can represent the ratio of actors taking part in studies of gamification in entrepreneurial education:

Figure 4: Representation of group of actors in the selected papers



Own source based on the selected papers

Figure 4. shows the ration of target audiences mentioned in the studies. There are an obvious inequality, which gives research gap in field of research on gamification.

Underrepresented fields of research

Through Figure 3, we can define also research gaps (Q3), which can validate new questions in present or further research trends, satisfying the function of mapping study (Kitcenham, 2011). If we look at the variables or the list of actors, studies did not name educator, lecturer or professor as an actor. We did not really find information about the changing role of them, which can be linked to the phenomena of the changing role of homo academicus in the entrepreneurial environment of universities. Advantages and disadvantages of homo oeconomicus has huge literature, while in these studies the focus is mostly the effect of application on students and professionals, who are able to apply the given tool on the classes or in other university operations like library usage. Relatively there are less studies, which are dealing with the needed skills and knowledge of using a gamification tool in in-class processes.
Variables of institutional elements are less represented. Based on this collection of papers a gap is indicated, which focus on the change of institutional elements of universities or their processes. Operational processes, structural changes can be assumed in many cases of implementing a gamification tool. For example, the usage of one software or online based application targets not only one department or course in a university. The implementation of the usage of the given tool requires horizontal management aspects in universities, where departments have high level of autonomy in each case, but they need to share sources in order to use this opportunity efficiently. Here is the question: what are the institutional requirements of an efficient implementation of a gamification tool?
As institutional variables were underrepresented in the extracted data, we can say that the overlap between the research of gamification and entrepreneurial university is not significant interest of researchers in this field. Perception of actors (students, professionals, entrepreneurs and software users- are the dominant focus of the selected articles.

CONTRIBUTION

Overall expectation of this analysis was to get some commonly examined trends both in topic of gamification and entrepreneurial university. We can say that logic of this mapping study process

was appropriate to find the final selected papers, which are the closest to the overlap of these topics.
The application of gamification tools in entrepreneurial teaching is presented as effect in the context of open innovation or technology-based entrepreneurship. Although there are growing number of analyzed studies in each topic, the overlap between applying gamification tools and examining their effects in entrepreneurial higher education institutions is relatively small comparing to the gamification effects on actors' perceptions or experience in them.
Through the application of mapping study, paper's theoretical contribution is visible. Applying the mapping study gives the novel result, that this method can contribute firstly to the reasonability of the synthesis of different research topics, which are linked together in one institution. Mapping study help to get an insight, whether different research topic is analyzed hand in hand or separately on empirical base. Secondly, this method also contributes to the validation of future research questions based on the logical screening and extraction of data of the papers.
This study was focusing on the fields, course types, gamification tools, ways of their implementation and examined variables. Classification behind this focus was enough to represent the actors of gamification in entrepreneurial universities with the exception of academic faculty members. It is an empirical gap, which was highlighted in this research.
Another empirical contribution we would like to mention is the underrepresented institutional aspects compared to actors' perception, requirements and experiences. Gamification tools are implemented in universities. Especially in case of online systems, most of the times, efficiency of them is based on the number of users. The higher the users number, the better is the given software or platform.
In case of this selection of the papers we could see that there is not unified vocabulary for gamification tools. In case of this mapping study the results can support the statement of Costa et al. (2017). Although mentionable research contribution can be that we extracted the gamification tools based on their functions. At the same time, we admit the difficulty of categorization of these tools:

Table 3. Tools and their functions of gamification in entrepreneurial education

TOOL	FUNCTION
Serious games	Learning
Info sharing (mobile) applications	Sharing information
TOOL WITH PLUS FUNCTIONS	PLUS FUNCTIONS
Open online courses	Recruiting
Virtual enterprise	Enterprising
Integrated, cumulative GPA	Testing
Flipped classrooms (social media)	Grouping
Global talent program	Sample tracking
Community awareness platforms	Grouping
Blogs	Testing
Gamified exams	Exam
Class badges	Rewarding

Own source based on the selected papers

Table 3 summarizes the functions of gamification in entrepreneurial education. As it was assumed previously, we realized overlap as serious games and info sharing applications have general functions in this grouping, which are represented by other tools listed under them.

RESEARCH LIMITS

In this research it was difficult to define each gamification tool. In many cases there were mixed tools. Offline ways of implementation were not visible in every cases. Categorization of gamification tools in this analysis lead to overlaps among gamification tools.
Other cross maps and Tables are available upon a request.

REFERENCES

Abeles P., T. (2001). “Partnering and cultural change”. *On the Horizon*, 9(4), 2-3. https://doi.org/10.1108/10748120110803761.

Adukaite, A., van Zyl, I., Er, Ş., & Canto, L. (2017). Teacher perceptions on the use of digital gamified learning in tourism education: The case of South African secondary schools. *Computers & Education*. https://doi.org/10.1016/j.compedu.2017.04.008

Antonaci, A., Dagnino, F. M., Ott, M., Belotti, F., Berta, R., De Gloria, A., Lavagnino, E., Romero, M., Usart, M., Mayer, I. (2014). A gamified, collaborative course in entrepreneurship: Focus on objectives and tools. *Computers in Human Behavior*. 2015(51), 276-283.

Antonio J. Sierra, Alvaro Martin-Rodríguez, Teresa Ariza, J. M.-C. and F. J. F.-J. (2015). Methodologies and Intelligent Systems for Technology Enhanced Learning. *LTI for Interoperating E-Assessment Tools with LMS*, 9. https://doi.org/10.1007/978-3-319-19632-9

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2015). Gamification for smarter learning: tales from the trenches. *Smart Learning Environments*, 2(1), 10. https://doi.org/10.1186/s40561-015-0017-8.

Belotti, F., Berta, R., Gloria, A. D., Lavagnino, E., Antonaci, A., Dagnio, F., Ott, M., Romero, M., Usart, M., Mayer, I.S. (2014). Serious games and the development of an entrepreneurial mindset in higher education engineering students. *Entertainment Computing*. 2014(5), 357-366.

Bianchi, L., Lió, P. (2017). Opportunities for community awareness platforms in personal genomics and bioinformatics education. *Briefings in Bioinformatics*, 18(6),1082–1090.□

Bodnar, C. A., Anastasio, D., Enszer, J., Burkey, D. (2016). Engineers at Play: Games as Teaching Tools for Undergraduate Engineering Students. *Journal of Engineering Education*. 105(1), 147-200.

Buckley, P. and Doyle, E. (2017). ‘Individualizing gamification: An investigation of the impact of learning styles and personality traits on the efficacy of gamification using a prediction market’. *Computers and Education*, 106, 43–55.

Ciupe, A., Ionescu, I., Meza, S., Orza, B. (2018). Towards Agile Integration within Higher Education: A Systematic Assessment. *Broad Research in Artificial Intelligence and Neuroscience.* 9(3), 69-87.

Clark, BR. (1998), “*Creating Entrepreneurial Universities: Organizational Pathways of Transformation.” Issues in Higher Education.* Elsevier Science Regional Sales, NY.

Costa, C. J., Aparicio, M., Aparicio, S., & Aparicio, J. T. (2017). Gamification usage ecology. *Proceedings of the 35th ACM International Conference on the Design of Communication* (p. 2). ACM.

Czakó, K. (2017) Notions of entrepreneurial university: some European models. *European Operations Management Association (EurOMA) (Ed.), Proceedings of 25th EurOMA Conference*. Budapest, Hungary (2018). 1-10.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining & quot; *Gamification & quot.* Retrieved from https://s3.amazonaws.com/academia.edu.documents/30609294/MindTrek_Gamification_PrinterReady_110806_SDE_accepted_LEN_changes_1.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1502650108&Signature=Wu4V9uWW2gLRdXg69jnba9F6fow%3D&response-content-disposition=inline%3B filename%3DFrom_game_design_elements_to_gamefulness.pdf., May 14.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society,* 18(3), 1-14.

Dziob, D. (2018) Board Game in Physics Classes - a Proposal for a New Method of Student Assessment. *Research Science in Education*. 2018 March. https://doi.org/10.1007/s11165-018-9714-y.

Faghihi, U., Brautigam, A., Jorgenson, K., Martin, D., Brown, A., Measures, E., Maldonado-Bouchard, S. (2014). How Gamification Applies for Educational Purpose Specially with College Algebra. *Procedia Computer Science.* 2014(41), 182-187.

Felker, K. (2014). Gamification in Libraries - The State of Art. *Reference and user service Quarterly.* 54(2), 19-23.

Fekete, D. (2015) *New regional economic development methods of European automotive centers.* Universitas, Hungary.

Fisher, D. J., Beedle, J., & Rouse, S. E. (2014). Gamification: a Study of Business Teacher Educators' Knowledge of, Attitudes Toward, and Experiences With the Gamification of Activities in the Classroom. *Journal for Research in Business Education*, 56(1), 1–16. Retrieved from https://reddog.rmu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=115099505&site=ehost-live&scope=site, May 10.

Galan, F. (2013). Introducing some gamification features in an undergraduate entrepreneurship course. *Edulearn13 Proceedings*. IATED, 5432-5437.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? - A Literature Review of Empirical Studies on Gamification. *2014 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Holotescu, C., Goţiu, L. O. L., Andone, D., Cismariu, L., Grosseck, G., Slavici, T. (2017). Entrepreneurship Learning Ecosystem for Smart Cities through MOOCs. *Broad Research in Artificial Intelligence and Neuroscience.* 8. (2),33-43.

Huang, W. H., & Soman, D. (2013). *Gamification Of Education.* University of Toronto - Rotman School of Management, 1–29. https://doi.org/10.1111/j.1467-8535.2011.01259.x

Kitchenham, B. A., Budgen, D., & Pearl Brereton, O. (2011). Using mapping studies as the basis for further research–A participant-observer case study. *Information and Software Technology*, 53(6), 638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Jason, V. (2013). *Technological Innovation: Perceptions and Definitions*. Library Technology Reports. USA: American Library Association.

Jessop, B. (2017). Varieties of academic capitalism and entrepreneurial universities. On past research and three though experiments. *Higher Education*. 2017(1), 853-871.

Martí-Parreño, J., Seguí-Mas, D., & Seguí-Mas, E. (2016). Teachers' Attitude towards and Actual Use of Gamification. *Procedia - Social and Behavioral Sciences*, 228, 682–688. https://doi.org/10.1016/j.sbspro.2016.07.104

Mascarenhas, C., (2017). "Entrepreneurial university: towards a better understanding of past trends and future directions", *Journal of Enterprising Communities: People and Places in the Global Economy*, Vol. 11(3), 316-338.

Minocha, S., Hristov, D., Leahy-Harland, S. (2018) Developing a future-ready global workforce: A case study from a leading UK university. *The International Journal of Management Education*. 2018(16), 245-255.

Naaji, A., Mustea, A., Holotecsu, C., Herman, C. (2015). How to Mix the Ingredients for a Blended Course Recipe. Broad Research. *Artificial Intelligence and Neuroscience.* 6(1-2), 106-116.

Normand, R. (2016). *The changing epistemic governance of European Education: The fabrication of the Homo Academicus Europeanus?* 2016, Sringer.

Oxforddictionaries (2017). *Gamification - definition of gamification in English* | Oxford Dictionaries. (n.d.). Retrieved from https://en.oxforddictionaries.com/definition/gamification. May 14.

Pelling, N. (2011). *The (short) prehistory of gamification*. Retrieved from http:// nanodome.wordpress.com/2011/08/09/the-short- prehistory-of-gamification/, May 16.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic Mapping Studies in Software Engineering. *12th International Conference on Evaluation and Assessment in Software Engineering,* 17, 10. https://doi.org/10.1142/S0218194007003112

Popoiu, M. C., Grosseck, G., Holotescu, C. (2012). What do we know about the use of social media in medical education? *Procedia - Social and Behavioral Sciences.* 2012(46), 2262-2266.

Reise, C., Müller, B., Seliger, G. (2014). Resource Efficiency Learning Game – Electric *Scooter Game. Proceedings of 21st CIRP Conference on Life Cycle Engineering*. 2014(15), 355-360.

Rippa, P., Secundo, G. (2018) Digital academic entrepreneurship: the potential digital technologies on academic entrepreneurship, *Technological Forecasting and Social Change*. 2018, July.

Sanip S, Abdul Rahman N., F. Integrated cumulative grade point average (iCGPA): benefits and challenges of implementation for the medical faculty. *Education in Medicine Journal*. 10(1), 69-80.

Severengiz, M., Roederb, I. Schindlerb, K, Seliger, G. (2018). Influence of Gaming Elements on Summative Assessment in Manufacturing Engineering *Education for sustainable Manufacturing. Procedia Manufacturing.* 2018(21), 429-437.

Slaughter, S., Leslie LL. (1997), *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*, The Johns Hopkins University Press, Baltimore.

Sousaa, M. J., Carmoa, M., Gonçalvesa, A. C., Cruza, R., Martins, J. M. (2019). Creating knowledge and entrepreneurial capacity for HE students with T digital education methodologies: Differences in the perceptions of students and entrepreneurs. *Journal of Business Research.* 2019(94), 227–240.

Seixas, L. R., Gomes, A. S., Melo Filho, I. J. (2016). Effectiveness of gamification in the engagement of students. *Computers in Human Behavior*. 2016(58), 48-63.

Strike, A.J. (2009). Academic staff's career pathway design in English pre-1992 universities: contemporary evolution or systematic de-construction of Homo Academicus? *Doctoral dissertation*. University of Southampton. 2009.

Troudt, E.E., Stuard, A.S., Winkler, C. (2017). An evolving Entrepreneurship Simulation as a Vehicle for Career and Technical Education, *New directions for Community Colleges.* 2017(178), 35-44.

Wintermeyer, A., Knautz, K. (2015). Information Literacy: Moving Toward Sustainability. *Communications in Computer and Information Science*, 552, 350–359. https://doi.org/10.1007/978-3-319-28197-1

Yen-Chun Jim Wu, Tienhua Wu, Yibin Li (2017). Impact of using classroom response systems on students' entrepreneurship learning experience. *Computers in Human Behavior*.2017(n.a.), 1-12.

Yildirim, I. (2017). The effects of gamification-based teaching practices on student achievement and students' attitudes toward lessons. *The Internet and Higher Education*, *33*, 86–92. https://doi.org/10.1016/j.iheduc.2017.02.002

Zaina, L. A. M., Álvaro, A. (2015). A design methodology for user-centered innovation in the software development area. *The Journal of Systems and Software*. 110 (2015), 155–177.

DECISION SCIENCES INSTITUTE

Chance-constrained Bi-objective Optimization Model for Hazardous Materials Distribution

Kamran S. Moghaddam
College of Business, Clayton State University, Morrow, GA 30260, USA
Email: KamranMoghaddam@clayton.edu

Farshid Azadian
College of Business, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA
Email: azadianf@erau.edu

ABSTRACT

In this research, a chance-constrained bi-objective optimization model for hazardous materials (hazmat) distribution problem is developed. The goal is to find optimal links and routes to obtain a trade-off between the safe and fast distribution of hazmat through a transport network under customers' demand and service time uncertainty along with unknown probabilities for hazmat incidents in a link-based model. We also develop a solution method based on a hybrid Monte-Carlo simulation and compromise programming to determine the optimal solutions. Computational results of a real case study are also provided to illustrate the effectiveness of the model and the solution method.

KEYWORDS: Hazardous material, vehicle routing and scheduling problem, demand uncertainty, service time uncertainty, chance-constrained programming, multi-objective optimization

INTRODUCTION

The inherent properties of Hazardous Materials (hazmats) are such that its production, warehousing, and transportation activities impose serious risks to both society and its surrounding environment. A large proportion of hazmats shipments that are transported in-land is carried by trucks. In the United States, about 60% of the tonnage of total hazmat shipments are reported to be transported by trucking industry (U.S. Department of Transportation, 2017). Commercial carriers in the United States transport over 3 billion tons of hazardous materials each year. Such hazmat shipments can be highly risky and highly sensitive and if improperly handled, labeled, or packaged could result in the loss of life, property damage, and harm to national security interests.

Compared to the number of shipments and the amount of hazmat transported through the transport networks, the number of hazmat incidents, fatalities, injuries and economic loss compared to the flow of hazmats are diminutive, approximately 10^{-6} per vehicle-miles traveled (Harwood *et al.*, 1993). As such, hazardous material transportation planning has been categorized as Low-Probability-High-Consequence (LPHC) problem. Although the number of hazmat transportation incidents is a very small fraction compared to the total number of shipments, due to nature of hazmats, there will be tremendous societal costs as well as environmental impacts in a case of an incident. These consequences make hazardous materials transportation, routing, and scheduling operations very important and sensitive logistical decisions.

Hazmat incidents are rare events as such estimating the likelihood of their occurrence based on historical data is challenging. Moreover, the road network is naturally a dynamic system. The changes in the existing roads, expansion of the network, and the development of new roads mean that reliable historical data may not be available to be used for estimating the likelihood of incidents. As such, we encounter a transportation network in which incident probabilities are unknown or uncertain. Therefore, unlike the majority of the traditional literature, this research focuses on unknown incident probabilities. Risk and time (or cost) are the two indispensable objective functions of the Hazardous materials Vehicle Routing and Scheduling Problem with Time Windows (HVRPTW) that govern both path choice and routing (Pradhananga *et al.*, 2014). Despite the practical significance, HVRPTW-related research is not widely practiced. The present study is an effort to fill in the gap of the literature and aims for the development of an efficient optimization model for the HVRPTW addressing the practical aspects of hazmat transportation.

The remainder of this paper is organized as follows: the relevant literature is reviewed in Section 2. The detailed problem description along with its difficulty aspects and assumptions is presented in Section 3, and it is mathematically formulated in Section 4. Computational results including the presentation of a case study, and the practicality of the optimal solutions are discussed in Section 5. Section 6 concludes the paper by highlighting the contributions and provides direction for future research.

LITERATURE REVIEW

The traditional approach to model the routing and/or scheduling decisions assume that incident probabilities are known. The alternative approach is that these probabilities are unknown. One of the early studies to deal with unknown probabilities of hazmat incidents is found in (Bell, 2007). In that study, a two-player non-cooperative game is considered between the network user seeking a path to minimize the expected trip cost and on the other hand, an “evil entity” choosing link performance scenarios to maximize the expected trip cost. The applications of the classic two-player non-cooperative game are also found in (Szeto, 2013) and (Szeto *et al.*, 2017). However, all of these studies used single-objective optimization models to formulate the two-player non-cooperative games. Other recent significant studies that employed single-objective mathematical models can be found in (Dadkar *et al.*, 2010), (Hu *et al.*, 2017), and (Kumar *et al.*, 2018).

Combined routing and scheduling of hazmat shipments under multiple objectives in networks that have time-dependent attributes has also been studied in the literature (Androutsopoulos and Zografos, 2010), and (Androutsopoulos and Zografos, 2012). Chang *et al.* (2005) developed a method for finding non-dominated routes for multiple routing objectives in transport networks in which the routing attributes are uncertain and the probability distributions that describe those attributes vary by time of day. Pradhananga *et al.* (2014) developed a Pareto-based bi-objective optimization of hazardous materials vehicle routing and scheduling problem with time windows for a realistic case study. Esfandeh *et al.* (2016) investigated dual-toll setting as a policy tool to mitigate the risk of hazmat shipment in road networks by formulating the dual-toll problem as a bi-level program in which the upper level aims at minimizing the risk, and the lower level explores the user equilibrium decision of the regular vehicles and hazmat carriers.

The hazardous materials distribution problem involves the determination of the sequence of deliveries and the corresponding paths assigned to each truck. Erkut and Alp (2007) considered an integrated routing and scheduling problem in hazardous materials transportation where accident rates, population exposure, and link durations on the network varied with time of day and tried to minimize risk in the form of accident probability. Ma *et al.* (2012) investigated vehicle routing problem with time windows and link capacity constraints to transport hazardous materials using a tabu search heuristic with an adaptive penalty mechanism (TSAP) to help manage company’s vehicle fleet. Bronfman *et al.* (2015), Bronfman *et al.* (2016) and Garrido and Bronfman (2017) studied the hazmat routing problem from an origin to a destination in an urban area in order to maximize the distance between the route and its closest vulnerable center, weighted by the center’s population.

Applications of stochastic models have also been found to the hazmat routing and scheduling problem. Jia *et al.* (2011) formulated a hazmat problem to minimize the risks from a natural accident and the uncertainty of a terrorist threat and developed a fuzzy-stochastic constraint programming method to deal with multiple uncertainties. Desai and Lim (2013) employed a Stochastic Dynamic Programming (SDP) approach to solving the hazmat route selection problem to determine the optimal routing policies in a stochastic dynamic network. In order to expedite the process of obtaining optimal routing policies they proposed three techniques for pruning stochastic dynamic networks as (1) use of static upper/lower bounds, (2) pre-processing the stochastic dynamic networks by using the start time and origin location of the vehicle, and (3) a mix of pre-processing and upper/lower bounds. Pamucar *et al.* (2016) proposed a new approach for cost and risk assessment in the multi-objective selection of routes for the transport of hazardous materials on a network of city roads based on the application of an Adaptive Neuro-Fuzzy Inference System (ANFIS).

PROBLEM DESCRIPTION

The proposed integrated routing and scheduling problem involves three underlying definitions: loss, expected loss, and total expected loss. Loss (expected loss) is defined as the number (expected number) of people affected in the event of accidents. Expected loss can be defined on the link, route, and network levels. The expected loss on a link (route) is the expected number of people affected in the event of accidents on that link (or route). The sum of the expected loss on all links or routes gives the total expected loss in the network. For each link *ij*, $loss_{ij}$ is measured by the population inside a circle of given impact radius, centered at any point on link *ij*, in the event of an accident from node *i* to node *j*. The impact radius would depend on the hazardous material under consideration, and can vary from several feet to several miles (Erkut and Verter, 1998).

Link failure is usually rare, so as previously mentioned there is likely to be insufficient data from which to estimate failure probabilities. Hence q_{ij}, the conditional probability of an incident on link *ij*, is unknown. We assume that the dispatcher plans on the basis of one link failure. Multiple link failures are excluded from consideration in this study as they are too unlikely. However, they can be considered in a case of an extensive natural disaster such as earthquake, flood, or hurricane. Aside from the total expected loss, the dispatcher is interested in finding the shortest paths in the distribution network that can start from any of the origins and end at any of the destinations. So the dispatcher's problem involves in finding the safest links and routes with regard to the risk of population exposure and simultaneously determining the shortest paths within the distribution network.

The necessary underlying assumptions are listed as follows:

- The numbers and locations of the depots and customers are predetermined;
- The capacities of depots are known and deterministic while the demand and service time of customers are associated with uncertainty;
- Each customer can be served by multiple depots and multiple vehicles;
- The quantity of total supplies is greater than the quantity of total demands;
- Customers' demand must be serviced within pre-specified time windows and waiting due to early arrival is possible, while late arrival is not allowed at all;
- All vehicles must start and end their routes at the depots. The time of starting and ending the operations of the vehicle must be within the time windows specified at the depot;
- The probability of an incident on a link is unknown;
- The probability of using a link is unknown;

- Simultaneous hazmat incidents will not happen. The reason is that a hazmat incident is a low-probability event. Naturally, multiple link failures are very unlikely.

PROBLEM FORMULATION

Notation

In this section, we introduce the notation used in the paper to formulate the problem mathematically.

Sets

N	set of nodes in the network *G*
A	set of links in the network *G*
O	set of origin nodes in the network *G*
D	set of destination nodes in the network *G*
F	set of objective functions

Indices

i	Index for node *i* in the network *G*, *i* = 1, …, *N*
j	Index for node *j* in the network *G*, *j* = 1, …, *N*
k	Index for node *k* in the network *G*, *k* = 1, …, *N*

Parameters

$Loss_{ij}$	Loss, or population exposure, is defined as the number of people impacted in the event of an accident on link *ij*, $i \in N, j \in N$
$Dist_{ij}$	Travel distance (or time) from node *i* to node *j*, $i \in N, j \in N$
Dem_i	Demand for hazmat at node *i*, $i \in D$, random variable with mean μ_{Dem_i}, standard deviation σ_{Dem_i}, and the cumulative probability distribution of F_{Dem_i}
Tme_i	Earliest allowed arrival time at node *i*, $i \in D$
Tml_i	Latest allowed arrival time at node *i*, $i \in D$
Tmv_i	Stop or visit (i.e., service) time at node *i*, $i \in D$, random variable with mean μ_{Tmv_i}, standard deviation σ_{Tmv_i}, and the cumulative probability distribution of F_{Tmv_i}
Tmpm	Travel time per mile
Tmax	Maximum time for any trip
Dmax	Maximum distance for any trip
Vcap	Capacity of vehicles
Maxveh	Maximum number of vehicles
ε_{Dem}	Pre-determined satisfaction level of probabilistic demand constraints
ε_{Tmv}	Pre-determined satisfaction level of probabilistic service time constraints

Decision variables

x_{ij}	Binary variable equals to 1 if a vehicle travels from node *i* to node *j*, $i \in N, j \in N$; otherwise it is equal to 0.
p_{ij}	Probability of link *ij* selected for shipment, $i \in N, j \in N$
q_{ij}	Conditional probability of an incident on link *ij* given that an incident occurs, $i \in N, j \in N$
u_i	Accumulated deliveries of hazmat at node *i*, $i \in D$
td_i	Total travel distance for node *i*, $i \in D$
tma_i	Arrival time at node *i*, $i \in D$

Development of Mathematical Models

The routing and scheduling problem described above is a bi-objective decision problem. The first objective is to minimization transport risk associated with the unknown probability of incidents and links selection. In order to formulate the objective function, we employ a game theoretic approach in which it is assumed there is a two-player game between a dispatcher, who seeks a least cost tour, and nature (e.g., demon), with the power to cause one link to fail. This zero-sum, non-cooperative, two-player game mimics the mental process of a risk-averse dispatcher seeking the best route by considering a sequence of worst-case scenarios and his responses in each case. As *m* tends to infinity, the link use and failure probabilities tend to the mixed strategy Nash equilibrium as shown in the following bi-level *minimax* objective function.

$$Min_p\ (Max_q(\sum_{i=1}^{N}\sum_{j=1}^{N} Loss_{ij}.p_{ij}\ .q_{ij})) \tag{1}$$

Additionally, the second objective function tries to minimize the total travel distance determining the shortest paths within the distribution network as formulated in equation (2).

$$Min\ f_2 = \sum_{i=1}^{N}\sum_{j=1}^{N} Dist_{ij}.x_{ij} \tag{2}$$

In order to complete the mixed strategy Nash equilibrium, equation (3) indicates that there is only one link failure in the network. The risk-averse dispatcher, in general, will not wish to use one route but rather a mix of links so generally ($\sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij} \geq 1$). Equations (4)-(6) impose the requirement of using a mixed strategy for link selection probabilities at the origin, destination, and intermediate points.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} q_{ij} = 1 \tag{3}$$

$$\sum_{j=1, j\neq i}^{N} p_{ij} = 1 \quad \forall\ i \in O \tag{4}$$

$$\sum_{i=1, i\neq j}^{N} p_{ij} = 1 \quad \forall\ j \in D \tag{5}$$

$$\sum_{i=1, i\neq k}^{N} p_{ik} - \sum_{j=1, j\neq k}^{N} p_{kj} = 0 \quad \forall\ k \notin O, k \notin D \tag{6}$$

Equation (7) requires that inbound and outbound transported hazmat for intermediate nodes to be equal. In order to mandate that a link is available to transport hazmat whenever the link is included in a mixed strategy (its probability of selection is greater than zero), constraint (8) uses a very large number, *M*, enforcing the equivalent if-then-constraint is met. In addition, equation (9) ensures that a vehicle does not travel within the same node.

$$\sum_{i=1, i\neq k}^{N} x_{ik} - \sum_{j=1, j\neq k}^{N} x_{kj} = 0 \quad \forall\, k \notin O, k \notin D \tag{7}$$
$$x_{ij} \leq M.p_{ij} \quad \forall\, i \in N, j \in N \tag{8}$$
$$x_{ii} = 0 \quad \forall\, i \in N \tag{9}$$

Equation (10) requires that a vehicle can enter a destination node *k* from node *i* if it comes either from a depot that indicates it has enough supply to provide or from a destination node in which the total demand at both nodes does not exceed the vehicle's capacity. Similarly, equation (11) enforces that a vehicle must leave node *k* after service to a node *j* with the condition of the next node is either a depot or the vehicle has enough supply to serve the next demand node.

$$\sum_{\substack{i=1, i\neq k,\\ i\in O \text{ or } Dem_i + Dem_k \leq Vcap}}^{N} x_{ik} = 1 \quad \forall\, k \in D \tag{10}$$
$$\sum_{\substack{j=1, i\neq k,\\ j\in O \text{ or } Dem_j + Dem_k \leq Vcap}}^{N} x_{kj} = 1 \quad \forall\, k \in D \tag{11}$$

Constraint (12) restricts the probability of the accumulated hazmat deliveries at destination points to be greater or equal to the demand and less than or equal to the vehicles' capacity. If destination node *k* is visited after destination node *i*, then the probability of the difference of accumulated hazmat delivered at each node, $u_k - u_i$, should be bounded to the vehicle's capacity and fulfilled stochastic demand as formulated in constraint (13). On the other hand, if destination node *k* is the first stop after the depot then the probability of the entire demand for the node can be fulfilled, $u_k = Dem_i$ by constraint (14) otherwise, this probability of the demand will be provided by multiple vehicles coming from other nodes as formulated by constraint (15).

$$P(Dem_k \leq u_k \leq Vcap) \geq \varepsilon_{Dem} \quad \forall\, k \in D \tag{12}$$
$$P(u_k \geq u_i + Dem_k - Vcap + Vcap.(x_{ki} + x_{ik}) - (Dem_k + Dem_i)x_{ki}) \geq \varepsilon_{Dem} \quad \forall\, i \notin O, k \in D, i \neq k \tag{13}$$
$$P(u_k \leq Vcap - (Vcap - Dem_k)x_{ik}) \geq \varepsilon_{Dem} \quad \forall\, i \in O, k \in D \tag{14}$$
$$P\left(u_k \geq Dem_k + \sum_{i=1, i\notin O}^{N} Dem_i.x_{ik}\right) \geq \varepsilon_{Dem} \quad \forall\, k \in D \tag{15}$$

Constraints (16) and (17) calculate the total travel distance from depots to the destinations nodes and from each destination node to other destination nodes. Constraint (18) restricts the total distance traveled by each vehicle from depots to be less than maximum distance allowed on a trip.

$$td_j \geq Dist_{ij}.x_{ij} \quad \forall\, i \in O, j \in D \tag{16}$$
$$td_j \geq td_i + Dist_{ij}.x_{ij} - Dmax(1 - x_{ij}) \quad \forall\, i \notin O, j \in D \tag{17}$$
$$td_i \leq Dmax \quad \forall\, i \in O \tag{18}$$

Constraints (19)-(21) represent the time window requirements for the arrival and departure of vehicles to the destination nodes. Inequality (19) defines the probability of the arrival time of a vehicle at node *k* after it visited or served previous destination node *i* to be greater than or equal to the latest allowed arrival time at node *i* plus the service time at node *i* and travel time required to go from node *i* to *k*. Constraint (20) restricts the decision variable of arrival time to be within allowed early and late arrivals in which a vehicle is allowed to wait in order to arrive no earlier than early arrival time. Constraint (21) enforces the probability of the triangle inequality requirement of maximum trip time for arrival time, service time, and travel time between the origin points and destination points in the network to be greater than a specified satisfaction level.

$$P(tma_k \geq tma_i + (Tmv_i + Tmpm.Dist_{ik})x_{ik} - Tml_i(1 - x_{ik})) \geq \varepsilon_{Tmv} \quad \forall\, k \in D, i \neq k \tag{19}$$
$$Tme_k \leq tma_k \leq Tml_k \quad \forall\, k \in D \tag{20}$$
$$P(tma_k + Tmv_k + Tmpm.Dist_{ki}.x_{ki} \leq Tmax) \geq \varepsilon_{Tmv} \quad \forall\, i \in O, k \in D \tag{21}$$

In order to calculate the minimum number of vehicles to transport the hazmat from depots to the destination locations, $Vehf$, the following probability inequality (22) is formulated. Since $Vehf$ may be fractional in value so it will be rounded to the next integer using equation (23).

$$Vehf \geq P\left((\sum_{i=1, i \notin O}^{N} Dem_i)/Vcap \right) \tag{22}$$
$$Vehr = int(Vehf + 0.999) \tag{23}$$

The dispatcher is required to send enough vehicles from the depots to the destination nodes to ensure that demand for hazmat will be fulfilled by the coming vehicles. This requirement is formulated by constraint (24). Furthermore, constraint (25) requires that the number of dispatched vehicles is always less than the maximum number of vehicles available.

$$\sum_{i=1, i \in O}^{N} \sum_{j=1, j \in D}^{N} x_{ij} \geq Vehr \tag{24}$$
$$\sum_{i=1, i \in O}^{N} \sum_{j=1, j \in D}^{N} x_{ij} \leq Maxveh \tag{25}$$

Finally, constraints (26) and (27) restrict the probability of link selection, p_{ij}, and the conditional probability of link failure, q_{ij}, to be in the range of [0,1]. Constraints (28) and (29) mandate the requirement for decision variables, x_{ij}, to be binary variables and the decision variables $u_i, td_i, \text{and } tma_i$ to be positive.

$$0 \leq p_{ij} \leq 1 \quad \forall\, i \in N, j \in N \tag{26}$$
$$0 \leq q_{ij} \leq 1 \quad \forall\, i \in N, j \in N \tag{27}$$
$$x_{ij} \in \{0, 1\} \quad \forall\, i \in N, j \in N \tag{28}$$
$$u_i, td_i, tma_i \geq 0 \quad \forall\, i \in N, j \in N \tag{29}$$

M is a very large number

CASE STUDY AND COMPUTATIONAL RESULTS

Network and Data Setting

Chance-constrained bi-objective hazmat routing and scheduling is a new problem and no test instances have been developed for the problem yet. Therefore, we have adopted a real distribution problem encountered by a chemical plant and designed a new test problem capturing our new stochastic features to demonstrate an application of the developed model and the solution method. The case study problem is inspired by a distribution of sulfuric acid by our industry collaborator from a chemical plant in Chicago, IL (labeled in red) for industrial use in 11 major US cities (labeled in blue) located in the Midwest and the West Coast as shown in Figure 1. Sulfuric acid is an important chemical commodity, and it has been suggested that a nation's sulfuric acid production is a good indicator of its industrial strength.

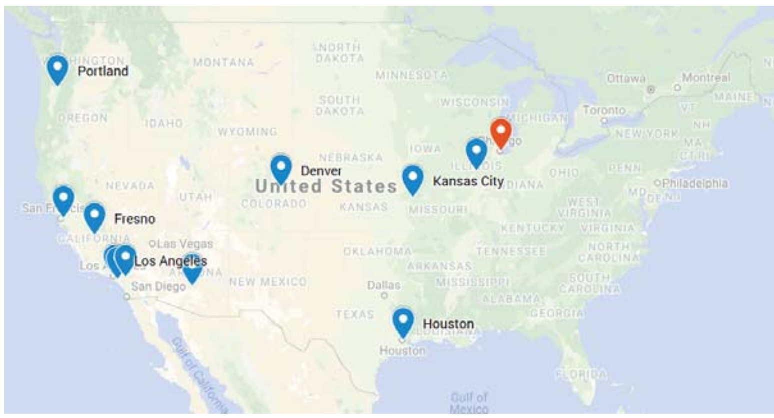


Figure 1. Location of the chemical plant (red label) and industrial facilities (blue labels)

The transportation network has 12 nodes and 66 bi-directional links which is equivalent to 132 unidirectional links (inside-the-city travel links are eliminated).The population of the cities was extracted from the estimates calculated by the United States Census for 2018. Demand for sulfuric acid at the industrial facilities is calculated based on the average consumption rate of the facilities. It also is assumed that there will be enough semi-truck tankers each with the capacity of 18,000 liters. Table 1 demonstrates the estimated population in cities, average and standard deviation of demand and service time, and the requirements for the earliest and latest allowed arrival time to the cities. All parameters are chosen in a realistic order of magnitude.

Table 1. Population and logistics data of the nodes in the distribution network

Node	City	Population	Average Demand (10000 L)	SD of Demand (10000 L)	Average TMV (hr.)	SD of TMV (hr.)	TME (hr.)	TML (hr.)
1	Chicago	9,533,040	0	0.00	0.00	0.000	0.00	1666.65
2	Denver	2,888,227	6	0.60	0.17	0.028	16.67	41.67
3	Fresno	972,297	3	0.33	0.20	0.040	46.67	48.33
4	Houston	6,313,158	7	1.68	0.17	0.040	33.33	50.00
5	Kansas City	2,159,159	7	1.05	0.17	0.030	30.00	65.00
6	Los Angeles	13,131,431	18	3.06	0.23	0.047	60.00	65.00
7	Oakland	425,195	4	0.48	0.17	0.037	58.33	65.00
8	Anaheim	352,497	5	0.50	0.17	0.017	40.00	63.33
9	Peoria	373,590	2	0.38	0.17	0.027	1.67	13.33
10	Phoenix	4,737,270	6	0.90	0.17	0.030	25.00	40.00
11	Portland	2,389,228	7	0.84	0.18	0.044	30.00	46.67
12	Riverside	303,871	2	0.28	0.20	0.036	31.67	63.33

In order to calculate the population loss in the distribution network, it is assumed that only a small fraction of the population would be impacted depending on how close and dense the population is scattered around the network's links. We used a gravity-based method to estimate the impacted population (i.e., population loss) in case of a catastrophe event. The travel distance between the cities was extracted from Google Maps.

For computational purposes, we can use either normal distributions or triangular distributions to represent demand and service time uncertainty. The reason to choose these two distributions in this study is that they are easy to understand and it is intuitive for practitioners to describe the uncertain demand and service times in these forms. The normal distribution is a common assumption and is easily described by two parameters, the mean and the variance. Like the well-known beta distribution used in PERT/CPM, the triangular distribution allows practitioners to describe the random demand and random service time in a limited region with three parameters; the minimum value, the most likely value, and the maximum value.

The average speed for semi-truck is assumed to be 50 mph (i.e., travel time per mile is 1.2 min/mile) and the maximum distance allowed on a trip for a vehicle is 5000 miles. Finally, we set $\varepsilon_{Dem} = \varepsilon_{Tmv}$ during the solution process. It is possible to explore the solutions with different ε_{Dem} and ε_{Tmv} values. However, the system's reliability level, described by different ε_{Dem} and ε_{Tmv} levels, is less meaningful for practitioners to interpret. We set $\varepsilon_{Dem} = \varepsilon_{Tmv} = 0.90$ to find the ideal solutions with respect to each objective function and Pareto optimal solutions for that level of reliability. LINGO optimization software (LINGO 18.0, unlimited version) is utilized to solve the compromise programming sub-model in the algorithm.

Properties of the Optimal Solutions for Individual Objective Functions

Table 2 shows the optimal routes and schedules with respect to the first objective function independently while ignoring the other objective function subject to the functional and operational constraints. It can be observed that the optimal solution for objective function (1) with the lowest possible value for the total expected loss, number of people affected, is 946 while the total travel distance is 29073 miles. In this case, it can be stated that the dispatcher tends to diversify the link selection probabilities as much as possible in order to reduce the negative effect of potential hazmat spillage due to possible link failures. It is found that the optimal solution consists of 58 (out of 132) candidate links with non-zero probability indicating the availability of a large selection poll. In addition, if a link has zero selection probability, it will not be advisable to use that link to transport any amount of hazmat. According to this optimal scenario, 20 candidate links are finally selected to transport sulfuric acid from the chemical plant to the industrial facilities. The modeling approach also enables the dispatcher to choose different vehicles to distribute hazmat shipments. Table 2 also illustrates that there are nine vehicles required through 20 different routes for safest distribution of sulfuric acid in the network while meeting demand and service time requirements.

Table 2. Optimal routes for the objective function of total expected loss ($f_{1,min}$ = 946 people and f_2 = 29073 miles)

Vehicle	Routes	Arrival Time (hr.)
1	Chicago → Denver	19.9
	Denver → Chicago	40.0
2	Chicago → Kansas City	30.0
	Kansas City → Houston	44.4
	Houston → Chicago	65.9
3	Chicago → Los Angeles	60.0
	Los Angeles → Chicago	101.3

Vehicle	Routes	Arrival Time (hr.)
4	Chicago → Oakland	65.0
	Oakland → Chicago	107.8
5	Chicago → Anaheim	61.5
	Anaheim → Chicago	102.7
6	Chicago → Peoria	3.0
	Peoria → Chicago	6.2
7	Chicago → Phoenix	34.3
	Phoenix → Chicago	68.7
8	Chicago → Portland	41.7
	Portland → Chicago	83.5
9	Chicago → Riverside	40.1
	Riverside → Fresno	46.7
	Fresno → Chicago	90.1

Table 3 shows the optimal routes and schedules with respect to second objective function independently while ignoring the first objective function (total expected loss) subject to the functional and operational constraints. The optimal solution with respect to the total travel distance results to the minimum 15,997 miles of travel but it affects 125,181 people of the population. Under this optimal scenario, the travel distance can be improved by 45% at most, while the total expected loss is increased by 132 times. The link selection probabilities in this situation reveal a different pattern as the optimal solution enforces the dispatcher against diversification as there are only 38 (out of 132) candidate links with non-zero probability. However, only 15 candidate links are finally selected to transport sulfuric acid from the chemical plant to the industrial facilities. Table 3 indicates that to meet the customers demand and their required service time restrictions only 4 vehicles are required through 15 different routes for the fastest distribution of sulfuric acid in the network. This pattern of optimal solution indicates that if the dispatcher goal is to only minimize the total travel distance, the optimal solution narrows down the dispatcher's choice to fewer candidate links from the distribution network, but these links will have a much higher likelihood for selection resulting to a significant increase of impacted population.

Table 3. Optimal routes for the objective function of total travel distance (f_1 = 125181 people and $f_{2,min}$ = 15997 miles)

Vehicle	Routes	Arrival Time (hr.)
1	Chicago → Denver	19.9
	Denver → Portland	46.7
	Portland → Oakland	65.0
	Oakland → Chicago	107.8
2	Chicago → Los Angeles	60.0
	Los Angeles → Chicago	101.3
3	Chicago → Peoria	3.0
	Peoria → Kansas City	30.0
	Kansas City → Houston	44.4
	Houston → Chicago	65.9
4	Chicago → Phoenix	34.5
	Phoenix → Fresno	48.3
	Fresno → Anaheim	53.5
	Anaheim → Riverside	60.1
	Riverside → Chicago	100.4

CONCLUSIONS AND FUTURE RESEARCH

This research developed an integrated routing and scheduling hazmat transportation problem with multiple supply and demand locations and unknown link incident probabilities. A new chance-constrained bi-objective network optimization model was formulated to minimize the total expected loss (based on the risk of population exposure) and the total travel distance for a hazmat distribution problem with time windows in which uncertain demand and service times are captured by well-known probability distributions. In order to illustrate the effectiveness of the developed mathematical model and the solution method in obtaining Pareto-optimal solutions, a real case study was solved and analyzed. For future research, exact and heuristic approaches that include novel mathematical programming components can be examined especially for very large-scale problems. In addition, possible extensions can be made to the window reduction and partition approach originally developed for the multiple traveling-salesman problems with time windows. Incorporating of some more realistic constraints (e.g., special-line transportation, temporary warehousing, fatigue driving, and government rules of banning road accessibility, among others) may also be taken into account to reflect more practical aspects.

REFERENCES

Androutsopoulos, K. N., Zografos, K. G., (2010) Solving the bicriterion routing and scheduling problem for hazardous materials distribution, *Transportation Research Part C: Emerging Technologies*, Vol. 18, No. 5, pp. 713-726.

Androutsopoulos, K. N., Zografos, K. G., (2012) A bi-objective time-dependent vehicle routing and scheduling problem for hazardous materials distribution, *EURO Journal on Transportation and Logistics*, Vol. 1, No. 1, pp. 157-183.

Bell, M. G. H., (2007) Mixed routing strategies for hazardous materials: decision-making under complete uncertainty, *International Journal of Sustainable Transportation*, Vol. 1, No. 2, pp. 133-142.

Bronfman, A., Marianov, V., Paredes-Belmar, G., A., L.-V., (2016) The maxisum and maximin-maxisum HAZMAT routing problems, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 93, pp. 316-333.

Bronfman, A., Marianov, V., Paredes-Belmar, G., Lüer-Villagra, A., (2015) The maximin HAZMAT routing problem, *European Journal of Operational Research*, Vol. 241, No. 1, pp. 15-27.

Chang, T.-S., Nozick, L. K., Turnquist, M. A., (2005) Multiobjective path finding in stochastic dynamic networks, with application to routing hazardous materials shipments, *Transportation Science*, Vol. 39, No. 3, pp. 383-399.

Dadkar, Y., Nozick, L., Jones, D., (2010) Optimizing facility use restrictions for the movement of hazardous materials, *Transportation Research Part B: Methodological*, Vol. 44, No. 2, pp. 267-281.

Desai, S., Lim, G. J., (2013) Solution time reduction techniques of a stochastic dynamic programming approach for hazardous material route selection problem, *Computers & Industrial Engineering*, Vol. 65, No. 4, pp. 634-645.

Erkut, E., Alp, O., (2007) Integrated routing and scheduling of hazmat trucks with stops en route, *Transportation Science*, Vol. 41, No. 1, pp. 107-122.

Erkut, E., Verter, V., (1998) Modeling of transport risk for hazardous materials, *Operations Research*, Vol. 46, pp. 625–642.

Esfandeh, T., Kwon, C., Batta, R., (2016) Regulating hazardous materials transportation by dual toll pricing, *Transportation Research Part B: Methodological*, Vol. 83, pp. 20-35.

Garrido, R. A., Bronfman, A., (2017) Equity and social acceptability in multiple hazardous materials routing through urban areas, *Transportation Research Part A: Policy and Practice*, Vol. 102, pp. 244-260.

Harwood, D. W., Viner, J. G., Russell, E. R., (1993) Procedure for developing truck accident and release rates for hazmat routing, *Journal of Transportation Engineering*, Vol. 119, No. 1, pp. 189-199.

Hu, H., Guo, S., Ma, H., Li, J., Li, X., (2017) A credibilistic mixed integer programming model for time-dependent hazardous materials vehicle routing problem, *Journal of Uncertain Systems*, Vol. 11, No. 3, pp. 163-175.

Jia, H., Zhang, L., Lou, X., Cao, H., (2011) A fuzzy-stochastic constraint programming model for hazmat road transportation considering terrorism attacking, *Systems Engineering Procedia*, Vol. 1, pp. 130-136.

Kumar, A., Roy, D., Verter, V., Sharma, D., (2018) Integrated fleet mix and routing decision for hazmat transportation: A developing country perspective, *European Journal of Operational Research*, Vol. 264, pp. 225-238.

Ma, H., Cheang, B., Lim, A., Zhang, L., Zhu, Y., (2012) An investigation into the vehicle routing problem with time windows and link capacity constraints, *Omega*, Vol. 40, No. 3, pp. 336-347.

Pamucar, D., Ljubojevic, S., Kostadinovic, D., Dorovic, B., (2016) Cost and risk aggregation in multi-objective route planning for hazardous materials transportation - A neuro-fuzzy and artificial bee colony approach, *Expert Systems With Applications*, Vol. 65, No. 1-15.

Pradhananga, P., Taniguchi, E., Yamada, T., Qureshi, A. G., (2014) Bi-objective decision support system for routing and scheduling of hazardous materials, *Socio-Economic Planning Sciences*, Vol. 48, No. 2, pp. 135-148.

Szeto, W. Y., (2013) Routing and scheduling hazardous material shipments: Nash game approach, *Transportmetrica B: Transport Dynamics*, Vol. 1, No. 3, pp. 237-260.

Szeto, W. Y., Farahani, R. Z., Sumalee, A., (2017) Link-based multi-class hazmat routing-scheduling problem: A multiple demon approach, *European Journal of Operational Research*, Vol. 261, No. 1, pp. 337-354.

U.S. Department of Transportation. (2017) Freight Facts and Figures 2017. In. Washington DC: Bureau of Transportation Statistics.

DECISION SCIENCES INSTITUTE

Anticipated Regret and Consumers' Smartphone Replacement Decisions——The Moderation of Consumers' Personal Traits

Yuying Zheng
National Chengchi University
Email: 107351503@nccu.edu.tw

Chien-Wei Chen
National Chengchi University
Email: cweichen@nccu.edu.tw

ABSTRACT

As a necessity of life, smartphones have become rapidly popular in China in recent years. Chinese consumers' never-ending demands for superior function, performance and appearance push them to replace smartphones more frequently, bringing in huge market opportunities for smartphone marketers. In contrast to the previous studies that focus on cognition, this research takes an affective approach to consumer's decisions of replacing consumer and analyzes the impacts of consumers' anticipated regret on their smartphone replacement decisions based on empirical data collected in China. The respective moderating effects of consumers' novelty seeking and need for uniqueness are subsequently examined.

KEYWORDS: Anticipated Regret, Intention to Upgrade, Intention to Switch Brands, Novelty Seeking, Need for Uniqueness, Moderation

INTRODUCTION

Nowadays, smartphones have become an indispensable part of people's life. According to a report quoted by US technology media Mashable, the smart phone penetration rate in China was 58% in 2016 and the figure reached 89% in 2018. Zenith's research report in 2017 showed that the number of smartphone users in China would reach 1.3 billion in 2018, ranking first in the world. As people become more and more dependent on smartphones, they become more demanding for functions, specifications and other attributes. As a result, people's consumption of smartphones is rapidly increasing, and the frequency of replacement is getting faster. With the increasing market saturation, the business competition becomes gradually fierce. Therefore, the decisions of consumers on smartphone replacement have become the focus of business and a challenge for all smartphone manufacturers. According to the "2017-2022 China Smart Phone Market In-depth Evaluation and Development Trend Research Report", consumers are not as sensitive to the monetary cost of replacing smartphones as they used to be. The frequency of smartphone replacement among consumers remains high, with 32% of consumers willing to replace their phones within 6 months.

For smartphone manufacturers, the huge smartphone replacement market has become a major battleground. There is a profound difference in marketing for manufacturers between attracting a new consumer and persuading an existing consumer to replace his/her product. Nevertheless,

the previous literature makes no distinction between the first purchase decision and the replacement decision in discussing consumer decision of purchasing smartphones. Most of existing studies address the cognitive aspect of consumer replacement and upgrade decisions (cf. Bellezza, Ackerman, & Gino, 2017). Fewer studies that discuss the replacement decision of consumers are mainly from the perspective of affective psychology, ignoring the fact that consumers may not be completely rational ones and they make decisions with affect. A direct relationship between affect and action may appear obvious (Frijda, 2010). Actually, affective judgments are fairly independent of, and even precede, the perceptual and cognitive activities in consumer behavior (Zajonc, 1980). It is evident that consumers' affect plays an important role in consumer behavior. For instance, Westbrook (1987) indicated that consumers' affect is highly associated with post-purchase behaviors such as complaint and word of mouth as well as with purchase intention (Ruth, Brunel, & Otnes, 2002). Affect has not been fully empirically studied and discussed in the durable replacement literature. This research incorporates the affective psychology of consumers, i.e. anticipated regret, into the analysis of consumer replacement decision. Specifically, it hopes to explore the impact of consumers' anticipated regret on their product upgrading and brand switching decisions. Novelty seeking and need for uniqueness are selected as moderators because these two factors are the most important consumer traits that may affect the affect- behavior linkage.
The main research objectives are summarized as follows:

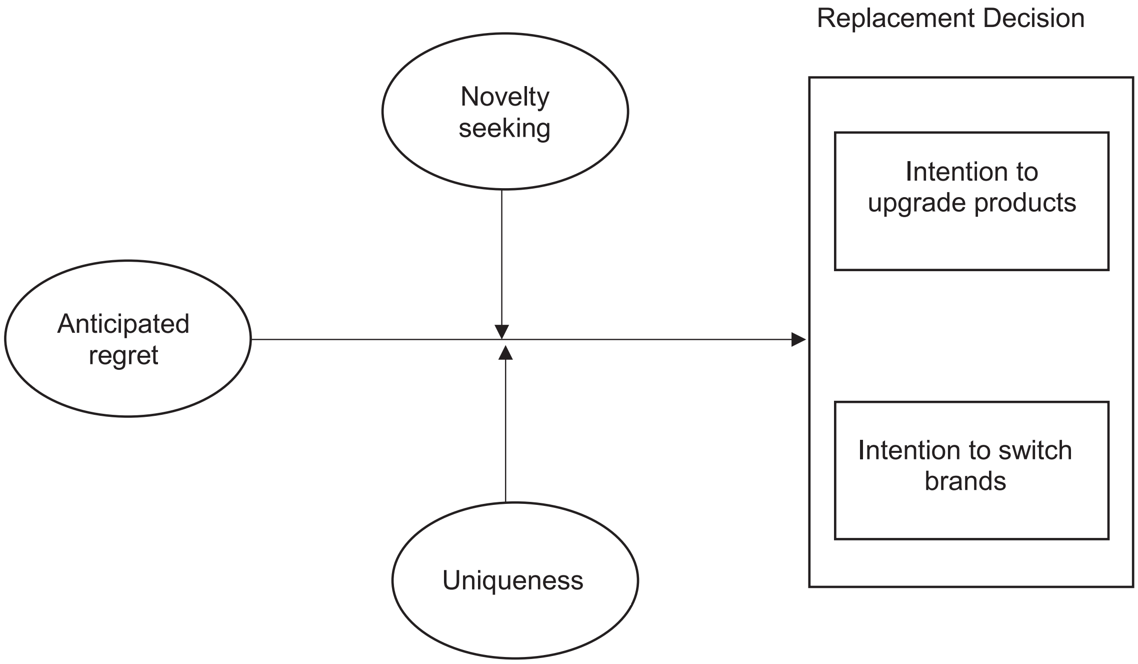
1. To investigate the impact of consumers' anticipated regret on smartphone upgrade and brand switching decisions.
2. To explore the moderating effects of consumers' novelty seeking and uniqueness on the relationship between anticipated regret and replacement decisions.

Empirically, it is hoped that this study can provide guidance for the marketing of smartphone.

LITERATURE REVIEW AND THEORETICAL DEVELOPMENT/MODEL

A research framework is therefore proposed, as shown in Figure 1, to analyze the relationships between consumers' anticipated regret and their intention to upgrade and switch brands. The relationships are expectedly moderated by novelty seeking and need for uniqueness. According to Dickson (1982), the interactionist perspective is the theoretical basis of the moderating effects, and it advocates that consumers' behaviors depend on the context, the individual differences and the interaction between them. These two moderators are used to reflect the contingency about the heterogeneity of consumers. Novelty seeking (Tse & Crotts, 2005) and need for uniqueness (Snyder & Fromkin, 1977) are both consumer traits that are the manifestations of individual differences and may have moderating effects on consumers' behaviors. At the same time, in order to avoid other confounding influences on dependent variables, this study uses consumer satisfaction as a control variable when analyzing the proposed direct and moderating effects.

Figure 1. Research Framework: Anticipated regret and the Intention to upgrade and to switch brands



Anticipated Regret and the Intention to Upgrade and Switch Brands

In the context of consumer purchase decision making, regret will not only occur when purchase is realized but may also appear before purchase. Researchers in psychology have found that, when people face counterfactual comparisons, they learn to anticipate the regrets they will have and try to minimize this kind of future regret in various ways (Meller et al., 1999; Zeelenberg & Beattie, 1997). This kind of regret of pre-purchase illusion is called anticipated regret, as opposed to experienced regret, which is felt through actual comparison after purchase. Janis and Mann (1977) used anticipated regret as an allusion to the main psychological effect of various worries about loss, which is, the materialized loss that troubles decision-makers before their decisions. Anticipated regret is usually divided into upward anticipated regret and downward anticipated regret (Roese, 1997). The upward anticipated regret usually restrains consumers' emotions when making their purchase decisions, which accordingly increases their preferences for conservative purchase. In his research on the decision-making behavior of consumers, Simonson (1992) found that after knowing that they made a wrong decision, consumers would like to buy guaranteed products and try to avoid regret in the future. For consumers, product upgrade decisions usually represent better performances and higher specifications of new products, which is a more secure conservative consumption. Therefore, consumers with stronger anticipated regrets are more willing to upgrade their products.

H1a: The consumers' intention to upgrade should increase as their anticipated regret increases.

In the process of brand switching, consumers will be faced with such problems as conversion cost and search cost, etc. These problems enable consumers to perceive the risks of brand switching and thus to reduce their intention of brand switching (Colomb & Morrison, 1989). When the anticipated regret of consumers is stronger, consumers are more reluctant to make risk-taking choices, thereby reducing the intention of brand switching in the face of the risks brought by brand switching.

H1b: The consumers' intention to switch brands should decrease as their anticipated regret increases.

Moderation of Novelty Seeking

Tse and Crotts (2005) indicates that human curiosity is a strong internal force that can drive people to explore the world around them and the changes among individuals. Curiosity is one of the main concepts of behavioral motivation, which can drive people to learn, execute, verify, explore and experience. In some literature related to behavioral science, novelty seeking is also regarded as a kind of curiosity drive, perceptual pursuit and exploration drive. Novelty is an inherent psychological trait of consumers. A logical explanation for a large difference in people's purchase behavior mainly comes from the individual's different desires for novelty (Bello & Etzel, 1985). Therefore, the hypothesis is put forward:

H2a: The positive relationship between anticipated regret and intention to upgrade increases as novelty seeking increases.

Novelty seeking is a key motive underlying the purchase of smartphones. Consumers are likely to buy them because they want to experience things different and innovative, satisfying their desire for novelty and excitement (Lee & Crompton, 1992). According to the relationship between novelty and exploratory behavior proposed by Berlyne (1966), the exploration behavior is a manifestation of the individual's curiosity awakened by the novel environment. The more time a person likes to spend in exploring novelty and excitement, the less time he will spend in those have been experienced before. That is, the stimulation of novelty heightens the exploration behavior. Therefore, the hypothesis is put forward:

H2b: The negative relationship between anticipated regret and intention to switch brands increases as novelty seeking increases.

Moderation of Need for Uniqueness

Snyder and Fromkin (1977) found that, when individuals feel that their own uniqueness is threatened, there will be a need to differentiate themselves from others and it will compete with other demands. In short, individuals want to be able to show others what they are different from others or beyond others in the real world. This kind of demand is called need for uniqueness. Uniqueness has various patterns and means of expression. For instance, individuals can manifest their uniqueness by displaying their belongings (Belk, 1988), by expressing their own interpersonal interaction style (Maslach et al., 1985), or by demonstrating their own knowledge and expertise in a certain field (Holt, 1995). As one of the important possessions of consumers, the upgrade decision of smartphone can reflect the consumers' professional knowledge and expertise in the field of electronic technology to a certain extent. So the hypothesis is:

H3a: Uniqueness strengthens the positive relationship between anticipated regret and intention to upgrade.

The theory of uniqueness (Snyder 1992) indicates that, under the premise of avoiding serious social punishment due to violation of social norms, some forms of material expression can satisfy the individual's need to express his uniqueness. Some customer behaviors, e.g., the purchase of certain products, showing others their unique consumption characteristics, have become a typical manifestation of individual in the context of the market economy. Burns and Warren (1995) also illustrated in their study that, consumers with higher need for uniqueness would choose shopping mall far away from home in order to purchase goods different from their neighbors. However, the development of a brand usually goes from less commonly used and less known to more and more popular, which is just contrary to the needs of consumers with strong uniqueness. Therefore, it is hypothesized that:

H3b: Need for uniqueness weakens the negative relationship between anticipated regret and intention to switch brands.

RESEARCH METHODS

This research adopted an online survey for testing the proposed model. Among a variety of surveys, online surveys enjoy low cost, high anonymity, high convenience, rapid diffusion and high response speed (Wright 2005). Smartphone is used as the object of this research. Anticipated regret is examined herein because a large proportion of consumers possess this particular affect when they purchase smartphone. Consumers on WeChat platform in China are legitimate respondents for they use smartphones daily. They were asked to fill out the questionnaire, which consists of all of the measurement items of the six constructs, along with behaviors and experience in using smartphones. After a series of snowball data collection, 470 usable questionnaires were secured for the subsequent data analysis.

Among the 470 usable samples, roughly 58.1% are unmarried people, followed by the married with children (34.5%), the married without children (5.3%), the divorced (1.3%), and the widowed (0.9%). In terms of age, only 3.2% of the samples are under 20 years old, 56.2% are between 21 and 30 years old, 9.8% are between 31 and 40 years old, 18.7% are between 41 and 50 years old, and 12.1% are above 50 years old. In terms of educational background, only 1.1% are junior high school or below, 7.9% are senior high school or technical secondary school, 12.1% are junior college, 37.9% have bachelor degree (the most), 36.4% have master's degree, and 4.7% have doctoral degree. As for occupation, white collar worker holds the largest proportion (47%) of the samples; students constitute 38.9%; the rest of respondents are selfoperated businessmen, retirees, freelancers that only account for a very small proportion. The mode of the respondents' monthly income is 8,000 yuan (about 1165 USD) or less, accounting for 81.4% of all samples, while 27% earn less than 2,000 yuan (about 291 USD). 32.1% of the samples are with monthly income between 2,000 yuan and 4999 yuan (about 727 USD), as opposed to 22.3% with monthly income between 5,000 yuan and 7999 yuan.

The most often used smartphone brands is Apple (48.5%), followed by Huawei (24%) and xiaomi (10.9%). In terms of operation systems, Android and iOS account for 50.4% and 48.1% respectively. As for replacing smart phone, 96% of the respondents have had replaced their phones, and the main reason for replacement is “the original mobile phone is broken”, accounting for about 44.3%. In addition, replacing the old version for “the new mobile phone specifications/performance” also accounts for about 21.3%, followed by about 15.7% for “the function of new mobile phone is better/more”. Finally, in terms of average inter-purchase interval, 18-24 months represent the largest proportion (33.2%), followed by 30-36 months (31.3%). Over 83.4% of the respondents choose to replace their smart phones within an average duration of 18-48 months.

We measure all the constructs in the research hypotheses by following Churchill’s (1979) approach. All responses were expressed on 5-point Likert scales. Six measures of interest were included in the final instrument. Anticipated regret is captured by a 16-item scale by adopting the study of Janis and Mann (1977). The intention to upgrade is adopted from the 6-item scale in Tseng and Lo’s (2000) research. The intention to switch brands is measured in terms of a 8item scale adopted from Jones and Tremblay (2000) and Söderlund and Julander (2003). As for moderators, novelty seeking and need for uniqueness are respectively measured by a 12-item scale and a 28-item scale. The former is based on the studies of Hirschman (1980) and Hsiao and Yang (2010), while the latter on those of Knight and Kim (2007) and Tian and McKenzie (2001). Satisfaction with the current brand is the control variable related to consumers’ purchase and use of smartphones.

After a series of confirmatory factor analyses (CFAs), convergent validity, reliability and discriminant validity are established. The CFAs use LISREL 8.80 to purify the reflective scales. Overall, all the CFA models show reasonably good fit, in terms of ^{2}/df (equaling to 3.377) and other model fit indicators as follows: Root Mean Square Error of Approximation (RMSEA) value is .074, Normed Fit Index (NFI) is .93, and both Comparative Fit Index (CFI) and Incremental Fit Index (IFI) are equaling to .95 (Bollen, 2014). The Cronbach’s Alpha values of each construct are greater than 0.8, indicating that questionnaires of the same construct are homogenous, and the reliability and unidimensionality of each construct are quite high (Nunnally, 1978). To be specific, the alpha values are 0.911 and 0.877 for the two sub-dimensions of novelty seeking, while 0.957, 0.851 and 0.954 for the three sub-dimensions of uniqueness. The intention to upgrade and the intention to switch brands have alpha values of 0.893 and 0.915 separately. Cronbach’s Alpha for the two lower level traits of anticipated regret are 0.950 and 0.940.

According to Sharma et al. (1981), the research uses the moderated regression analysis to test the hypotheses mentioned above. This model consists of the main effects of one control variable, one independent variable, two moderator variables, together with two relevant independent variable × moderator interactions. The satisfaction with current brand is the control variable in the regression analysis. The scores for all constructs are calculated by use of the corresponding questionnaire responses in the equally weighted scales. Before creating the interaction terms, the independent variables and moderator variables, anticipated regret, novelty seeking and uniqueness, have been mean-centered, so multicollinearity will not be a serious problem in this research.

Table 1 shows the results of the moderated regression model. In Model 1, the control variable accounts for 0.4% of the variance in current brand satisfaction (F=1.680, p> .10). Model 2 means the variance of current brand satisfaction is 19%, with the F-value of 27.285 (p< .01).

Between Model 1 and Model 2, there is a significant difference in R^2 with F equaling to 35.695 and p< .01, which means that the main effect makes sense. The effect of anticipated regret on intention to upgrade is positive and significant (= .643, p < .05), so Hypothesis 1a is supported.

Table 1 Results of Moderated Regression Analysis

Dependent Variable	Intention to Upgrade			Intention to Switch Brands		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Control Variable						
Satisfaction with Current Brand (SCB)	.060	.039	.038	-.222***	-.238***	-.237***
Independent Variable (Main Effect)						
Anticipated Regret (AR)		.105**	.643**		-.096*	-.779***
Moderator Variables (Main Effects)						
Novelty Seeking (NS)		.415***	.772***		.076	-.378**
Need of Uniqueness (NU)		-.039	-.037		-.242***	-.246***
Moderating Effects						
AR NS			-.731**			.929**
AR NU			.132**			-.175***
R2	.004	.190	.202	.049	.125	.145
R2		.187	.012		. 076	. 020
F-value	1.680	27.285***	19.506***	24.178***	16.573***	13.063***
F		35.695***	3.388**		13.398***	5.413***

*Significant at p < .10; **Significant at p < .05; ***Significant at p < .01

This research uses an incremental F-test to test the moderating effects (Aiken et al., 1991) in Hypotheses 2a and 3a. The approach is to compare Model 3 with Model 2. Between Model 2 and 3, the F-value is 3.388, with p < .05. Both of two interaction terms are statistically significant according to the test of individual interaction terms. In the H2a above, we predict that the positive relationship between anticipated regret and intention to upgrade increases as novelty seeking increases. As the regression coefficient of interactive term in H2a is significantly negative (= -.731, p < .05), the hypothesis is rejected. As for H3a, it predicts that need for uniqueness increases the positive relationship between anticipated regret and intention to upgrade. The corresponding interaction term, which is anticipated regret uniqueness, is positive and significant (= .132, p < .05), so H3a is supported.

As for Model 4, 5 and 6, the same approach is used to test. In Model 4, the control variable accounts for 4.9% of the variance in current brand satisfaction, with the F-value equals to 24.178 and p < .01. Model 5 means the variance of current brand satisfaction is 12.5% (F=16.573, p < .01). There is a significant difference in R^2 with F equaling to 13.398 (p < .01) between Model 4 and 5. Thus, the main effect is certainly present. Hypothesis 1b is supported according to the findings, because the effect of anticipated regret on intention to switch brands is significantly negative (=- .779, $p < .01$).

The F-value is 5.413 ($p < .01$) between Model 5 and Model 6. The test of individual interaction terms shows that both of these two interaction terms are statistically significant. Consistent with the expectation of H2b, the negative relationship between anticipated regret and intention to switch brands increases as novelty seeking increases. The data findings support it because the regression coefficient of interactive term is positive and significant (= .929, $p < .05$). Hypothesis 3b predicts that uniqueness decreases the negative relationship between anticipated regret and intention to switch brands. Finally, it is confirmed because the corresponding interaction term (anticipated regret need for uniqueness) is significantly negative, with = -.175, $p < .01$.

CONCLUSIONS

This research shows that the need for uniqueness has a significant moderating effect on the relationship between anticipated regret and the intention to upgrade and switch brands, while novelty seeking does not. Alternative explanations can be considered for such seemingly conflicting results. The development of smartphone in China has reached such a relatively mature stage that smartphones can mostly meet the daily needs of consumers in this market because of the fast speed of upgrading products. In China, manufactures often launch new products within only a few months and many updated functions and technologies for consumers become dispensable. As a result, even for novelty-seeking consumers, their anticipated regret psychology won’t affect their intention to replace any more. Furthermore, smartphones becomes a necessity partly due to the rapid growth of consumers’ personal income and their consumption level in China. The decisions to purchase and replace smartphones are not difficult for Chinese consumers. The economic value of smartphones is no longer critical in the purchase/repurchase decisions, especially for novelty-seeking consumers, who seldom worry about value declining. As a result, the moderation of novelty seeking on consumers’ affect-replacement linkage becomes relatively nonsignificant.

The findings of this study confirm the proposed model of consumers’ upgrade and brand switching decision, based on affective psychology. The research can make up for the gap in our understanding of anticipated regret, which refers to an allusion to the main psychological effect of various worries about loss, the materialized loss that troubles decision-makers before their decisions (Janis & Mann, 1977), and the impact of anticipated regret on consumers' intention to upgrade products and switch brands. Therefore, emotional process, instead of cognitive process, may drive the purchase behavior of consumers (Verhagen & van Dolen, 2011). It is clear, however, that current research on consumer decision-making is largely rooted in the cognitive domain, with little regard for affect. Therefore, the research content of this study shifted from general consumer cognition to consumer affect, and explored the mechanism of consumers' decision to upgrade and switch brands with the influence of anticipated regret in high-tech durable products. An investigation into the main effect of anticipated regret on consumers' intention to upgrade products and switch brands has been done, along with the moderating effects of novelty seeking and uniqueness. Studying the bahaviors of Chinese smartphone consumers may offer some theoretical and practical insights. The empirical results will help to better understand consumers' replacement decisions for high-tech durable products, even predict their behaviors in upgrading current products and switching to other brands. Several managerial guidelines can be provided for managing and launching high-tech consumer products. According to the data collected from the questionnaires in the paper, 96% of the samples have replaced their smartphones, so it can be judged that the current smartphone market is basically saturated, and the most important customers of smartphone are consumers who need to replace their products. However, in the current market with numerous smartphone brands, manufacturers should encourage to shorten the holding time, stabilize their own consumers and attract potential consumers from other brands, which is the key to their success. Therefore, the current research concludes that smartphone manufacturers should formulate appropriate and purposeful marketing strategies based on the characteristics, personality and behavior patterns of different consumers for the consideration of the uniqueness demand of consumers. Manufacturers can provide exclusive products, purchasing channels and characteristic advertisements for specific target groups to meet the personalized needs of consumers.

With limited human, material and financial resources, the research tries to complete the research step by step in a rigorous way. However, there are still rooms for improvement. There are more than 1.3 billion people in China. Such an extremely large population represents an obstacle for random sampling. In this study, questionnaires were distributed through the WeChat platform, from which the samples are inevitably students and young white collar workers. As a result, demographic variables such as age and occupation of the respondents could not be completely controlled .Furthermore, consumers are affected by many situational factors when making replacement decisions for smartphones. A logical step to advance our knowledge is to extend our study by incorporate both cognitive and affective aspects into the explanation and prediction of consumers’ replacement decision.

REFERENCES

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, Calif.: Sage Publications.

Berlyne, D. E. (1966). Curiosity and exploration. *Science*, 153(3731), 25-33.

Bollen, K. A. (2014). *Structural equations with latent variables*. New York: John Wiley & Sons.

Belk, R. W. (1988). Possessions and the extended self. *Journal of consumer research*, 15(2), 139-168.

Bellezza, S., Ackerman, J. M., & Gino, F. (2017). “Be careless with that!” Availability of product upgrades increases cavalier behavior toward possessions. *Journal of Marketing Research*, 54(5), 768-784.

Bello, D. C., & Etzel, M. J. (1985). The role of novelty in the pleasure travel experience. *Journal of Travel Research*, 24(1), 20-26.

Burns, D. J., & Warren, H. B. (1995). Need for uniqueness: Shopping mall preference and choice activity. International Journal of Retail and Distribution Management, 23(12), 4-12.

Churchill, G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research*, 16(1), 64-73.

Colombo, R. A., & Morrison, D. G. (1989). A brand switching model with implications for marketing strategies. *Marketing Science*, 8(1), 89-99.

Dickson, P. R. (1982). Person–Situation: Segmentation’s Missing Link. *Journal of Marketing*, 46(4), 56–64.

Frijda, N. H. (2010). Impulsive action and motivation. *Biological psychology*, 84(3), 570-579.
Holt, D. B. (1995). How consumers consume a typology of consumption practices. *Journal of Consumer Research*, 22(1), 1-16.
Hirschman, E. C. (1980). Innovativeness, novelty seeking, and consumer creativity. *Journal of Consumer Research*, 7(3), 283-295.
Hsiao, C. H., & Yang, C. (2010). Predicting the travel intention to take high speed rail among college students. *Transportation Research*, 13(4), 277-287.
Janis, I. L., & Mann, L. (1977). *Decision making – A psychological analysis of conflict, choice, and commitment*, New York: The free press.
Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review*, 7(3), 550-558.
Knight, D. K., & Kim, E. Y. (2007). Japanese consumers' need for uniqueness effects on brand perceptions and purchase intention. *Journal of Fashion Marketing and Management*, 11(2), 270280.
Lee, T. H., & Crompton, J. (1992). Measuring novelty seeking in tourism. *Annals of Tourism Research*, 19(4), 732-751.
Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, 128(3), 332-345.
Maslach, C., Stapp, J., & Santee, R. T. (1985). Individuation: Conceptual analysis and assessment. *Journal of Personality and Social Psychology*, 49(3), 729-738.
Nunnally, J. C. (1978). *Psychometric Theory, 2d Ed*. New York: McGraw-Hill.
Roese, N. J. (1997). Counterfactual thinking. *Psychological bulletin*, 121(1), 133.
Ruth, J. A., Brunel, F. F., & Otnes, C. C. (2002). Linking thoughts to feelings: investigating cognitive appraisals and consumption emotions in a mixed-emotions context. *Journal of the Academy of Marketing Science*, 30(1), 44-58.
Snyder, C. R. (1992). Product scarcity by need for uniqueness interaction: A consumer catch-22 carousel. *Basic and Applied Social Psychology*, 13(1), 9-24.
Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research*, 19(1), 105-118.
Sharma, S., Durand, R. M., & Gur-Arie, O. (1981). Identification and analysis of moderator variables. *Journal of Marketing Research*, 18(3), 291-300.
Snyder, C. R., & Fromkin, H. L. (1977). Abnormality as a positive characteristic: The development and validation of a scale measuring need for uniqueness. *Journal of Abnormal Psychology*, 86(5), 518-527.
Söderlund, M., & Julander, C. R. (2003). The variable nature of services: An empirical examination of trust and its effects on customers' satisfaction responses to poor and good service. *Total Quality Management & Business Excellence*, 14(3), 291-304.
Tse, P., & Crotts, J. C. (2005). Antecedents of novelty seeking: International visitors' propensity to experiment across Hong Kong's culinary tradition. *Tourism Management*, 26(6), 965-968.
Tseng, F. M., & Lo, H. Y. (2000). Antecedents of consumers' intentions to upgrade their mobile phones. *Telecommunications Policy*, 35(1), 74-86.
Tian, K. T., & McKenzie, K. (2001). The long-term predictive validity of the consumers' need for uniqueness scale. *Journal of Consumer Research*, 10(3), 171-193.
Verhagen, T., & van Dolen, W. (2011). The influence of online store beliefs on consumer online impulse buying: A model and empirical application. *Information & Management*, 48(8), 320-327.

Westbrook, R. A. (1987). Product/consumption-based affective responses and postpurchase processes. *Journal of marketing research*, 24(3), 258-270.
Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of computer-mediated communication*, 10(3), JCMC1034.
Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American psychologist*, 35(2), 151-175.
Zeelenberg, M., & Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. *Organizational Behavior and Human Decision Processes*, 72(1), 63-78.

DECISION SCIENCES INSTITUTE

Regulatory Uncertainty in Operating Environments: Slack as a Coping Mechanism

Matthew W Ford
Northern Kentucky University
fordmw@nku.edu

Bertie M. Greer
Wayne State University
Bertie.greer@wayne.edu

ABSTRACT

This paper considers how operations decision makers employ excess organizational resources known as slack to manage regulatory uncertainty. Slack will be oriented toward either preserving or altering workflows depending on perceptions about the direct and indirect influences of regulatory uncertainty. Findings highlight important contingencies associated with slack-related decisions made in regulatory contexts.

KEYWORDS: Resource Management and Allocation, Environmental Management, Innovation Process, Contracts, Outsourcing, Process Change, Supply chain management, Conceptual modeling

INTRODUCTION

To cope with uncertainty in external environments, organizations develop sources of strategic flexibility that enhance adaptive capability (Aaker & Mascarenhas, 1984; Collis, 1992; Wernerfelt & Karnani, 1987). The resilience of workflows that produce critical organizational resources is a common focus of these efforts. Improving the adaptability of core workflows often involves building reserves of flexibility known as slack. Slack is a collection of excess resources that is not used in routine production but can be marshalled when necessary to facilitate adaptation in uncertain settings (Azadegan et al., 2012; Bourgeois, 1981).

Slack enables operations to navigate uncertain environments through two seemingly contradictory mechanisms. Slack can *preserve* existing workflows by cushioning an organization from external shocks. Inventory, for example, is a form of slack that buffers operations from unforeseen volume-related changes that might otherwise motivate sweeping modifications to production processes (Thompson, 1967; Galbraith, 1973). Enabling the maintenance of status quo workflows in uncertain settings is an attractive feature of slack because of the high cost of altering work processes—particularly those that employ specialized assets. Slack can also *alter* workflows when external change forces cannot be buffered. For instance, slack in the form of research and development activities offers creative resources for innovation and experimentation that support reconfiguration of work when status quo operations are unable to satisfy changing market demands (Cyert & March, 1963).

The level and nature of uncertainty present in operating environments influences slack-related decisions. Generally, higher levels of environmental uncertainty motivate the creation of more slack (Ford, 2017). Because uncertainty stimulates adaptive response (Lawrence & Lorsch, 1967), managers employ slack resources in turbulent settings to increase maneuverability. Failure to fortify operations with enough slack to cope with external uncertainty can foster rigid, maladaptive work processes. Slack's precise shape depends on the nature of uncertainty in the operating environment (Ford, 2017). For example, market forces create uncertainty about supply chain volume requirements as well as about technological requirements necessary for production (Walker & Weber, 1984). When volume uncertainty dominates the environment, slack is likely to be oriented toward preserving workflows in the event of volume-related shocks. When technological uncertainty dominates, slack is oriented toward altering workflows that would otherwise be rendered obsolete in the face of technological change.

Because production and trade are commonly subject to government oversight, regulatory uncertainty is often present in operating environments as well. Processes that develop regulations in addition to the regulatory outcomes themselves create uncertainty that requires attention when making operating decisions (Duncan, 1972). Regulatory uncertainty can be amplified by lack of working knowledge about the bureaucratic machinations that shape regulatory settings (Art, 1973). Because penalties associated with regulatory non-compliance can be complex and apportioned by politically interested actors, the outcomes of regulatory processes and their consequences on operating environments are often difficult for managers to discern (Besley & Coate, 1998; deLeon, 1992; Pettigrew, 1973). It is also possible that regulations, once implemented, fail to achieve their intended or anticipated purposes (Carrigan, 2017) which could consequently introduce additional episodes of uncertainty that must be managed.

Extending slack's utility for managing uncertainty into regulatory environments seems intuitive. Yet, the literature is unclear about how operations develop and deploy slack to cope with regulatory uncertainty. A sizeable stream of research suggests that, in uncertain regulatory settings, organizations generally prefer to postpone investments and assume a "wait-and-see" posture until uncertainty is resolved (e.g., Bittlingmeyer, 2001; Fuss et al., 2009; Marcus, 1981; Yang et al., 2004). On the other hand, some researchers propose that investments made during periods of regulatory uncertainty, particularly those that focus on reversible commitments using flexible resources, can create valuable strategic options for producers as the uncertainty subsequently dissipates (Aragon-Correa & Sharma, 2003; Doh & Pearce, 2004; Hoffman et al., 2009; Rugman & Verbeke, 1998). If investments in slack are in fact made, then questions arise about how that slack should be shaped to address regulatory uncertainty. Because slack can generally be oriented toward preserving workflows or toward altering them (Bourgeois, 1981; Ford, 2017), the shape of slack may depend on the particular nature of regulatory uncertainty at hand. Research, however, has yet to focus on this possibility.

In the operations management literature, there is little discussion that managers should possess competence for coping with unstable institutional settings. Regulatory issues are often seen as beyond the control of operating managers. Although employing slack to manage external uncertainties is intuitive, researchers have largely ignored the idea in the context of operating in hostile political and regulatory environments.

This paper advances a contingency-based framework to clarify the relationship between regulatory uncertainty and slack employed by operations to manage that uncertainty. Building on work from Hoffman et al., (2008), we conceptualize regulatory uncertainty along two dimensions—one related to uncertainty about the direct characteristics and consequences of regulations themselves, and the other related to uncertainty about the indirect consequences of regulations once implemented. We then develop a series of propositions linking the two dimensions of regulatory uncertainty to the development of different slack orientations for purposes of either preserving workflows or altering them. Because it increases the cost of process modification and change (Williamson, 1975), asset specificity is considered as a moderator in the relationship between regulatory uncertainty and slack. Our primary contribution is the development of an improved conceptual foundation for how operations decision makers employ slack to cope with regulatory uncertainty.

DIRECT AND INDIRECT REGULATORY UNCERTAINTY

Regulation is the imposition of rules by government, often backed by penalties for noncompliance, meant to guide or control economic or social behavior (Jones & Thompson, 1984). Regulations can be narrow or broad in scope. In some cases, regulations target specific industries, production processes, or types of output, as in the case of particular air, water, and product quality and safety rules. In other cases, regulations touch many industries and impact producers across borders, as in the case of tariff, exchange rate, and central bank policies. Viewed from a supply chain perspective, regulations commonly influence the broad sequence of events from upstream raw material development and product assembly through downstream transportation, distribution, and sale of finished goods. Table 1 shows representative regulations often present in an operation's environment.

Managers consider regulation among the socio-political factors of the external environment to be taken into account when making decisions (Duncan, 1972). When information is scarce, time lags are long, and relationships between variables are difficult to grasp, uncertainty about the external environment increases, reflected by the inability to confidently assign probabilities to future events or to predict outcomes of forthcoming decisions (Lawrence & Lorsch, 1967; Duncan, 1972; Knight, 1921). In decision-making contexts, environmental uncertainty is generally considered to be perceptual in nature because decisions are based mainly on what individual decision makers sense about their surroundings (Downey & Slocum, 1975; Hambrick et al., 2005; Milliken, 1987; Perrow, 1970; Weber & Mayer, 2014).

Table 1: Representative Regulations in an Operation's Environment

Location	**Representative Regulations**
Upstream (inputs from suppliers)	Material requirements and restrictions Laws that restrict sourcing choices Input quality, inspection, and testing
Production processes	Work safety standards Minimum wage laws Hiring requirements (e.g., anti-discrimination laws) Environmental regulations and standards Required production practices Limits to what can be produced (e.g., market power restrictions)
Downstream (output to customers)	Product safety Product quality, inspection, and testing Product purity and cleanliness Pricing laws and restrictions Limitations on selling, who can be customers Information/disclosure requirements to buyers
General environment	Tariffs Trade restrictions and sanctions Laws governing mergers and acquisitions

Licensing/qualification requirements
Transportation rules and regulations
Exchange rate policy
Interest rate policy

Perceived uncertainty related to regulatory aspects of the environment can be tied to Milliken's (1987) notion of state uncertainty. State uncertainty involves the inability of managers to understand how factors in the external environment are changing. Managers may be unable to identify key events or changes in the environment, or they may be uncertain about the likelihood that future events or changes will occur. In the regulatory context, managers may not know, for instance, what regulations incoming political regimes might support, or be able to confidently predict whether proposed regulatory schemes will subsequently be approved. State uncertainty can also involve incomplete understanding of interrelationships between elements in the environment. For example, managers might be uncertain about the impact of newly imposed regulations on supply markets, or about how competitors will respond to those regulations.

Regulatory uncertainty, then, is a form of state uncertainty that manifests in perceptions about processes that develop regulations and the outcomes resulting from those processes. Hoffman et al. (2008) argued that regulatory uncertainty can be separated into two components: one related to uncertainty about the direct characteristics and consequences of regulations themselves, and the other related to uncertainty about the indirect consequences of regulations once implemented. Although we label these dimensions differently than Hoffman et al. (2008) for purposes of clarity, we define them similarly. *Direct regulatory uncertainty* is the inability of managers to predict the future state of the regulatory environment. It is linked to particular regulatory policies and their characteristics such as legislative direction, specific rules and measures, implementation plans, and interdependencies with other regulations (Lopez et al., 2017). *Indirect regulatory uncertainty* is the inability of managers to predict the state of the nonregulatory environment caused by a regulation. It is consistent with Milliken's (1987) observation that state uncertainty can sometimes involve incomplete understanding of interrelationships between elements in the environment. In the regulatory context, uncertainty about the state of the regulatory environment (i.e., direct regulatory uncertainty) often clouds knowledge about how regulatory changes will touch aspects of the non-regulatory environment (i.e., indirect regulatory uncertainty). The specter of new tariff impositions, for example, may create doubt not only about the precise shape and scope of the tariffs themselves, but also about how those tariffs might affect market-oriented dimensions of the environment such as those related to monetary exchange rates and trade patterns.

Although indirect regulatory uncertainty flows from regulatory uncertainty brought about by a changing regulatory state, it is possible that these two types of uncertainty may not coexist (Hoffman et al., 2008). Direct regulatory uncertainty is subject to abrupt, discontinuous resolution as segments of the regulation development process are completed. Sticking with our above tariff example, doubts about the shape and scope of upcoming tariffs are likely to fade once those tariffs are approved and implemented. The "punctuated equilibrium" nature of direct regulatory uncertainty differentiates it from market-related uncertainties that tend to diminish gradually over time. While direct regulatory uncertainty is often resolved abruptly, indirect regulatory uncertainty may persist for considerable time (Hoffman et al., 2008). While direct regulatory uncertainty associated with our tariff example is likely to decline once the details of the tariff have been determined and enacted, the indirect regulatory uncertainty associated with the tariff's effects on trade and other aspects of the non-regulatory environment may linger for great length of time after the tariff is implemented.

SLACK AND REGULATORY UNCERTAINTY

The temporal, multi-dimensional nature of regulatory uncertainty presents unique managerial challenges. Direct regulatory uncertainty threatens status quo production and process management with the prospect of rules and requirements that can force reconfiguration of workflows, alteration of work procedures, and installation of abatement equipment. Managers are unlikely to preemptively alter their production processes while direct regulatory uncertainty is present. They will be reluctant to do so not only due to the cost of making process changes, which could involve substantial irreversible commitments, but also because an abrupt end to direct regulatory uncertainty can be anticipated. Once the regulatory development process concludes, the direct consequences of the regulatory outcomes on production processes will become apparent. The risk of making a poor remedial decision is reduced at that point.

This does not mean that managers will necessarily assume a classic “wait-and-see” stance where no investment is made until the direct regulatory fog lifts. Because they may need to react quickly to regulatory outcomes once they become known, managers will favor a state of operational flexibility that permits maintenance of status quo operations while loosening core rigidities (Leonard-Barton, 1992) that might otherwise impair adaptation should it be mandated by the new regulatory environment. Reinforcing managerial preference for flexibility is the realization that, while direct regulatory uncertainty is likely to dissipate with the conclusion of the regulatory development process, doubts about the effects of newly imposed regulations on aspects of the non-regulatory environment will be prone to persist and perhaps even increase. The enduring presence of indirect regulatory uncertainty motivates managers to maintain a state of operational flexibility in order to facilitate adaptation until clarity can be achieved.

Creating a state of operational flexibility to cope with the multi-dimensional nature of regulatory uncertainty should precipitate the use of slack. Slack represents what Ghemawat and del Sol (1998) termed usage-flexible resources. Usage-flexible resources can be employed in various ways and applications. For example, usage-flexible resources can be used to change production volumes in response to demand fluctuations, alter the attributes of existing products, or to introduce new products (Ghemawat & del Sol, 1998). Slack’s usage flexibility provides strategic options for coping with regulatory uncertainty (e.g., Aragon-Correa & Sharma, 2003; Doh & Pearce, 2004; Hoffman et al., 2009; Rugman & Verbeke, 1998). Because managers can shape it at their discretion, slack can be deployed to correspond with particular junctures in the regulation development process. If that particular form of flexibility is no longer deemed necessary or effective, then those slack resources can be redeployed toward other uses. For example, once the regulatory development process concludes, then slack previously deployed to manage direct regulatory uncertainty might be reshaped to address lingering indirect regulatory uncertainty.

Higher levels of regulatory uncertainty should invite greater use of slack. Regulatory environments perceived as benign and predictable are unlikely to motivate investment in usage flexible resources since such flexibility is unnecessary in stable settings. On the contrary, slack deployed in stable environments can be a source of bureaucracy and waste (Bourgeois, 1981; Cheng & Kesner, 1997). As the regulatory environment becomes more unpredictable, however, managers will sense more value in accumulating usage-flexible resources to increase adaptive capacity. As regulatory uncertainty grows, so should the intensity of slack utilization.

Proposition 1: The higher the regulatory uncertainty, the greater the utilization of slack to cope with that uncertainty.

SLACK ORIENTATION AND DIMENSIONS OF REGULATORY UNCERTAINTY

As a usage-flexible resource, slack can, by definition, assume various forms. This variety can be categorized according to the two primary orientations that slack assumes when employed as a strategic environmental coping mechanism. Slack can be oriented toward either preserving workflows or toward altering them. In the section that follows, we describe these two categories of slack in some detail and propose how managers employ them to address direct and indirect regulatory uncertainty.

Preserving Orientations

When its purpose is to maintain status quo workflows, slack assumes a preserving orientation. Preservation-oriented slack insulates or protects workflows from external shocks. A traditional form of preservation slack is the operating buffer. Operating buffers, such as inventory and spare capacity, act as shock absorbers to cushion production processes from volume-related uncertainties that would otherwise disrupt workflows (Bourgeois, 1981; Galbraith, 1973). For example, managers might build inventories in advance of prospective regulations that threaten to shut down production facilities so that pollution abatement equipment can be installed.

In some cases, however, operating buffers may be unable to preserve status quo operations in the face of regulatory uncertainty. Excessive buffering can create suboptimal structural arrangements that align more with personal managerial preferences rather than with economic efficiency (Child, 1972; Yasai-Ardekani, 1986). Additionally, the excess buffering can slow managerial response to external conditions that demand quick action (Cheng & Kesner, 1997). Moreover, because uncertainty generally threatens the flow of resources between interdependent organizations in the supply chain (Oliver, 1991; Pfeffer & Salancik, 1978), operating buffers may not sufficiently protect the interorganizational linkages that govern resource exchange. To protect those linkages, managers will seek to create negotiated environments that govern exchanges in manners that secure those resources (Cyert & March, 1963; Dyer, 1997; Eisenhardt, 1989; Poppo & Zenger, 2002; Williamson, 1991). In the regulatory context, a necessary resource for unencumbered exchange includes government permission to engage in trade in a manner that other vital resources can be obtained. Managers can exploit the usage-flexible nature of slack to create preservation-oriented forms that foster negotiation and governance of interorganizational exchange so that resource flows can be stabilized in uncertain regulatory times.

Preservation-oriented slack that governs interorganizational exchange can be either formal or informal in nature (Ford, 2017). Formal exchange governance features the use of contracts that buttress exchanges to withstand adverse effects of unforeseen change (Williamson, 1991). In market contexts, contractual agreements are traditionally established between buyers and sellers in supply chains of goods and services. In regulatory contexts, contractual agreements can also be established between producers and government. For example, producers might provide personnel to sit on panels that co-develop new regulations with government bureaucrats. To develop formal governance mechanisms with trading partners, operations require slack resources that can be deployed for contracting activities and their subsequent enforcement. These preservation-oriented resources can be used in activities that enable the establishment of formal agreements and managing across organizational boundaries to stabilize resource exchange and maintain status quo workflows.

Formal governance of interorganizational transactions may not be effective in all environmental contexts (Ford, 2015; Poppo & Zenger, 2002). For instance, prospective regulatory changes may threaten to upset measures and standards that form the basis for control and improvement, making it difficult to enforce formal agreements over significant periods of time and driving up re-contracting costs (Handley & Benton, 2012; Heide & John, 1990). Uncertainties that arise in some environments can also widen information asymmetries between exchange partners that increase risk of strategic behavior and opportunism

(Eisenhardt, 1989; Nilikant & Rao, 1994), leading to less willingness to enter into formal commitments.

As an alternative to formal governance, slack resources can also be applied toward informal, relational governance. Relational governance mechanisms foster trust, parallel expectations, joint action, and procedural fairness that build cooperative, flexible alliances between exchange partners (Benton & Maloni, 2005; Dyer & Singh, 1998; Geyskens et al., 2006; Heide, 1994; Kanter, 1994). Slack deployed in relational governance manners commonly assumes the form of joint planning and steering committees, inter-organizational problem solving teams, and customer/supplier training programs (Ford, 2017). In regulatory contexts, slack could similarly sponsor special studies, white papers, lobbying and other efforts aimed at influencing institutional participants in the regulation development processes (Bonardi & Keim, 2005; Hillman & Hitt, 1999; Kingsley et al., 2012). Slack that builds relational governance capability, such as proficiencies for boundary spanning and interorganizational teamwork, helps stabilize resources flows and strengthen competitive positions in technologically uncertain contexts. Because they do not depend on particular standards or specifications, relational governance mechanisms are capable of enduring high levels of environmental uncertainty while encouraging mutually beneficial collaboration on matters affecting resource stability. Table 2 includes the various forms of preservation-oriented slack and some examples of each.

Table 2: Slack Orientations and Examples in the Regulatory Context

Preservation Orientation	Altering Orientation
Operating Buffers:	*Scanning:*
Inventory	Monitoring of regulatory media
Spare capacity	Networking with regulatory officials
	Performing competitive analysis
Formal Exchange Governance:	
Contract development and oversight	*Workflow Reconfiguration:*
Regulatory committee participation	Research and development
	Human resource training
Relational Governance:	Project management skills
Coalitions, trade groups	Team-based problem problem-solving
Interorganizational planning/problem solving	
Lobbying	
Conducting special studies	
Regulatory committee participation	

Higher levels of direct regulatory uncertainty should invite use of more preservation-oriented slack. As direct uncertainty increases, doubts grow about the impact of prospective regulations on workflows. Use of slack to protect workflows should intensify as managers create a state of flexibility until they can gain clarity about what durable remedies for the new regulatory environment might look like.

Proposition 2: The higher the direct regulatory uncertainty, the greater the utilization of slack oriented toward preserving existing workflows.

Altering Orientations

Although preserving workflows is generally preferred due to the high cost of process change, several factors may motivate managers to consider the possibility of modifying core workflows under conditions of regulatory uncertainty. As uncertainty increases, preservation-oriented slack may not provide enough capacity for buffering external forces from upsetting resource flows. On the other hand, because excessive buffering capacity could render decision-makers less responsive to environmental stimuli (Cheng & Kesner, 1997; Modi & Mishra, 2011), managers might consider large measures of preservation-oriented slack as imprudent. Of course, it is also possible that managers simply perceive the nature of regulatory uncertainty to be such that status quo workflows will not enable adaptation and consequently must be modified (Rindova & Kotha, 2001).

When its purpose is to modify or change workflows, slack assumes an altering orientation. Because environmental uncertainty creates discrepancies and challenges to the skills upon which operating tasks are based (Fredrickson & Mitchell, 1984), dynamic capabilities are necessary that enable the alteration of production processes (Teece et al., 1997; Winter, 2003). Dynamic capabilities facilitate workflow modification by increasing organizational capacity for innovation and adaptation (Helfat & Winter, 2011; Rindova & Kotha, 2001; Teece, 2014). Since slack can provide resources for creative behavior (Bourgeois, 1981), dynamic capabilities for innovation are furnished by slack with an altering orientation.

Dynamic capabilities involve skills for scanning environments and reconfiguring workflows (Cohen & Leventhal, 1990; Collis, 1994; Pavlou & El Sawy, 2011; Teece, 2007). In the regulatory context, alteration-oriented slack must develop aptitude for scanning regulatory environments for opportunities, threats, and changes. Aptitude must also be acquired for experimentation and for modifying work processes to better meet external demands so that vital resources can be secured. Table 2 includes the various forms of alteration-oriented slack and some examples of each.

Due to their preference for maintaining status quo workflows until outcomes of regulation development processes are known, managers are unlikely to employ large amounts of alteration-oriented slack in response to direct regulatory uncertainty. Instead, managers should view altering orientations as more attractive as indirect regulatory uncertainty grows and lingers. Higher levels of indirect regulatory uncertainty increase doubt about how prospective regulations will impact the non-regulatory environment and everyday production processes. Because this uncertain state may persist for some time, managers are likely to orient more slack toward the possibility that workflows may require modification in order to adapt.

Proposition 3: The higher the indirect regulatory uncertainty and the longer it persists, the greater the utilization of slack oriented toward altering existing workflows.

ASSET SPECIFICITY

In economic contexts, specialization is the extent to which individuals, groups, or organizations perform a narrow range of activities (Schilling et al., 2003). The benefits of specialization on productivity have been recognized for some time (e.g., Fayol 1916; Smith, 1776; Taylor 1911), and include learning effects from repetition and efficiency gains from low switching costs. Most operations adopt at least some degree of specialization to realize a portion of these benefits.

Specialized production fosters specificity in labor as well as in assets such as machinery, facility location, and brand name capital (Williamson, 1991). Asset specificity refers to the degree to which an asset can be redeployed to alternative uses without loss of productive value (Williamson, 1975). The redeployment costs associated with highly specific assets make them riskier than general purpose assets. Because they require commitment to specific assets that can be difficult to modify or trade, specialized operations are inherently rigid and unresponsive in uncertain environments (Ghemawat, 1991; Leonard-Barton, 1992; Williamson, 1975). Asset specificity forces operations decision makers to protect existing production by increasing slack due to the limited ability of an alternative use for specialized assets.

Degree of asset specificity should moderate the relationship between regulatory uncertainty and use of slack. When responding to a given level of regulatory uncertainty, highly specialized operations are likely to employ more slack than operations with lower degrees of specific assets. The rigid nature of specialized workflows requires additional resources to create a state of flexibility to cope with both direct and indirect regulatory uncertainty. Greater use of slack increases the likelihood that specialized operations will adapt in uncertainty regulatory settings.

Proposition 4: At a given level of regulatory uncertainty, higher degrees of asset specificity will be associated with greater utilization of slack.

DISCUSSION

This study examines the relationship between regulatory uncertainty and slack employed by operations decision makers. Generally, higher levels of regulatory uncertainty should motivate greater utilization of slack. When direct regulatory uncertainty is elevated, managers are likely to prefer slack oriented toward preserving workflows. When indirect regulatory uncertainty increases, managers are more likely to employ slack oriented toward altering workflows. Specialized operations, which contain high degrees of specific assets that are difficult to modify and trade, should utilize slack more in order to cope with regulatory uncertainty than should operations composed of general-purpose assets. Our primary contribution is the development of an improved, contingency-based foundation for how operations employ slack to cope with regulatory uncertainty.

In the operations management literature, the notion that managers must possess competence for coping with unstable institutional settings remains largely novel. Regulatory settings are often seen as beyond the control of operating managers. Although employing slack to manage external uncertainties is intuitive, researchers have largely ignored the idea in the context of operating in hostile political and regulatory environments. Hopefully, this paper will encourage dialogue and work in this direction.

The conceptual framework proposed in this study must be tested and validated. Empirical research might extend Lopez et al.'s (2017) work that considered how direct and indirect regulatory uncertainty affected corporate investment decisions. Because supply chains continue to lengthen across country borders, future studies should also investigate situations where managers must cope with a portfolio of regulatory uncertainties that accrue from operating in multiple political regimes simultaneously.

From a practical standpoint, previous work has alerted managers to contingencies linked to employing slack for managing environmental uncertainty. Ford (2017), for example, suggested that how slack is oriented to address uncertainty is likely to depend on whether volume or technological uncertainty dominates market environments. This study extends the contingency notion to the external environment. Managers will be prone to orient slack toward either preserving or altering workflows depending on direct and indirect regulatory uncertainties that threaten the production status quo.

REFERENCES

Aaker, D.A. & Mascarenhas, B. (1984). The need for strategic flexibility. *Journal of Business Strategy, 5*(2), 74-82.

Aragon-Correa, J.A. & Sharma, S. (2003). A contingent resource-based view of proactive corporate environmental strategy. *Academy of Management Review, 28*, 71-88.

Art, R.J. (1973). Bureaucratic policies and American foreign policy: A critique. *Policy Sciences, 4*(4), 467-490.

Azadegan, A., Patel, P.C. & Parida, V. (2012). Operational slack and venture survival. *Production and Operations Management, 22*, 1-18.

Benton, W.C. & Maloni, M. (2005). The influence of power driven buyer/seller relationships on supply chain satisfaction. *Journal of Operations Management, 23*, 1-22.

Besley, T. & Coate, S. (1998). Sources of inefficiency in a representative democracy: A dynamic analysis. *American Economic Review, 88*(1), 139-156.

Bittlingmeyer, G. (2001). Regulatory uncertainty and investment: Evidence from antitrust enforcement. *Cato Journal, 20*, 295-325.

Bonardi, J.P & Keim, G. (2005). Corporate political strategies for widely salient issues. *Academy of Management Review, 30*, 555-576.

Bourgeois, L.J. III (1981). On the measurement of organizational slack. *Academy of Management Review, 6*, 29-39.

Carrigan, C. (2017). *Structured to fail? Regulatory performance under competing mandates*. Cambridge: Cambridge University Press.

Cheng, J.L.C & Kesner, I.F (1997). Organizational slack and response to environmental shifts: The impact of resource allocation patterns. *Journal of Management, 23*, 1-18.

Child, J. (1972). Organizational structure, environment, and performance: The role of strategic choice. *Sociology,* 6, 2-22.

Cohen, W.M. & Levinthal, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35,* 128-152.

Collis, D. (1992). The strategic management of uncertainty. *European Management Journal, 10*, 125-135.

Collis, D.J. (1994). How valuable are organizational capabilities? *Strategic Management Journal, 15*, 143-152.

Cyert, R.M. & March, J.G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.

deLeon, P. (1992). Policy formulation: Where ignorant armies clash by night. *Policy Studies Review, 11*(3/4), 389-405.

Doh, J.P. & Pearce, J.A. (2004). Corporate entrepreneurship and real options in transitional policy environments: Theory development. *Journal of Management Studies, 41*, 645-664.

Downey, H.K. & Slocum, J.W. (1975). Uncertainty: Measures, research, and sources of variations. *Academy of Management Journal, 18*, 562-577.

Duncan, R.B. (1972). Characteristics of organizational environments and perceived environmental uncertainty. *Administrative Science Quarterly, 17*, 313-327.

Dyer, J.H. & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review, 23*, 660-679.

Dyer, J.H. (1997). Effective interfirm collaboration: How firms minimize transaction costs and maximize transaction value. *Strategic Management Journal, 18*, 535-556.

Eisenhardt, K.M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*, 57-74.

Fayol, H. (1916). *Administration industrielle generale*. Paris: Dunod.

Ford, M.W. (2015). Supply chain quality management and environmental uncertainty: A contingency perspective. *Quality Management Journal, 22*(4), 54-65.

Ford, M.W. (2017). Specialization, slack orientation, and adaptive capacity in uncertain environments. *Journal of Strategic Innovation and Sustainability, 12*(1), 55-67.

Fredrickson, J.W. & Mitchell, T.R. (1984). Strategic decision processes: Comprehensive and performance within and industry with an unstable environment. *Academy of Management Journal, 27*, 399-423.

Fuss, S., Johansson, D.J.A., Szolgayova, J., & Obersteiner, M. (2009). Impact of climate policy uncertainty on the adoption of electricity generating technologies. *Energy Policy, 37*, 733-743.

Galbraith, J.R. (1973). *Designing complex organizations*. Reading, MA: Addison-Wesley.

Geyskens, I., Steenkamp, J.B.E. & Kumar, N. (2006). Make, buy, or ally: A transaction cost theory meta-analysis. *Academy of Management Journal, 49*, 519-543.

Ghemawat, P. (1991). *Commitment: The dynamic of strategy*. New York: The Free Press.

Ghemawat, P. & del Sol, P. (1998). Commitment versus flexibility? *California Management Review, 40*(4), 26-42.

Hambrick, D.C., Finkelstein, S. & Mooney, A.C. (2005). Executive job demands: New insights for explaining strategic decisions and leader behaviors. *Academy of Management Review, 30*, 472-491.

Handley, S.M. & Benton, W.C. (2012). The influence of exchange hazards and power on opportunism in outsourcing relationships. *Journal of Operations Management, 30*, 55-68.

Heide, J.B. & John, G. (1990). Alliances in industrial purchasing: The determinants of joint action in buyer-seller relationships. *Journal of Marketing Research, 27*, 24-36.

Heide, J.B. (1994). Interorganizational governance in marketing channels. *Journal of Marketing, 58*, 71-85.

Helfat, C.E. & Winter, S.G. (2011). Untangling dynamic and operational capabilities: Strategy for the (n)ever changing world. *Strategic Management Journal, 32*, 1243-1250.

Hendricks, K., Singhal, V. & Zhang, R. (2009). The effect of operational slack, diversification, and vertical relatedness on the stock market reaction to supply chain disruptions. *Journal of Operations Management, 27*, 233-246.

Hillman, A. & Hitt, M. (1999). Corporate political strategy formultation: A model of approach, participation and strategy decisions. *Academy of Management Review, 24*, 825-842.

Hoffman, V.H., Trautmann, T., & Hamprecht, J. (2009). Regulatory uncertainty: A reason to postpone investments? Not necessarily. *Journal of Management Studies, 46*, 1227-1253.

Hoffman, V.H., Trautmann, T., & Schneider, M. (2008). A taxonomy for regulatory uncertainty: Application to the European emission trading scheme. *Environmental Science & Policy, 11*, 712722.

Jones, L.R. & Thompson, F. (1984). Risks for efficiency: Comprehensive reform of direct regulation. *Academy of Management Review, 9*, 746-756.

Kanter, R.M. (1994). Collaborative advantage. *Harvard Business Review, 74*(4), 96-108.

Kingsley, A.F., Vanden Bergh, R.G. & Bonardi, J. (2012). Political markets and regulatory uncertainty: Insights and implications for integrated strategy. *Academy of Management Perspectives, 26*(3), 52-67.

Knight, F. (1921). *Risk, uncertainty, and profit.* Boston: Houghton-Mifflin.

Lawrence, P.R. & Lorsch, J.W. (1967). *Organization and environment*. Cambridge, MA: Harvard University Press.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal, 13*, 111-125.

Lopez, J.M.R., Sakhel, A., & Busch, T. (2017). Corporate investments and environmental regulation: The role of regulatory uncertainty, regulation-induced uncertainty, and investment history. *European Management Journal, 35*, 91-101.

Marcus, A.A. (1981). Policy uncertainty and technological innovation. *Academy of Management Review, 6*, 443-448.

Milliken, F.J. 1987. Three types of perceived uncertainty about the environment: State, effect, and response uncertainty. *Academy of Management Review, 12*, 133-143.

Modi, S.B. & Mishra, S. (2011). What drives financial performance – resource efficiency or resource slack? Evidence from U.S. based manufacturing firms from 1991 to 2006. *Journal of Operations Management, 29*, 254-273.

Nilikant, V. & Rao, H. (1994). Agency theory and uncertainty in organization: An evaluation. *Organization Studies, 15*, 649-672.

Oliver, C. (1991). Strategic responses to institutional processes. *Academy of Management Review, 16*, 145-179.

Pavlou, P.A. & El Sawy, O.E. (2011). Understanding the elusive black box of dynamic capabilities. *Decision Sciences, 42*, 239-273.

Perrow, C. 1970. *Organizational analysis: A sociological view*. Belmont, CA: Wadsworth.

Pettigrew, A.W. (1973). *The politics of organizational decision-making*. London: Tavistock.

Pfeffer, J. & Salancik, G.R. (1978). *The external control of organizations: A resource dependence perspective.* New York: Harper & Row.

Poppo, L. & Zenger, T. (2002). Do formal contracts and relational governance serve as substitutes or complements? *Strategic Management Journal, 23*, 707-725.

Rindova, P.V. & Kotha, S. (2001). Continuous “morphing”: Competing through dynamic capabilities, form, and function. *Academy of Management Journal, 44*, 1263-1280.

Rugman, A.M. & Verbeke, A. (1998). Corporate strategies and environmental regulations: An organizing framework. *Strategic Management Journal, 19*, 363-375.

Schilling, M.A., Vidal, P., Ployhart, R.E. & Marangoni, A. (2003). Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science, 49*, 39-56.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Staats, B.R. & Gino, F. (2012). Specialization and variety in repetitive tasks: Evidence from a Japanese Bank. *Management Science, 58*, 1141-1159.

Taylor, F. (1911). *The principles of scientific management*. New York: Harper & Brothers.

Teece, D.J. (2007). Explicating dynamic capabilities: The nature and micro foundations of (sustainable) enterprise performance. *Strategic Management Journal, 28*, 1319-1350.

Teece, D.J. (2014). The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives, 28*(4), 328-352.

Teece, D.J., Pisano, G. & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*, 509-533.

Thompson, J.D. (1967). *Organizations in action*. New York: McGraw Hill.

Walker, G. & Weber, D. (1984). A transaction cost approach to make or buy decisions. *Administrative Science Quarterly, 29*, 373-391.

Weber, L. & Mayer, K. (2014). Transaction cost economics and the cognitive perspective: Investigating the sources and governance of interpretive uncertainty. *Academy of Management Review, 39*, 344-363.

Wernerfelt, B. & Karnani, A. (1987). Competitive strategy under uncertainty. *Strategic Management Journal, 8*, 187-194.

Williamson, O.E. (1975). *Markets and hierarchies: Analysis and antitrust implications.* New York: Free Press.

Williamson, O.E. (1991). Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly, 36*, 269-296.

Winter, S.G. (2003). Understanding dynamic capabilities. *Strategic Management Journal, 24*, 991-995.

Yang, B., Burns, N.D., & Backhouse, C.J. (2004). Management of uncertainty through postponement. *International Journal of Production Research, 42*, 1049-1064.

Yasai-Ardekani, M. (1986). Structural adaptations to environments. *Academy of Management Review, 11*, 9-21.

DECISION SCIENCES INSTITUTE

The Institutions Influence in an Uncertain Environment: A Coevolutionary Perspective

Kenyth Alves de Freitas
Fundação Getulio Vargas
Email: kenyth.freitas@gmail.com

Ely Laureano Paiva
Fundação Getulio Vargas
Email: ely.paiva@fgv.br

ABSTRACT

Multinational Enterprises (MNEs) with operations in emerging countries seek to adjust to institutional uncertainty. This paper seeks to understand how MNEs respond to market changes caused by the uncertain environment through a coevolutionary perspective, guided by the following research question: how does the uncertain institutional environment affect capabilities' coevolution process? This study follows a multiple case study research strategy in seven MNEs. The results suggest that a high level of uncertainties affect companies directly or indirectly in this institutional environment. In critical levels of uncertainty severity, result in new capabilities development (exploitation strategy), or strengthen existing ones (exploration).

KEYWORDS: Coevolutionary theory, Uncertain environment, Ambidexterity, MNEs, and Multiple case study

INTRODUCTION

In the management field, the coevolutionary theory seeks to understand the selection and adaptation processes of firms in their environments (Mckelvey, 1997). Regarding a firm's environment, this theory assumes it as a complex set of relationships amongst supply chain partners, industry competitors, entities, and institutions that influence firms' corporate strategy. Therefore, the aim of firms is not only to adapt themselves to the environment and actors they interact with but also to influence it (Cantwell, Dunning, & Lundan, 2010).

To be capable of influence or to be influenced by their environment, firms need to make decisions about their internal resources. They can address this issue by creating new practices and capabilities which can impact on their business. One way is by improving internal resources and capabilities, and/or by being able to put them together to get influence over its competitors, partners, and institutions.

We assume that uncertain institutional environments are more susceptible to market volatility, changes in legislation and regulation, theft, fraud, and corruption. Regarding the former, information asymmetry arose by the uncertain environment, mainly during transitions, making way for high levels of opportunism when dealing with third-parties (Peng, Lee, & Wang, 2005). Firms may suffer from opportunistic, unfair or unlawful behaviors (Su, Peng, & Xie, 2016), as well as vulnerability to sudden institutional disruption, market reforms and political backlash (Yamakawa, Peng, & Deeds, 2008). Companies also tend to rely on relationships and networks more intensely if institutions are weak, since formal rules are fragile, forcing them to develop strategies to face uncertainty and institutional idiosyncrasies (Meyer & Peng, 2016).

Mair, Marti, and Ventresca (2012) argue that the institutional environment affects business' infrastructures to the extent that it can be considered as market failures, influencing all stages of the supply chain. For instance, emerging markets have institutional idiosyncrasies that are not usually found in advanced economies, leading them to an uncertain environment, thus, forcing companies to adapt their strategy to the institutions they are subject to, and also to influence them. In that sense, we question: “How does the uncertain institutional environment affect operational capabilities’ coevolution process?”.

To address the question proposed in this study, we follow a qualitative approach by multiple case studies (Eisenhardt, 1989; Yin, 2014), and seek to understand how companies develop capabilities to respond and adapt to externalities caused by the uncertain institutional environment in a coevolutionary perspective. The use of a case study is justified because this phenomenon is not very well understood, also it is indicated to analyze complex subjects (Meredith, 1998). In this research, seven companies were selected; leaders in the national and global market, of the automotive (3 companies), beverage (2 companies), cosmetic (1 company), and tobacco (1 company) sectors.

This research seeks to understand the strategies adopted by MNEs in uncertain institutional environments to reduce the impacts on their operations. For that matter, this study uses coevolutionary theory to understand the process of interaction between companies and the external environment, and the concept of ambidexterity to identify the influences that companies generate in the external environment.

The sequence of this paper follows the subsequent order: the second section presents a literature review of Coevolutionary Theory and Institutional Theory, and how institutional uncertainty context generated changes in capabilities coevolution. The third section describes the methodological steps used in the case study based on Eisenhardt (1989). The fourth section presents the results of the within-case and cross-case analyses, followed by a section to discuss our findings with the literature. Finally, the main conclusions and orientations for future studies are presented in the last section.

LITERATURE REVIEW

Coevolutionary Theory

The coevolutionary theory was born from natural science arguing a joint evolution between organisms, species of animals and plants (Ehrlich & Raven, 1964). The coevolutionary approach was brought to the strategy field to understand the selection and adaptation of companies to a competitive environment (McKelvey, 1997). According to Lewin and Volberda (1999), the coevolutionary process would be shaped by a large and complex process of coevolution that would involve influences through different levels of management.

These influences could come from the corporate, the supply chain, the industry and the country (or region) levels (Lewin & Volberda, 1999). However, they differ regarding the (immersion) embeddedness level of a company in an environment and its relationship level with those players. Thus, this approach allows analyzing the company's choices at micro, meso or macro levels, which creates a possibility to broaden the understanding of the different strategies adopted in operations. Simultaneously, the company’s capabilities can be influenced by competitive and institutional environments, and the company itself, through an innovative resource or process, can affect the external environment from a new capability (Rodrigues & Child, 2003). From this perspective, the organizations’ strategy is not only to adapt themselves to the institutions, but also to influence them (Cantwell, Dunning, & Lundan, 2010).

Multinational enterprises (MNE’s) are ideal for coevolutionary theory application because they integrate a set of distinct environments and contexts. They put together the role of institutions for creating new policies and the effect of corporate strategies over subsidiaries (macro-environment), the role of suppliers, customers and competitors and their activities (meso-environment), and the set of resources and capabilities of subsidiaries (micro-environment) (Madhok & Liu, 2006).

According to Lampel and Shasie (2003), the coevolutionary theory presupposes that the mobilization of resources and the creation of competences are self-determined. The concept of ambidexterity considers that resources exploration can result in two distinct processes. The first of them, exploitation, the company mobilizes its internal resources to fulfill a short-term change (Kristal, Huang, & Roth, 2010). However, in the second process, exploration, the company uses its resources with a focus on the future, generating new capabilities (O'Reilly & Tushman, 2004). According to Birkinshaw and Gupta (2013), these two processes are not necessarily excluding, but acting together can offer present gains and anticipation of future changes.

For that matter, the coevolutionary theory suggests that are four strategies that companies can respond to external changes. We related each one of that strategy with ambidexterity concept.

- Naive selection - A random and unconscious choice of practices occurs, these being little connected with the institutional or competitive environments. The selection of a new competence is only to minimize the impacts of some change in the external environment at its current level of performance. It would be equivalent to Stage 1 proposed by Wheelwright and Hayes (1985) and focus on the exploitation process.
- Managed selection - Occurs in complex internal environments in which it is necessary to select practice or routine in front of the others in the market. An analysis of the times and resources spent in practice or routine is carried out. In this case, it would be equivalent to Stage 2 proposed by Wheelwright and Hayes (1985) and a major focus on the exploitation process.
- Hierarchical Renewal - The choice of practices and routines are strategic elements of the business. Thus, the process is proactive, well-defined and with clear objectives to achieve strategic goals, however, all guidance is top-down. It would be equivalent to Stage 3 proposed by Wheelwright and Hayes (1985) and a combined exploitation and exploration processes.
- Holistic Renewal - There is collective learning at the different levels of the company, in the top-down and bottom-up perspectives, which aims at cyclically renewing various aspects of the organization, including beliefs, cultures, and practices. In this case, it would be equivalent to Stage 4 proposed by Wheelwright and Hayes (1985) and focus on the exploration process.

Institutional Uncertain Environment

According to Meyer and Rowan (1977), the organizational structure is not only composed of available resources but is the result of the interaction of these resources with the current institutional environment. The institutional environment determines and limits the opportunities of new business for the companies (Bruton, Ahlstron, & Li, 2010). According to Scott (2008), the institutional environment can be defined as "rules, norms, and belief systems [which] undergird all stable social systems, including economic systems." Government actions and social norms, therefore, can create a restrictive environment, which becomes an obstacle to the development of a new business or even to the maintenance of current business performance.

According to DiMaggio and Powell (1983), this environment generates external pressures on these organizations, through norms, coercion or adequacy. However, organizations cannot be seen as passive agents in this process. For Oliver (1991), economic agents interact with the institutional environment. Their reactions, nonetheless, are not only restricted to adjusting to external pressures, but they can also avoid, challenge, compromise and/or manipulate these pressures.

The differences between institutions of emerging and developed economies are noticeable (Meyer et al., 2009). In the former, information asymmetry arose by the uncertain environment, mainly during transitions, making way for high levels of opportunism when dealing with third-parties (Peng, Lee, & Wang, 2005). Firms may suffer from opportunistic, unfair or unlawful behaviors (Khanna & Palepu, 1997; Su, Peng, & Xie, 2016), besides vulnerability to sudden institutional disruption, market reforms and political backlash (Yamakawa, Peng, & Deeds, 2008). Companies also tend to rely more intensively on relationships and networks if formal rules are fragile, developing strategies to face uncertainty and institutional idiosyncrasies (Meyer & Peng, 2016).

Emerging economies have less sophisticated institutional frameworks and inconsistent institutional pressures sometimes (Meyer & Peng, 2016). The lack of certain market-supporting institutions leads to the existence of “institutional voids” (Khanna & Palepu, 1997), reducing the tendency for isomorphism (Meyer & Peng, 2016). Those “voids” can be filled or lessened by Government support, for example (Su, Peng, & Xie, 2016). When present, firms should redesign their organizational forms (Meyer & Peng, 2016) and also ponder decision-making processes to face such “institutional voids”.

Not all emerging markets are the same. Meyer et al. (2009) classify the emerging economies institutional framework into two categories: developed (strong) and underdeveloped (weak). For them, those countries "show substantial variation in formal and informal institutions (Meyer et al., 2009: 62), which makes them an interesting subject of analysis.

In this complex scenario, emerging economies are significant markets, due to their supply of raw materials, labor at low prices or market size, although they might present many peculiarities (Rottig, 2016). International business scholars have been comparing MNEs’ operations in both home and host country in several aspects (Ramamurti, 2012; Estrin, Meyer, & Pelletier, 2018). Recently He, Brouthers, and Filatotchev (2018) point out that when an MNE from an emerging country acquiring facilities in a developed country it needs to develop and to upgrade its capabilities to the new market. It means that capabilities development occurs in both directions when moving to a different institutional context.

On the other hand, investing in emerging economies may be much more challenging for companies than they usually expect, particularly, in their supply chain activities (Karamchandani, Kubzansky, & Nishant, 2011). Companies must deal with the occurrence of institutional voids in its different forms (labor, regulatory and contracting mechanisms) in the raw material, manufacturing, distribution and marketing activities of the value chain (Parmigiani & Rivera-Santos, 2015).

The lack of efficient institutions demands firms to make strategic choices to operate in institutional voids environments (Peng, 2003). From an institutional perspective, these emerging markets have considerable institutional idiosyncrasies that usually are not found in advanced economies (Rottig, 2016). Doh, Rodrigues, Helmhout, and Makhija (2017) argue that institutional voids might enhance opportunism and corruption, among other institutional failures.

METHODS

As presented in the introduction, this research question guided our study: “How does the uncertain institutional environment affect operational capabilities’ coevolution process?”. This study follows a qualitative approach through multiple case studies (Yin, 2014), and seeks to understand how MNEs develop capabilities to respond to institutional uncertainties in emerging markets. The use of this research strategy is justified because this phenomenon is not very well understood (Meredith, Raturi, Amoako-Gyampah, & Kaplan, 1989), also that it is indicated to analyze complex subjects. According to Barratt, Choi, & Li (2011: 330), a case study based on the “how” question takes an explanatory approach because: [..] “the context and experiences of actors are critical and especially the experiences of managers so as to increase the practical relevance of the findings."

This section was structured according to the five stages of a study case, initially proposed by Eisenhardt (1989) and continually improved (Eisenhardt, 1991; Stuart, Mccutcheon, Handfield, Mclachlin, & Samson, 2002). According to this model, the analysis of the case study should follow these steps: research question (already presented); case selection, data gathering; data analysis; and replication. In this way, the next subsections were structured following the one proposed by this author.

Case selection

We decided to select MNEs that have greater power of adapting capabilities and also the power of influencing their respective markets. According to Eisenhardt (1989), a multiple case study should select between four and ten cases to be considered effective. In this research seven companies were selected, leaders in the national and global market, of the sectors automotive (3 companies), beverages (2 companies), cosmetics (1 company), and tobacco (1 company). After identifying the most appropriate cases to fulfill the purposes of our research, the managers of these companies were invited to participate by e-mail or telephone. Then, the objectives and scope of this research were presented, and, when requested, the semi-structured interview protocol also. Also, a confidentiality agreement was sent indicating that we would not disclose the companies or the participants' names.

Therefore, the following cases were selected:

- Case 1 - AUTO1 - An automotive multinational that operates in a host country. In addition to Brazil, this company operates in more than 40 countries with a global annual revenue above US$ 100 billion.
- Case 2 - AUTO2 - An automotive multinational firm that operates in a host country. It has assembly, manufacturing, distribution, office, or warehousing operations in more than 60 other countries with a global annual revenue above US$ 140 billion.
- Case 3 - AUTO3 - An automotive multinational firm that operates in a host country. The company manufactures popular and luxury cars, sports cars, trucks and commercial vehicles with a global annual revenue above US$ 200 billion.
- Case 4 - BEER1 - A home country beverage multinational firm. The production and distribution of beer is the main activity but also operates in the non-alcoholic soft drinks with a national revenue above 13 billion.
- Case 5 - BEER2 - A beverage multinational firm operating in a host country. The company produces beers, distillates, wines and soft drinks with global revenue above US$ 20 billion.
- Case 6 - COSM1 - A home country cosmetic multinational firm. The company manufactures skincare, fragrances, and bath products with annual revenue above US$ 3 billion.
- Case 7 - TOBC1 – A tobacco multinational firm that bought a national company. The company produces cigarettes in more than 200 countries. In Brazil, firm´s market share is 80% of the legal market (Bloomberg, 2018).

Data gathering

Data gathering was performed through in-depth interviews from a semi-structured protocol (Barratt et al., 2011). Managers from several operational functions (purchasing, logistics, IT, product development, among others) were invited, and at least three interviews were conducted per company, that provided a broader view of the institutional pressures perceived in each organization. This resulted in greater reliability in the information collected, reducing the risk of limited understandings by the areas of performance or subjective factors of the interviewee.

Table 1: Interviews done

Case	Company	Interviewee ID	Operational functions	Time
Automotive	AUTO1	AUTO1E1	Product development	60 minutes
		AUTO1E2	Product development	70 minutes
		AUTO1E3	Logistics	40 minutes
		AUTO1E4	Information management	40 minutes
Automotive	AUTO2	AUTO2E1	Product development	60 minutes
		AUTO2E2	Purchasing	40 minutes
		AUTO2E3	Purchasing	40 minutes
		AUTO2E4	Logistics	40 minutes
		AUTO2E5	Logistics	30 minutes
Automotive	AUTO3	AUTO3E1	Planning	30 minutes
		AUTO3E2	Planning	50 minutes
		AUTO3E3	Product development	70 minutes
Beverage	BEER1	BEER1E1	Logistics	40 minutes
		BEER1E2	Transport management	40 minutes
		BEER1E3	Manufacturing	30 minutes
Beverage	BEER2	BEER2E1	Distribution	40 minutes
		BEER2E2	Distribution	40 minutes
		BEER2E3	Logistics	40 minutes
		BEER2E4	Process management	60 minutes
		BEER2E5	Process management	60 minutes
		BEER2E6	Process management	60 minutes
		BEER2E7	Distribution	40 minutes

Table 1: Interviews done				
Case	Company	Interviewee ID	Operational functions	Time
Cosmetics	COSM1	COSM1E1	Supply chain management	40 minutes
		COSM1E2	Logistics	30 minutes
		COSM1E3	Product development	30 minutes
Tobacco	TOBC1	TOBC1E1	Logistics	60 minutes
		TOBC1E2	Logistics	40 minutes
		TOBC1E3	Transport management	40 minutes
		TOBC1E4	Security management	60 minutes

Thus, we interviewed 29 professionals working in these seven companies, with an average duration of approximately 40 minutes per interview (Table 1). Of these interviews, 24 were permitted to be recorded, while the other 5 were not allowed by the interviewees. In these, the record was made through field notes. The data collection period occurred between August 2017 and February 2019. After the data collection, the recorded interviews were transcribed.

Data Analysis

Thereafter, all transcriptions and notes were coded and analyzed according to Corbin and Strauss (2008) by open coding strategy through Atlas IT software, using the techniques of individual and cross-case analysis (Eisenhardt, 1989). These techniques presuppose that the researcher follows these steps: detailed transcription and notes of the interviews, the codification of concepts, tabulation, analysis of the data, and, finally, interpretation of the results. The inductive method was adopted as a way of exploring the phenomena in this research. This considers the process of theorization from the emergence of constructs coming from the field under study, and not being guided by any theoretical model (Randall & Mello, 2012). Three steps were used to compile the data analysis. First, the researchers identified the categories present in the interviews and their respective subcategories through the Atlas TI software. Second, according to these categories, they were refined to follow a convergence in the presentation of the results. Third, the microanalysis of relationships and meanings was used to report the phenomenon. Thus, a unique analysis model emerged, allowing the phenomenon to be understood more broadly, simultaneously, reinforcing the internal validity of the analysis (Patton, 2002).

Replication

One concern of the researcher in conducting a case study is to ensure the accuracy of the information collected. According to Yin (2014), this process begins before the researcher goes to the field. The construction of the research protocol and case selection are factors used to ensure the reliability of the study results. We also conducted some qualitative criteria to ensure the quality of the multiple case study (Table 2), in a positivistic approach (Lincoln & Guba, 1985).

Table 2: Quality criteria application		
POSITIVISTIC CRITERIA	DEFINITION	HOW TO COPE WITH THE CRITERION
Internal validity	The degree to which findings correctly map the phenomenon in question (Lincoln & Guba, 1985).	- Seven leading MNEs from four important industries; - Interviews with professionals from different operational areas, with at least three interviews per company; - Propositions presented; - The theoretical framework presented.
External validity	The degree to which findings can be generalized to other settings similar to the one in which the study occurred (Lincoln & Guba, 1985).	- Specification of the unit of analysis and the context; - Dense context description (within-case analysis)
Reliability	The extent to which findings can be replicated or reproduced by another investigator (Lincoln & Guba, 1985).	- Research protocol provided in the Appendix, including the semi-structured questionnaire; - Documentation of all the procedures performed during the case studies; - Data Analysis through proper software (Atlas TI); - Transcription of the interviews; - Use of multiple researchers.
Objectivity	The extent to which findings are free from bias (Lincoln & Guba, 1985).	- Raw material presented, including interviewees quotation; - Create a case selection criterion.

Thus, we selected companies from different sectors representative of the Brazilian economy. To avoid that the findings in the survey did not represent the entire company, we tried to do at least three interviews per case. Besides, we looked for profiles from different areas of operations. Finally, the research protocol used in this study was made available in the Appendix.

RESULTS

The data analysis of this research was divided into two sections following Eisenhardt (1989): the first one, a within-case analysis, each of the six cases context were deeply analyzed; and in the second section, a cross-case analysis will be presented, where the main categories and subcategories that emerged from the interviews were evaluated, and an analytical model and interaction between constructs will be presented. Citations throughout this section are referenced at the end by a code that refers to the code of the respondent of origin.

Within-case Analysis

<u>AUTO1 - Multinational of the Automotive Industry</u>

Historically, this company has defined the location of its operations based on the tax incentive policies offered by local governments in less developed areas of the country. Unlike other

companies, that settled in the more industrialized region, it was necessary to create a supplier base of their own.
AUTO1 today is […] off the map of most suppliers, it has a belt of (near) suppliers here, it is a risk because with the falling market as we had a while ago, suppliers could break. It could paralyze the production line and everything, but it is a risk assessed as low for people compared to competitors, even because it is off-axis there in São Paulo, Rio, and other regions, it is isolated here and we have more control over it (suppliers) (AUTO1E1).
With the opening of new plants, more distant geographically, the risk of rupture increased dramatically, which led the company to build facilities for its suppliers within the plant. That strategy was a way to mitigate the risk of rupture of the production line, also the company created a supplier evaluation system to monitor the financial health of local suppliers.
When it was opened here (a place far from the industrialized area), it was decided to build what we call of "Supply Park". The Supply Park is a total of 9 strategic suppliers, which we bring here. So, we set up the plant, […], we gave them these buildings and they got the technology (AUTO1E1).
The company deals with a series of pressures from the national institutions that regulate its activities. To mitigate these external pressures, several systems have been created that mitigate uncertainties in regulation. Therefore, it takes a great effort to adjust to government changes.
Every time the government changes the legislation we have to make a systemic adjustment and the legislation changes, literally speaking, every day. To give you an idea, 25% of the commercial area effort is these adjustments, (the government) is the largest single client of our company (AUTO1E4).
Another important influence is the market instability generated by the national economic crisis, which changes the profile of the final consumer. Despite a generalized drop in sales, there was a steeper decline in cheaper products, with the luxury goods segment gaining importance in this time of crisis. This has brought internal changes by valuing the quality of the product more than seek for reducing costs.
I think the market has dropped a lot. But it also changed a lot, cars getting more expensive, AUTO1 betting more on quality, so when I joined the group was already undergoing this modification. Today it is much stronger. When we went to the maintenance of material we were under a very great pressure that was to choose the cheapest things. Today, we have a much greater pressure to choose the things that work, it does not matter so much. Obviously, that it has cast management, but we are not as tied to the cost (AUTO1E2).
Therefore, the company uses different processes of exploration resources to remain competitive. The exploitation process is used by the company to accommodate some direct pressures from the institutional environment such as regulations. Thus, the company creates systems and routines that can mitigate these changes, however, without generating improvement in competitiveness. On the other hand, the installation of factories in remote regions has generated the company's need to develop a strategy to mitigate logistics problems. Thus, the creation of the structure for the suppliers near the factory generated productivity gains, by offering greater control of the company production process. Therefore, the company uses both exploitation and exploration processes in the hierarchical renewal strategy.

AUTO2 - Multinational of the Automotive Industry

This company was greatly affected by the 2008 economic crisis. Because of this, it underwent a long process of global restructuring, including processes, products, and costs in its all subsidiaries. Since the Brazilian national crisis, the company has taken the lead in production and has already announced several investments in the country. Despite this, political instability is still viewed with caution.

I see the AUTO2 appearing, we learned from mistakes (that lead to crises in 2008) and now we are the market leader and we will not give up easily, we will fight tooth and nail. We are investing a lot of money […]. We have had a crisis in recent years in Brazil and we are announcing a lot of investment in Brazil. Take advantage of the opportunity of having a crisis, to have an idle capacity of the supplier, tooling supplier and manage to negotiate better prices for us because we have money. And he who has the money to invest ends up doing good business. […] (AUTO2E4).

The disruption in supply has become the biggest pressure for this company. As a result of the Brazilian economic crisis, several suppliers went bankrupt or became indebted. The market has shrunk approximately 70% in the last 3 years, and the local suppliers were very threatened by this drop in the market.

For example, there is a hole in the car and you have to plug the hole to not let in water, so it's this type of plastic or air ducts. So, the plastic industry, for example, is an industry with few players and many (of those) players with financial difficulties. This is a risk, it is a delayed chain compared to outside of the country (AUTO2E2).

To mitigate this pressure, the supplier's financial instability, the company has been conducting its evaluation through an external consultancy. Those suppliers who are in trouble are indicated for a recovery plan. As an alternative, the company seeks new suppliers, national and international. Therefore, as a way to maintain its national operations, the company increases its dependence on international suppliers.

Those (suppliers) who are at financial risk we do have (evaluation system). We have a score, we hire a consultancy that does this evaluation. We try to develop some plan to recover this supplier or to develop a plan B (another supplier), a secondary source. Or eventually, in certain cases, we develop a final alternative source and we leave that supplier (AUTO2E2).

Today the risks we have are lack of qualified suppliers. [...] then today here in Brazil we have less productivity, some suppliers of medium quality, compared (to international suppliers) ... Of course, all of the level, but out there the level is higher than here and productive is much higher. So, the risk today is the supplier in the chain, a (quality) little below the market out there (AUTO2E2).

The main pressure of the external environment is suppliers' bankruptcy. The company uses the exploration process by creating supplier assessment mechanisms and financial assistance. Thus, the company uses internal resources to strengthen the supply chain, gaining competitiveness, adopted a managerial selection strategy.

AUTO3 - Multinational of the Automotive Industry

This company was one of the first to install facilities in Brazil in 1950, and still has its main operations in the most industrialized region of the country. One of its plants works with a model called modular, in which strategic suppliers were invited to develop the plant together, which ended up generating managerial skills aimed at flexibility. This allows the company to deals better with the instability of the domestic market.

We have the flexibility/agility to tailor our volume up and down. When you have a need for fitness, you divide that responsibility into 8 partners. So, you work with agility to solve the issues very fast and on the other hand, you have a constant struggle, in a good way, to reduce costs, because everyone wants to. In fact, reduce costs with the same qualities to be able to generate, for the customer, a value-added perception of our product. This as competitiveness is a big differential (AUTO3E2).

This type of model is seen by the interviewees as something irreversible, that is, the skills generated in this plant cannot be copied in the other plants owned by the company. Despite that, the financial health of national suppliers is considered a great pressure for this company also.

We have suffered from 2 or 3 suppliers there because of financial risk. It turns out, the guy stops giving you the piece and says "I do not have the money to buy the raw materials". It's almost a knife in your neck, from where we come in with money to buy the raw material for the guy, some suppliers are at that level. Of course, you start a plan to change supplier because you cannot live like this (AUTO3E3).

For the interviewees, the contribution of foreign capital differentiates the capacity of the multinational in overcoming an economic crisis in front of the local suppliers. Thus, multinationals can deal with market fluctuations in a given country, since the other subsidiaries do not suffer directly from national crises in other countries.

A problem of the economic crisis here, the corporate headquarters inject capital, they make a capital contribution to hold the fight. A normally global supplier also has this benefit. I worked with a supplier, I say this with propriety, in a moment of crisis there was a capital injection, so you can ride the wave in the moment of crisis, taking damage, but you will not close because the market oscillates. Today you are losing and tomorrow you win again. That's what happens with automakers, but there are many domestic suppliers where the guy has no cash flow and many suppliers have broken into many crises (AUTO3E3).

The factory modular model offers the company the ability to flex production quickly, which brings efficiency to deal with abrupt market reductions, which fits as an exploration process. Regarding supplier bankruptcy, no supplier assistance mechanisms were identified, only evaluation mechanisms, identifying an exploitation process. Thus, we can identify that a hierarchical renewal strategy has been adopted.

<u>BEER1 - Brazilian Multinational in the Beverage Industry</u>

The case presents a company that has an operational competence in costs based on a standardized process, as well in logistics of distribution in the interred country. However, the risks production stoppages caused by suppliers increase over time but was reported that occur mainly because of the low quality of the inputs. Therefore, it was necessary to create a national supplier assessment system, which mitigated this pressure.

Employing this evidence, the company developed a ranking system of the suppliers, to mitigate the risks of delivery in terms of inputs quality. This system is centralized in the matrix, although there are local rankings due to variations of deliveries by regions and quality delivered from the same supplier. On the one hand, the ranking helped the company to monitor the inputs received; on the other hand, this ranking has now demanded a higher level of quality management and supplier deliveries (BEER1E3).

Also, this company noticed a movement of small companies boosting the beverage market from the consumption of the artisanal brewery. As a response to this market, the company started to acquire an ample amount of handcrafted companies located in Brazil, the United States, and Europe. With an already established distribution system, the company now has handicraft products in addition to traditional mass-produced products.

But, finally, [...] there are several sectors, but as the Brazilian profile is changing more for craft beer we started to invest more in craft beer and premium beer. [...] If you have a market you have to invest. Until, for example, they are thinking ahead, I do not know if they have even started selling, for example, hops malt for handmade producers (BEER1E2).

Additionally, this company deals with problems related to the lack of public security, such as cargo theft. To do this, the company contracts insurance and makes use of monitoring systems that were initially purchased to reduce distribution costs. Despite this, there is no more active mitigation by the company, due to the wide distribution network and the costs associated with the use of armed escort, for example.

BEER1 does not have such aggregate value within the product. So, I think the cost of an armed escort and the risk it brings to people is too great. We work with insurance for freight, for distribution and transfer to the center. We do armed escorts at the factory park. In the distribution, it doesn't. It is not worth financially and not worth in for the operations. We have a very large operation to have an armed escort for everyone (BEER1E2).

To deal with cargo theft pressure, the company adopted a strategy of selecting systems and practices that were not initially developed to deal with this problem. Thus, the company does not gain competitiveness in the distribution over its competitors, using an exploitation process and a naive selection strategy.

BEER2 - Multinational of the Beverage Industry

The second case in the beverage industry deals with a multinational company that has aggressively entered the national market in the last five years. This company strategy was to buy local breweries, besides bringing their original brands to the national market. Compared with direct competitors, the company shows high-quality products. However, the merger gives rise to some pressures concerning the brand, the processes, and quality levels.

We talk a lot about departments and each department has a strategy and sometimes this strategy is a bit out of alignment with reality, and now that it is merging, there is still a lot of internal alignment problems of just in terms of production, which product is produced and which is not (BEER2E1).

Given the risk of reputation (image to the customer), the company started to develop its processes with rigid standard procedures, enlargement and greater participation of the laboratory in the control of the process to guarantee normality in production batches over time. Considering a set of biological elements involved in the manufacture of beverages, the laboratory-raised the quality control of the process, and in turn, allowed to raise the quality management of the final product (BEER2E4; BEER2E5; BEER2E6).

Since the acquisition by the multinational, the national plants have developed a more effective preventive maintenance in order to ensure a reduction in unforeseen losses and higher standards of product quality. This allowed the company to affirm its image of product quality in the Brazilian market.

Despite this, the company handles problems common to the national beverage market. According to an interviewer (BEER2E1), there are cases of product counterfeiting in the industry where lower value products are changed with the more expensive product labels. This practice, although not very significant, is very difficult to restrain, since clients and employees of the company do not perceive the falsification through simple verification of the product.

If we discover a product of ours, as (brand of the multinational), which is a widely used brand (for counterfeiting). It is the most used brand. In fact, they buy our product to transform into another. The guy buys (brand of the same group) to turn it into (a premium brand of the same group), or in (a mainstream brand of BEER1). What happens? Selling a cold beer in the bar is impossible to detect. We say it's impossible because we have a tasting panel there, trained people, master brewer, who have difficulty differentiating one beer from the other. It is very easy this forgery, the perception of the end customer is very difficult. We help, always have police investigation, etc., we contribute, obviously do not agree, but this happens in the market (BEER2E1).

The product counterfeiting does not aggressively affect the company, no new practices or system was adopted to mitigate this pressure. Thus, the company does not affect the external environment, adopting a naive selection strategy.

COSM1 - Brazilian Multinational of the Cosmetics Sector

This case deals with a company that acts with high levels of sustainability. With their suppliers located in several regions of the country, there is a strong pressure related to the climatic issue. This pressure does not only affect the probability of a disruption of the production flow, but also the distribution based on direct sales. "[...] A region that has large levels of flooding, or regions that are subject to landslides, this makes it difficult for the consultant to deliver" (COSM1E1). Quality is seen as the company's main performance as well as its delivery capability. To ensure the quality of the inputs, the company started to buy inputs from local suppliers in the Amazon Forest. As these inputs are part of its market positioning, the company seeks to develop these suppliers. Thus, the company sought to open a plant near these suppliers. To this end, it sought fiscal benefits from the local government because the region has a low level of development and a large demand for investments and employment.
That's why I told you, not only the tax aspect, the monetary aspect, in the first place and the rest of the world that gets screwed up (in deciding where to install a factory). It has a variable that the company looks at. We had a small base in (Amazonian city) where we processed all the oils that came from Amazonia. [...] In this ecosystem, we collected rainwater, the treatment of the water is made entirely natural, only with the plant, has no chemical, a whole structure on top of that factory. [...] So we went to the government and said, "Look, we have this strategic vision of a factory in [the Amazonian city], near the forest, then minister, what profit do we get?" It is not that I build the factory that the incentive will be given (COSM1E2).
However, this geographic dispersion also generates problems related to the strategic positioning of the company as a sustainable organization. Thus, any activity that increases the company's carbon emissions goes against the internal guidelines, which causes the company to seek new ways to distribute.
This indicator (related to sustainability) for us greatly affects the environmental impact ... we already have goals for that. The annual employee bonus is tied to the percentage of carbon emissions, cost, quality, and normal supply chain (which allows) to know how we can reduce environmental impact. Then there is the concern if I send something to Bahia (northeast state) by boat or by car or I deliver in the region with electric car or bicycle and not with a Fiorino (a type of truck) (COSM1E2).
The strategy of positioning as a sustainable company has generated a series of long-term internal changes. In this way, the company's routines had to deal with changes to suit the company's new standards. The result was that the company gained prestige in the market, using products from the Amazon Forest produced by local communities. Thus, we can understand that the company used an exploration process adopting a hierarchal selection strategy.

TOBC1 - A Tobacco Multinational Firm that Bought a National Company

This is a national origin company that was acquired by an MNE. It is the leader in the tobacco industry with over 80% of the market share. However, the legal market represents only 60% of tobacco sales in the country. The other products are illegally imported by the Brazilian border with Paraguay. Thus, the company had to develop systems that help the national security forces to counter this external pressure.

These management systems have become very useful when cargo theft started to increase in the industry. In particular, the cities of Rio de Janeiro and São Paulo, this pressure became so strong that it was necessary to create a system of convoys protected by an armed escort. However, there were conflicts between thieves and security agents that resulted in death. Because of that, this system had to be reviewed in the city of Rio de Janeiro.
The company, therefore, has the power to influence the institutional environment by working together with the government agency to mitigate the problems related to cargo theft. These strategies, along with enhancing distribution security, improve your positioning against competitors. Thus, the company adopts the strategy of hierarchical renewal, as it manages to develop systems and new routines to deal with this pressure. Also, it manages to include internal partners and government agencies in its strategy.

Cross-case Analysis

After analyzing the cases individually, an analysis of all cases was carried out (Corbin & Strauss, 2008). Thus, we began the analysis of the open coding of the 29 interviews, and of the more than 100 different codes found. It evolved into a final model with three main categories, six properties, and twelve dimensions, present in the following tables. The main results pointed to a concern of the managers regarding the competences adaptation to the uncertain institutional environment. In the view of the managers, there is a general uncertainty with the actions taken by the government, especially in public policies and regulations. This uncertainty has become more serious when the country has entered an economic crisis in recent years.
Therefore, the competencies of these multinationals needed to adjust to the unstable environment, and this has become an important competitive advantage for the success of these companies' strategies in that country. Thus, we decide to pursue this path of understanding how the process of coevolution takes place. To interpret the results, we proposed three relevant categories: (1) Institutional environment influence; (2) Capability adaptation; and (3) Companies' influence.
This section continues to analyze in-depth each of the three categories encountered, and the relationships between their respective properties and dimensions. The tables summarize this information based on a sample of proof quotes. From these relationships, the prepositions of this study emerge in the next section (Discussion and Conclusions).

<u>Institutional Environment Influence</u>

The institutional environment is the source of several influences for the companies located in emerging countries (DiMaggio & Powell, 1983; Sapienza, Autio, Geoge, & Zahra, 2006). The institutions of these countries are distinct (Meyer et al., 2009), and may generate different pressures in MNEs (Meyer & Peng, 2016). Lewin and Volberda (1999) argue that the corporate strategy is defined by the influence of these pressures.
Uncertainties can be perceived differently by different industries because they generate pressures that affect them differently. For example, cargo theft strongly affects the tobacco company (TOBC1), has been identified as less harshly in the cosmetics industry and has little influence on the automotive industry. Infrastructure, however, seems to affect all companies located in regions far from the economic center of the country (AUTO1; BEER1; BEER2; COSM1). These factories are attracted by tax benefits offered by the home state or by the central government to develop the poorer regions. However, in addition to increasing the distance to the consumer market, these regions are lacking in quality infrastructure.
Once the pressures were identified, it was possible to identify which of them had more influence, and the level of impact on operations. Some institutional pressures, such as union

power (BEER2), had little power of influence and impact on operations (AUTO1). In contrast, public policies had a high level of influence and impact on operations. An example of this was in the automotive sector, which received tax incentives and increases in capacity. After the national economic crisis, the government couldn't maintain this policy, the demand was reduced, and the companies are now overcapacity (AUTO1E3).

Table 3: Institutional Environment Influence, its properties, and dimensions

Properties	Dimensions	Proof quotes
Uncertainty level	Low level	*In some regions, we do have unions that are more active for everyone and everything, and some localities have a stronger alignment. I have an alignment very close to the union, and one of the ways I managed to resolve the workday was to bring the union into the Distribution Center […] That is very much the way it is handled and how it is negotiated, but it has serious interventions (by the union) in some regions, especially in the transport area* (BEER2E3). *Generally, what is always being discussed in these (union) protests that I see is usually always related to salary. I've never seen any kind of advertisement, no kind of protest regarding non-labor rights. [..] I've never seen any kind of fight, nothing offensive, police, nothing* (AUTO1E2).
	High level	*... and you talk about financial risk, yes because (tax legislation) changes so much, it's very complex, we buy more than 6, 7 thousand pieces to make a car. So, you have (different legislation in) several countries around the world including here from Brazil, and in Brazil states that have different legislation* (AUTO2E4). *Inovarauto's (national public policy for the automotive sector) energy efficiency program aimed to have more than 15% improvement in energy efficiency. [..] Unfortunately, the government linked the program to steps in the productive process and this gave a feeling of a kind of protection to the local market. Which caused this measure to be questioned in WTO and we lost* (AUTO1E4). *We do not have tax benefits and every market works like this. One minute you have it, and the next it is gone. Then it's a problem. It may also be as the country is, in crisis, increasing the tax on beer. It is possible too, I am speaking more about the past. In fact, many tax increases were for alcoholic beverages. Alcoholic beverages and soda. In fact, these products are not essential for survival, so it makes sense for the government to raise taxes (BEER2E6).*
Influence impact	Direct impact on operations	*[..] law labor risk is one of the things that most change the factory. There is a case here that is emblematic, a woman came to work in high hills. Then she stumbled onto a sidewalk in the factory. [..] (Government) work safety time came and lowered a rule so no one else could use high hills (in the factory). We had to change all the sidewalks in the facility. So, labor risk always influences and drives a lot of big companies and, especially, an automotive company*

Table 3: Institutional Environment Influence, its properties, and dimensions		
Properties	Dimensions	Proof quotes
		(AUTO1E4). *There are tax classification rules that determine the percentage of nationalization of an item accordingly... and each has a tax rate, so ascertaining this correctly also is a significant job (AUTO1E4).* *I will give you an example, the government has created a program called the employment support program, which is a PES that allows companies to negotiate, with their class unions, a reduction in working hours and wages, and subsidized 50% of this reduction. [..] This alternative is one that allows you to make the journey more flexible, making the volume of production more flexible at an attractive cost for you, the employee and the government. For the government, because it guarantees employability, for us we can adjust the volume of production according to the market, at an attractive cost, and for the employee who has the guaranteed job and continues to receive more than he works. This is an example* (AUTO3E2).
	Indirect impact on operations	*So, it is extremely critical (supplier bankruptcy). What happens is that in Brazil there are many national suppliers. Nationals (suppliers) have a strong cash flow problem. Recently AUTO3E4 even promoted a meeting between suppliers and BNDES (government national bank) to encourage it to grant credit to these suppliers* (AUTO3E4).

However, the pressure does not impact all sectors in the same way. For example, TOBC1 and BEER2 identified that cargo theft was a risk affecting their distribution. Thus, through the routing, the most recurrent places of this event were identified. The same risk was not considered relevant by the companies AUTO1, AUTO2, and AUTO3, therefore, they do not carry out actions of the same nature.

In general, institutional pressure with a lower degree of severity tends to have a timely monitoring process, and any mechanism for effective monitoring is implemented. Unlike risks that impact the operations and flow of the company's products, monitoring is constant and, in general, accompanied by other mitigation mechanisms. For example, AUTO3 identified that outdated infrastructure was a lifetime risk for its products, causing the engineering team to increase the accuracy of quality tests before putting the trucks on sale (AUTO3E1o). According to Cuervo-cazurra et al. (2018), the MNEs evolve their practices due to the pressures of different institutional environments, which suggests a capability development.

The severity of the operational impacts is evaluated by managers to define their adaptation efforts (Carney, Dieleman, & Taussig, 2016). For example, the economic crisis had a very strong impact on suppliers in the automotive sector, with the risk of supplier failure becoming critical. Thus, several actions were carried out, such as the financial monitoring of suppliers. The same risk was identified as low by the beverage and cosmetic sectors, with little effort in that direction.

Capability Adaptation

The process of capability adaptation happens when a company performs actions to deal with the identified pressures. However, different strategies to deal with unique pressure could lead to different results. It was identified that some companies agreed to set up their factories in isolated regions of the country to receive a tax incentive from the state that promotes this policy (AUTO1; BEER2; COMS1). Due to this, these companies reacted to a government incentive without taking measures to reduces the logistics risk (BEER2). This not necessarily contributing to an improvement of their strategic positioning or a capability creation, but it reduces operational cost.

Table 4: Capability Adaptation, its properties, and dimensions

Properties	Dimensions	Proof quotes
Capabilities adaptation level	Low level	*AUTO1 engineering we develop radio, an instrument panel that are the ones we say are the original factory, but the same person who develops these radios too develops alternative radios that we call (incomprehensible). Then we develop the originals and know that they will be in the cars and everything else and we also develop other brands. Then AUTO1 itself develops the radios that it sells as alternatives. So maybe it's even to avoid counterfeiting in the parallel market and everything. The same engineer who developed the original, he also develops a similar cheaper* (AUTO1E2).
	High level	*Obviously, the road infrastructure is a problem for the trucker. In general, our products undergo the most rigorous tests here, and some of them simulating rough and/or poor paving, and the idea is always to have a product fit to work with the real conditions of Brazil. Obviously, everything has its limits. So, I would say that, in a general way of infrastructure, we have authorized stations, we have dealerships spread all over Brazil and our products undergo the most rigorous tests before they enter the field. The field is also faithfully accompanied by our group of technical and quality assistants, and when necessary, the engineering team also enters to be able to make the evaluations (AUTO3E1).*
Uncertainty severity	Little impact	*[…] the radios are assigned to work in a specific car. If you take one car and put it in another it does not work, so this ends up inhibiting the interest of theft. […] So, it happens, but it's punctual or something that worries us. I can tell you that it's one or two trucks that happen a year and the guy (thief) sometimes even sees the load and gives up the theft* (AUTO2E4). *Yes, but I think it is not in such a large proportion (forgery of products). We have already seen some seizures here and it is not a volume that justifies so much action on this (BEER2E3).*
	Critical impact	*We already had there in 2011, 2012 reduction of IPI that gave a boom in the market and I think even harmful because we had a significant increase in sales, increase in productive capacity, in*

Table 4: Capability Adaptation, its properties, and dimensions		
Properties	Dimensions	Proof quotes
		investments and it did not stay like that. Today we have idle capacity. AUTO1, for example, works with 60% of total capacity (AUTO1E3). *The Brazilian market was already the 5th largest car market in the world and was turning into 4th when [the factory] came to be. The market experts predicted that by 2018 we would reach 5 (million) units, this was much more than the installed production capacity of all manufacturers at the time, and everyone started to build factory after factory. Unfortunately, all forecasts were wrong, and we went 12 years back in terms of production volume, with several factories becoming idle (at the production level) as well* (AUTO1E4).

The same incentive, fiscal benefit, was used by COSM1 to strengthen its cosmetics production in the Amazon. Thus, the plant was installed close to its local suppliers, through the strategy of positioning itself as a sustainable company and defender of the Amazon Forest (COSM1E2). Here, the company acted actively in the use of public policy to reinforce its strategy, creating an institutional capability (Carney, Dieleman, & Taussig, 2016). On the other hand, AUTO1 mitigate the supply risk by created facilities for strategic suppliers in its own factories, which brings them more control to the entered process (AUTO1E1).

Companies' Influence

As predicted by the coevolutionary theory, firms do not act as passive agents to the environment but influence it (Cantwell, Dunning, & Lundan, 2010). This influence can only be reacting to the environmental pressures to contain these influences (exploitation), but it can undergo profound changes in its competences that have generated competitive advantages (exploration).

Table 5: Companies' Influence, its properties and dimensions		
Properties	Dimensions	Proof quotes
Ambidexterity process adopted	Exploitation	*For example, our factory in [distant state] today has its incentives and it's worth it for us to be there. If we are there, it is because there is a reason, certainly when we talk about costs. In the first place of any project, BEER2 considers taxes. I think for everyone, there is no longer an option to overlook taxes, tax incentive, concerning to planning a project* (BEER2E5).
	Exploration	*Suppliers are divided into groups. When [the factory] was opened [in a distant state], it was decided to build what we call supply park. Supply park is a total of 9 strategic suppliers, which we bring here. So, we set up the plant, built and supplied them with*

Table 5: Companies' Influence, its properties and dimensions		
Properties	Dimensions	Proof quotes
		the structures, and they came in with the technology (AUTO1E1). *The guy (supplier) is here on your side, the guy stays inside our park. So, he goes on to say so, our rules, such as uniform is the same pattern like ours, follows our standard of quality, the restaurant they eat is the same as ours… Such interaction is very great for us. So, it is very easy to solve a problem, it gave a problem you automatically in 2 minutes is already in the supplier's factory to solve. He has a great relationship* (AUTO1E1).
Influence impact	Direct	*It is an institutional relationship. We visit the police stations [..] present what we have done for safety and what we need to support them, in a reaction force, in research work* (TABC1E4).
	Indirect	*No, so much so that the government's role ... we always treat the government as a partner for the company in terms of activity, we had a financing with the BNDES recently that is still in force, that is important for the development of new products, in short, the sector regulation, as I said before, it goes through discussions in ANFAVEA, that is, they put an intermediary to meet or discuss market actions and in general I would for operation, government participation is more in terms of financing or something of that nature* (AUTO3E1).

This process is associated with each company's absorptive capacity strategy. Short-term decisions that do not imply future benefits, such as reducing operational costs with fiscal incentive, are associated with exploration strategies. However, when the company reacts to external pressure and manages to create a capability that brings competitive advantages, this is associated with the exploitation strategy (Birkinshaw & Gupta, 2013).
The influence of corporate strategies on the institutional environment can happen indirectly or directly. Indirectly, in general, it happens in sectors that have stronger associations, such as the automotive sector. This sector defines jointly with the government public policies through its national association, the ANFAVEA. In sectors in which a company has great market share, like tobacco, it acts individually, defending only the corporative interests of that company.

DISCUSSION AND CONCLUSIONS

Abrupt changes in legislation, economic crises, or changes in government public policies can aggravate information asymmetry and long-term strategy (Meyer & Peng, 2016; Yamakawa, Peng, & Deeds, 2008). Companies with operations in an uncertain institutional environment, as emerging countries, seek to adjust practices and strategies to respond to those pressures (Oliver, 1991).
To avoid disruptions in the companies' activities, all cases show some perception of potential institutional pressures. These pressures are perceived by the company according to their level

of operational impact. Institutional pressures can lead to capabilities adaptation (Carney, Dieleman, & Taussig, 2016). However, your findings suggest that pressure considered more important (high level of impact) is associated with greater effort to adaptation. Therefore, the choice of the process of exploitation of resources occurs in the function of the impact that the pressure in the company. Thus, it is possible to establish the following propositions:

Proposition 1. Uncertainties in the institutional environment are highly associated with direct and indirect impacts on the company.

Proposition 2. The choice of resource exploration process is highly associated with the level of impact of uncertainties.

The results of the research also suggest that companies exert influence in the institutional environment in two ways. This influence can occur directly, in which the company plays a decisive role in the promotion of public policies such as financing (AUTO2) or security strategy (TOBC1). However, companies may use industry associations to exert influence on public policies, such as the automotive sector. This suggests that the composition of the market can determine the channels of influence, therefore, very representative companies in their industry (TOBC1 and BEER1) tend to seek direct channels. Thus, it is possible to establish the following proposition:

Proposition 3. The type of market composition is highly associated with the direct or indirect influence of the company in the institutional environment.

In this research, we investigated how MNEs responded to pressures caused by institutional uncertainty, identifying how the influence between companies and institutional environment affects capability development in emerging countries. The main findings suggest that the impact level of uncertainty leads companies to develop new capabilities (exploration) or strengthen existing ones (exploitation).

As a limitation of this paper, the analysis does not allow for speculation on the difference of pressure response between national and international MNEs. Only in the beverage industry the selected cases reflecting this dichotomy. Another possibility is to understand how incoming companies adapting their capabilities to foreign markets in the first years. Another possibility for future research would be the monitoring of a single company during a longitudinal period, observing the creation and adaptation of its operational capabilities. These non-answered gaps could be subjects to be explored in future related research.

APPENDIX

Research Protocol

1) Research proposal

This paper aims to understand how the capabilities development process in MNEs is affected by the uncertain environment in a Coevolutionary Theory approach.

2) Search question

The research question of this study was: how does the uncertain institutional environment affect operational capabilities' coevolution process?

3) Theoretical bases of the research

- Coevolutionary Theory;
- Ambidexterity
- Institutional Theory.

4) Selection of cases

The main criteria for the selection of companies were:

- MNEs that have greater power of influencing in their respective markets;
- Leaders in the national and global market, of the automotive (3 companies), beverages (2 companies), cosmetics (1 company), and tobacco (1 company) sectors.

5) Data collection

- Interviews with managers recorded and transcribed;
- At least three interviews by each company.

6) Data Analysis

- Definition of codes in an open way through the technique of content analysis;
- Analysis based on within-case and cross-case analysis.

Semi-structured Questionnaire

Introduction
Full name
Position
The interviewee history in the company
How do you see the company in the market: Leader? In what categories? Who are the main competitors?

Part 1 - INTERNAL FACTORS
1) What are the main internal risks that the company manages?
2) What does the company do internally to manage them?
3) What competencies were created/adapted from risk management? Examples.
4) What motivates the company to develop such skills? Describe internal and external factors (internal factors: managers, processes, supply chain, etc.).

Part 2 - EXTERNAL FACTORS
5) What factors outside the firm motivate the creation/adaptation of company skills? What is the role of each? (Competitors, suppliers, sector, institutions, regulatory etc.).
6) How does the company respond to external influences?

Part 3 - PERFORMANCE
7) Regarding the operational performance such as cost, quality, flexibility, and delivery, which of them do you consider that your company performs better than your competitors? Why?
8) How did the risk capability development contribute to this performance? How did internal and external factors contribute to this?
9) In those who think that they perform better, which practices do they support? How do you use such practices? Give examples.
10) How the risks described influence the performance of the business or the business unit.

Part 4 - INSTITUTIONAL ASPECTS
1) How does labor legislation affect the competitiveness of the company? Is there a risk to trade unions?
2) Is there any risk concern to tax regulation?
3) How the sector legislations are defined? Does the company have an active voice in this definition?

4) Has there been any change in your operations to deal with crime? (Theft of cargo; counterfeiting).
5) How does the issue of infrastructure affect your operations?

REFERENCES

Barratt, M., Choi, T.Y. and Li, M. (2011). Qualitative case studies in operations management: trends, research outcomes, and future research implications, *Journal of Operations Management*, Vol. 29 No. 4, pp. 329-342.

Birkinshaw, J., & Gupta, K. (2013). Clarifying the distinctive contribution of ambidexterity to the field of organization studies. *Academy of Management Perspectives*, 27(4), 287-298.

Bloomberg L.P. (2006). Stock price graph. Retrieved Dez. 8, 2018 from Bloomberg database.

Bruton, G. D., Ahlstrom, D., & Li, H. L. (2010). Institutional theory and entrepreneurship: where are we now and where do we need to move in the future?. *Entrepreneurship Theory and Practice*, *34*(3), 421-440.

Cantwell, J., Dunning, J. H., & Lundan, S. M. (2010). An evolutionary approach to understanding international business activity: The co-evolution of MNEs and the institutional environment. *Journal of International Business Studies*, 41(4), 567-586.

Carney, M., Dieleman, M., & Taussig, M. (2016). How are institutional capabilities transferred across borders?. *Journal of World Business*, *51*(6), 882-894.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research*: Techniques and procedures for developing grounded theory.

Cuervo-Cazurra, A., Ciravegna, L., Melgarejo, M., & Lopez, L. (2018). Home country uncertainty and the internationalization-performance relationship: Building an uncertainty management capability. *Journal of World Business*, *53*(2), 209-221.

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48, 147-160.

Doh, J., Rodrigues, S., Saka-Helmhout, A., & Makhija, M. (2017). International business responses to institutional voids. *Journal of International Business Studies*, 48 (3), 293–307.

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 18(4), 586-608.

Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review*, v. 14, n. 4, p. 532-550.

Eisenhardt, K. M. (1991). Better stories and better constructs: The case for rigor and comparative logic. *Academy of Management review*, *16*(3), 620-627.

Estrin, S., Meyer, K. E., & Pelletier, A. (2018). Emerging Economy MNEs: How does home country munificence matter?. *Journal of World Business*, *53*(4), 514-528.

He, X., Brouthers, K. D., & Filatotchev, I. (2018). Market orientation and export performance: the moderation of channel and institutional distance. *International Marketing Review*, *35*(2), 258-279.

Karamchandani, A., Kubzansky, M., & Lalwani, N. (2011). THE GLOBE-Is the Bottom of the Pyramid Really for You?-Tackling the world's most challenging market is not for every company—but some innovative businesses are finding ways to serve low-income consumers profitably. *Harvard Business Review*, 107.

Khanna, T., & Palepu, K. (1997). Why focused strategies may be wrong for emerging markets. *Harvard Business Review*, 75 (4), 41–51.

Kristal, M. M., Huang, X., & Roth, A. V. (2010). The effect of an ambidextrous supply chain strategy on combinative competitive capabilities and business performance. *Journal of Operations Management*, 28(5), 415-429.

Lampel, J., & Shamsie, J. (2003). Capabilities in motion: New organizational forms and the reshaping of the Hollywood movie industry. *Journal of Management Studies*, 40(8), 2189-2210.

Lewin, A.Y., & Volberda, H. (1999). Prolegomena on Coevolution: A Framework for Research on Strategy New Organizational Forms. *Organization Science*, 5, 519-534.

Lincoln, Y. S., & Guba, E. G. (1985). Establishing trustworthiness. *Naturalistic Inquiry*, 289, 331.

Mair, J., Marti, I., & Ventresca, M. J. (2012). Building inclusive markets in rural Bangladesh: How intermediaries work institutional voids. *Academy of Management Journal*, *55*(4), 819-850.

Madhok, A., & Liu, C. (2006). A coevolutionary theory of the multinational firm. *Journal of International Management*, 12(1), 1-21.

Marquis, C., & Raynard, M. (2015). Institutional strategies in emerging markets. *The Academy of Management Annals*, 9(1), 291-335.

McKelvey, B. (1997). Perspective—Quasi-natural organization science. *Organization Science*, 8(4), 351-380.

Meredith, J. (1998). Building operations management theory through case and field research. *Journal of Operations Management*, *16*(4), 441-454.

Meredith, J.R., Raturi, A., Amoako-Gyampah, K. and Kaplan, B. (1989). Alternative research paradigms in operations, *Journal of Operations Management*, Vol. 8 No. 4, pp. 297-326.

Meyer, K., Estrin, S., Bhaumik, S., & Peng, M. (2009). Institutions, resources, and entry strategies in emerging economies. *Strategic Management Journal*, 30(1), 61-80.

Meyer, K., & Peng, M. (2016). Theoretical foundations of emerging economy business research. *Journal of International Business Studies*, 47 (1), 3-20.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340-363.

Oliver, C. (1991). Strategic Response to Institutional Processes. *Academy of Management Review*, 16(1), 145-179.

O'Reilly, C. A., III, & Tushman, M. L. (2004). The ambidextrous organization. *Harvard Business Review*, 82(4), 74.

Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, *1*(3), 261-283.

Parmigiani, A., & Rivera-Santos, M. (2015). Sourcing for the base of the pyramid: Constructing supply chains to address voids in subsistence markets. *Journal of Operations Management*, *33*, 60-70.

Peng, M. (2002). Towards an institution-based view of business strategy. *Asia Pacific Journal of Management*, 19 (2-3), 251-268.

Peng, M. (2003). Institutional transitions and strategic choices. *Academy of Management Review*, 28 (2), 275-296.

Peng, Mike W.; Lee, Seung-Hyun; Wang, & Denis YL. (2005). What determines the scope of the firm over time? A focus on institutional relatedness. *Academy of Management Review*, v. 30, n. 3, p. 622-633.

Peng, M., Wang, D., & Jiang, Y. (2008). An institution-based view of international business strategy: a focus on emerging economies. *Journal of International Business Studies*, 39 (5), 920-937.

Ramamurti, R. (2012). What is really different about emerging market multinationals? *Global Strategy Journal*, 2(1), 41-47.

Randall, W. S., & Mello, J. E. (2012). Grounded theory: an inductive method for supply chain research. *International Journal of Physical Distribution & Logistics Management*, *42*(8/9), 863-880.

Rodrigues, S., & Child, J. (2003). Co-evolution in an institutionalized environment. *Journal of Management Studies*, 40(8), 2137-2162.

Rottig, D. (2016). Institutions and emerging markets: effects and implications for multinational corporations. *International Journal of Emerging Markets*, *11*(1), 2-17.
Sapienza, H. J., Autio, E., George, G., & Zahra, S. A. (2006). A capabilities perspective on the effects of early internationalization on firm survival and growth. *Academy of Management Review*, *31*(4), 914-933.

Scott, W. R. (2008*). Institutions and organizations*: Ideas and interests. Sage.

Sodhi, M.S., Tang, C.S., (2009). *Managing supply chain disruptions via time-based risk management*. In: Wu, T., Blackhurst, J. (Eds.), Managing Supply Chain Risk and Vulnerability. Springer, London, pp. 29–40.

Stuart, I., Mccutcheon, D., Handfield, R., Mclachlin, R. and Samson, D. (2002). Effective case research in operations management: a process perspective, *Journal of Operations Management*, Vol. 20 No. 5, pp. 419-433.

Su, Z., Peng, M. & Xie, E. (2016). A strategy tripod perspective on knowledge creation capability. *British Journal of Management*, 27 (1), 58-76.

Wheelwright, S. C., & Hayes, R. H. (1985). Competing through manufacturing. *Harvard Business Review*, 63(1), 99-109.

Wu, S. J., Melnyk, S. A., & Flynn, B. B. (2010). Operational Capabilities: The Secret Ingredient. *Decision Sciences*, 41(4), 721-754.

Yamakawa, Y., Peng, M., & Deeds, D. (2008). What drives new ventures to internationalize from emerging to developed economies? *Entrepreneurship: Theory and Practice*, 32 (1), 59.

Yin, R. K. (2014). *Case study research, design, and methods* (4th edition). London: Sage Publications.

DECISION SCIENCES INSTITUTE

Optimum Buffer Size of Steel Strip Coils in a Continuous Pipe Production System

Md. Shahriar J. Hossain
Northwestern State University
Email: hossains@nsula.edu

Bhaba R. Sarker
Louisiana State University
Email: bsarker@lsu.edu

ABSTRACT

The performance of a continuous steel pipe manufacturing system can be improved by solving different operational problems in the system. One of the important problems involves determination of buffer capacity of the coil accumulator which acts as an inter-stage buffer storage of steel strip that slowly feeds into the rolling benders to form the steel pipes. This research work is oriented in modeling this part of coil-feeding operation before forming the pipe and solving this by ensuring minimum costs for building and maintaining the sufficient buffer storage space to keep up the normal flow of raw materials to the system.

KEYWORDS: Buffer Inventory, Cost minimization, Operations management, of Continuous production system.

INTRODUCTION

Continuous seam welded pipes are produced from steel strip coils. These steel strips are first uncoiled and straightened. Then the steel strip is welded with the end of preceding strip to maintain a continuous flow of it as needed by processes. In order to avoid any disruption of production due to this joining operation a buffer inventory is held in between joining and forming stations [Figure 1] (Hossain and Sarker 2018). This buffer zone is known as coil accumulator which can hold steel strips for future use as a safety stock.

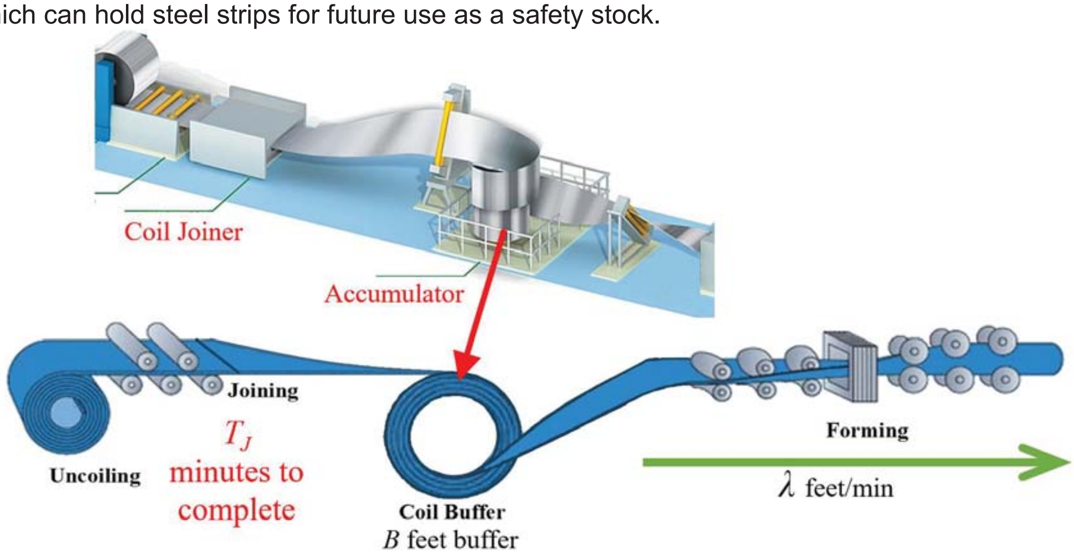


Figure 1. Buffer storage (accumulator) before the forming station.

The coil feeding system comprises of coil joining and coil accumulation. While a new steel coil is fed into the line through a welding joint, the process needs not to be stopped, rather the steel strip is continuously delivered to the main manufacturing line from a coil accumulator. The size of the allowable buffer inventory (buffer capacity) directly depends on the strip joining time. When feed rate of the line (roughly the pipe production rate) is given by λ feet/hour and the coil accumulator (buffer) can accommodate *B* feet (buffer size) of steel strip, the operators at the joining station gets maximum B/λ hours to complete their job. If they fail to complete the joining process by this time, the whole line must be stopped.

In order to avoid the line stoppage or downtime, the runout time of steel strips at the coil accumulator should be kept within an acceptable limit. High buffer inventory at the coil accumulator provides higher flexibility of time to carry out the joining operation. However, an increased buffer capacity results extra holding and operational cost, in addition to the initial fixed cost for buffer accumulator construction. So, an appropriate buffer capacity can be traded off between the buffer holding cost for maintaining the strip flow and downtime cost for potential loss of production. On the other hand, the time required for coil joining process is not deterministic rather shows some variabilities. Now, the problem is, therefore, to find an appropriate buffer capacity of the coil accumulator that can minimize the total cost of coil feeding facility. The buffer runout probability which in fact directly affects the buffer runout cost, eventually affects the value of optimum buffer capacity. Hence, determining an acceptable limit for buffer runout probability is another objective of this research work.

LITERATURE SURVEY

When needed, the in-line buffer storage is an intermediate stock of materials in the line that helps maintain a continuous flow of the material to the next subsequent station. An in-line buffer provides flexibility to avoid the shortage of raw material supply (here, strip of coil) for a station, and hence, reduces the disruption of the continuous process. Oversized buffer causes higher cost for holding the materials (in addition to initial large erection/commission cost), whereas a small buffer can cause station idleness due to lack of material leading to lower throughput and efficiency. The optimum size of the buffer capacity varies depending on the process and system parameters. Several research works were conducted in last few decades, in order to investigate the effect of the buffer storage on the line efficiency and to obtain an optimum size and location of the buffer in different production systems.

Effect of Buffer Storage on the Line Efficiency

The buffer storage in the multistage series production systems has been analyzed for more than fifty years. In the early stage, Hillier and Boling (1966) presented a detailed discussion and mathematical analysis about the effect of buffer storage on the production rate of two, three or four station production line. Next year, Buzacott (1967a) claimed that the efficiency of an automatic transfer line can be increased by dividing the line into a number of stages and providing buffer stocks between the stages. He (1967b) presented a model for predicting the line efficiency and showed how the line efficiency depends on the buffer capacity. Knott (1970) also considered a production line as a series of finite buffer at each station and developed a simple formula for determining the inefficiency of the line. N. P. Rao presented analytical studies for several aspects of two-stage (Rao 1975) and three-stage (Rao 1976a) production systems with inter-stage buffer storage. He presented the mathematical analyses for Erlang service times (Rao 1976b), as well as generalized distribution of service times (Rao 1975, 1976a). In a study, Panwalkar and Smith (1979) devised an expression to predict the average output rate of a series of multiple workstations with inter-stage finite buffers and exponential service times.

The intermediate buffer storage in a continuous process is analyzed in literatures for last few decades. Alvarezvargas *et al.* (1994) discussed about some simulation results to compare the production rate of the continuous flow model as an approximation of the asynchronous model. They considered manufacturing flow lines with deterministic processing times, unreliable machines, and finite buffers are considered. Tan and Gershwin (2009, 2011) used the fluid flow models to evaluate the performance of a two-stage production line with an intermediate finite buffer. They determined the production rate and the expected buffer level from the steady-state distributions. However, Bierbooms *et al.* (2013) analyzed continuous production lines consisting of a number of machines in series with a finite buffer between each pair of machines. They precisely estimated the throughput and the mean buffer content of the production line. Ozdogru and Altiok (2015) proposed a decomposition approximation method to study the steady-state behavior of a continuous material flow system operating in bulk port marine terminals, in terms of down time probabilities and the system throughput.

Buffer Allocation and Optimization

The optimization of buffer capability and location was one of the important topic in several researches. Anderson and Moodie (1968) analyzed the steady-state production line systems using simulations, to develop a total cost model. They derived an expression for an optimum capacity of the buffer storage that minimizes the total cost. Basu (1977) also developed a total cost model, for exponential service time, in order to optimize the buffer storage capacity. Sarker (1984) presented a critical review and a comparative study of some design aspects of the series production systems with a special emphasis on the design of inter-stage buffer stocks. Seong *et al.* (2000) formulated a non-linear programming for the buffer allocation problem in a continuous flow production and proposed an algorithm to solve the problem. Fu and Xie (2002) derived estimators of throughput sensitivity to changes in buffer capacity for continuous flow models. Rajaram and Tian (2009) formulated a non-linear integer program for optimizing the location and size of buffers in semi-continuous manufacturing processes.

Turki *et al.* (2013) presented continuous and discrete flow models for a single-product manufacturing system with a buffer storage and stochastic demand. The objective of their study was to find optimal buffer level, which minimizes the total expected cost. Kolb and Gottlich (2015) dealt with the buffer allocation problem consists of a production process combined with stochastic processing times. They derived a time-continuous model supplemented with a stochastic process. Gebennini *et al.* (2015) dealt with the performance evaluation of a two-machine one-buffer production line with restart policy modeled as a continuous time Markov process. The mathematical model was described along with its analytical solutions. Fitouhia *et al.* (2017) considered a two-machine continuous flow manufacturing system with a finite buffer capacity. They developed methodology to analyze the trade-off between the preventive maintenance of the machine and the buffer size. Hosseini and Tan (2017) presented a mixed-integer linear programming approach for determining the minimum buffer capacity that achieves a desired production rate and maximizes the profit for a continuous-flow production system. On the other hand, Dolgui *et al.* (2017) presented an optimization problem for determining the size of discrete product buffers to minimize the total cost of a production line. In that constrained optimization problem, they considered average production rate, costs for installing buffers, and the inventory cost, as process parameters. They solved the problem with metaheuristics.

Limitations of Previous Researches on Buffer Capacity

Though several research works are found on buffer storage optimization in continuous process flow, the available models are either approximated from discrete product flow or focused in performance evaluation of the process. In the continuous steel pipe manufacturing process, the buffer storage of steel strips in the line involves huge expenses for space, power and maintenance. It also involves, acceleration, deceleration of the input, random joining time and line stoppage due to buffer stock out. None of the available literature specifically addressed these issues in optimization of buffer capacity, especially matching with the continuous pipe production system. Hence, an analytical formulation of a buffer storage (coil accumulator) with random processing time, is needed to be developed and solved optimally.

ASSUMPTIONS

The mechanics of buffer operation at the coil accumulator is further clarified with some specific assumptions summarized below.

(a) The coil joining time follows a normal distribution with known mean and standard deviation.
(b) An acceptable limit for line stoppage or down time is specified by the level of significance for buffer runout.
(c) All coils that are fed into the line, are of the same strip length.
(d) Input rate to the coil accumulation is higher than the output rate from the accumulator.
(e) The strip input process involves the same value (α) for acceleration and deceleration, which incurs a power consumption cost proportional to the absolute value of α .
(f) The steel strip accumulator can be built in any dimension.

NOTATIONS

The buffer accumulation problem is formulated mathematically using some notations. Before illustrating the problem formulation, these notations are defined in four categories.

(a) System Parameters

λ	Strip output rate from the coil accumulator (feet/hour)
c_J	Cost of coil joining (dollars/hour)
c_λ	Unit cost for strip feeding to the coil accumulator [dollars/hour/ (feet/hour)]
c_α	Unit cost for acceleration/deceleration change [dollars/hour/ (feet/hour^2)]
h_B	Buffer strip holding cost at the coil accumulator (dollars/feet/year)
c_B	Operational cost for the buffer storage in coil accumulator (dollars/feet/year)
c_{td}	Line stoppage or downtime cost ($/hour)
D_c	Demand of coils (coils/year)
f_{T_J}	Probability density function (pdf) of T_J
L_c	Length of a coil (feet/coil)
T_J	Coil joining time (hours/coil)
$\bar{T}_J$	Mean time to join a coil (hours/coil)
σ_T	Standard deviation of joining time (hours/coil)

λ_{in}	Rate of strip feeding to the coil accumulator (feet/hour)
α	Acceleration for strip input to the coil accumulator (feet/hour^2)

(b) Intermediate variables

λ_{net}	Resultant rate of coil accumulation (feet/hour)
$\bar{I}_B$	Average buffer inventory level (feet)
T_c	Cycle time (hours/coil)
F_{T_J}	Cumulative density function (cdf) of T_J

(c) Variables

B	Buffer capacity (feet)
β	Level of significance for buffer runout

(d) Performance measures

TC^B	Expected total cost for the buffer storage system (dollars/year)

MATHEMATICAL MODEL FOR TOTAL COST OF COIL FEEDING SYSTEM

The coil feeding system includes fixed operational cost and the cost of capital for the coil accumulator, holding cost for the buffer, coil joining cost, power consumption cost involved in acceleration-deceleration and input-rate of steel strips to the coil accumulator, and downtime cost for strip runout.

The time gap between two consecutive coil joining operations determine one feed-cycle of the coil accumulation process. Throughout the cycle, the output rate (λ) of steel strip from the coil accumulator remains constant, which is the same as the feed rate of steel strips to the main manufacturing line. In a complete feed-cycle, the coil accumulation (in the coil accumulator) process can be segmented into six different timeslots over the time scale. Each of these timeslots are indicated graphically in Figure 2. The feed-cycle in Figure 2 starts at time $t = 0$ and end at time $t = t_6$, where the points of time t_1 , t_2 , t_3 , t_4 , t_5 and t_6 make partitions for the six time-slots containing six different patterns of coil accumulation.

(i) *Acceleration*: At time $t = 0$ the joining of previous coil is completed and input of a new steel strip to the coil accumulator is started. An acceleration (α) of coil input (to the accumulator) is started at $t = 0$ and continues until a desired coil input rate (λ_{in}) is achieved at the time point t_1 . Thus, when $0 < t \le t_1$ the resultant coil accumulation rate is found as $\lambda_{net} = \alpha t - \lambda$.

(ii) *Constant accumulation*: In the second time-slot of $t_1 < t \le t_2$, the strip input to the accumulator follows a constant rate λ_{in} . At the same time the output rate (λ) of the strip is consistently maintained. Hence, the resultant accumulation rate becomes $\lambda_{net} = \lambda_{in} - \lambda$.

(iii) *First deceleration*: A deceleration ($-\alpha$) of strip input is maintained in the time-slot $t_2 < t \le t_3$. At the end of this deceleration process, the input rate becomes same as output rate, resulting $\lambda_{net} = 0$.

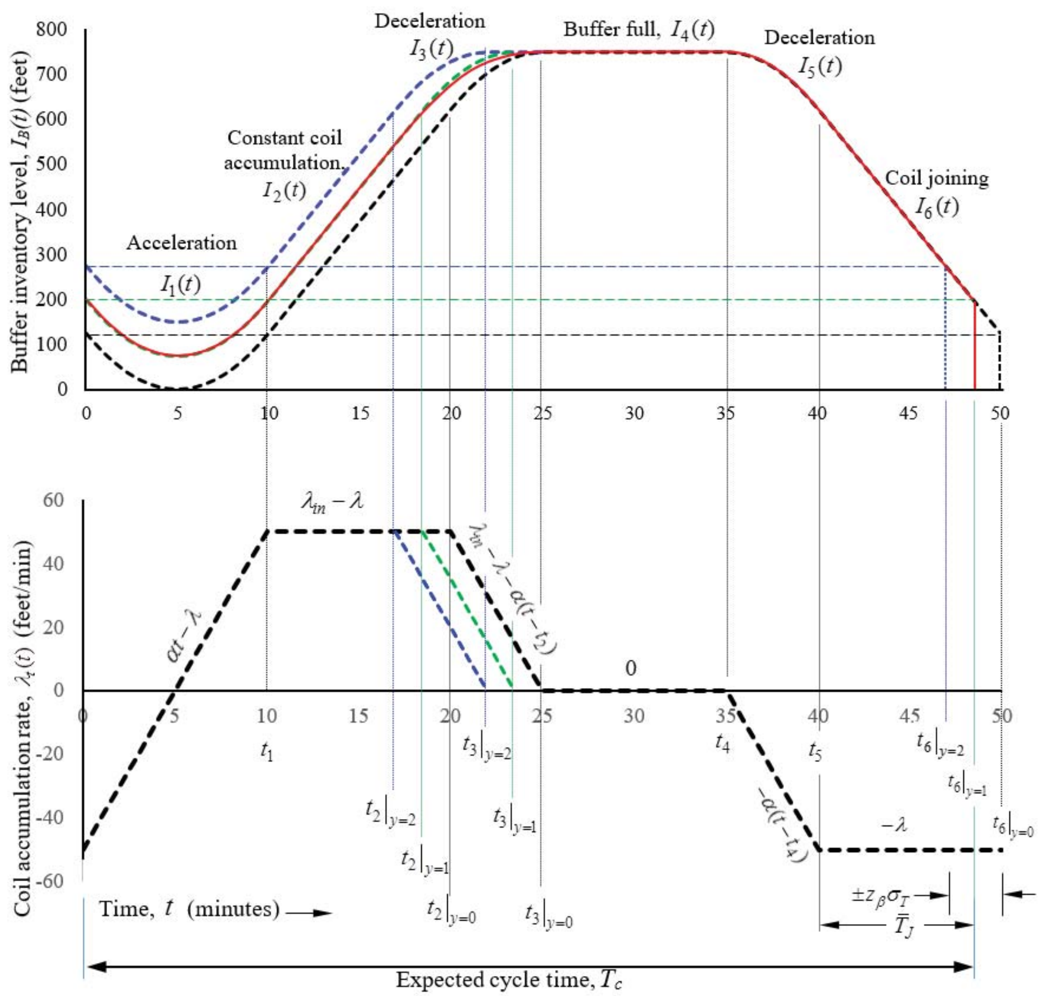


Figure 2. Buffer inventory level and coil accumulation rate over time.

(iv) *Buffer full*: At the time point t_3 input rate become equal to the output rate λ; hence, the resultant rate of coil accumulation become $\lambda_{net} = 0$. This steady state condition continues until the end of the next time segment $t_3 < t \le t_4$.

(v) *Second deceleration*: Now, in order to join a new coil with the continuing steel strip in the line, the feeding of strips to the coil accumulator has to be stopped. So, a deceleration is imposed. The deceleration continues until the input rate reaches to zero. This stage of deceleration results $\lambda_{net} = -\alpha\left(t - t_4\right)$, in the time-slot $t_4 < t \le t_5$.

(vi) *Coil joining*: At the last timeslot $t_5 < t \le t_6$, the feeding of the steel strip to the coil accumulator remains ceased, and the joining operation is conducted. This joining operation takes T_J time units. Once the joining of a new coil is completed the cycle re-starts again.

The complete expression for the resultant rate of coil accumulation, λ_{net} is expressed in Eq. (4.1).

$$\lambda_{net} = \begin{cases} \alpha t - \lambda, & \text{when } 0 < t \le t_1 & : \text{Acceleration} \\ \lambda_{in} - \lambda, & \text{when } t_1 < t \le t_2 & : \text{Constant accumulation} \\ \lambda_{in} - \lambda - \alpha\left(t - t_2\right), & \text{when } t_2 < t \le t_3 & : \text{First deceleration} \\ 0, & \text{when } t_3 < t \le t_4 & : \text{Buffer full} \\ -\alpha\left(t - t_4\right), & \text{when } t_4 < t \le t_5 & : \text{Second deceleration} \\ -\lambda, & \text{when } t_5 < t \le T_J + t_5 & : \text{Coil joining} \end{cases} \tag{1}$$

Now, expressing the size of the coil accumulator, B in terms of the length (feet) of steel strip, the total *operational* cost for the coil accumulator becomes $c_B B$ dollars/year, where c_B (dollars/feet/year) is the unit operation cost for the accumulator. Again, the *holding* cost for the buffer storage at coil accumulator depends on the expected buffer inventory level, $\bar{I}_B$. Hence, total expected holding cost can be written as $h_B \bar{I}_B$ dollars/year.

When the demand of coils for the pipe manufacturing plant is given by D_c coils/year, and the joining cost is denoted as c_J (dollars/hour), total *joining* cost in one year can be written as $D_c c_J \bar{T}_J$ where $\bar{T}_J$ (hours/coil) is the expected joining time. On the other hand, *feeding* of steel strip to the coil accumulator continues for the time-slot $0 < t \le t_4$, for which a cost of $c_\lambda \lambda_{in} t_4$ is accessed. Again, the input is accelerated and decelerated at a rate of α in the time-slots $0 < t \le t_1$, $t_2 < t \le t_3$ and $t_4 < t \le t_5$. For the *acceleration* and deceleration of the feeding, a cost of c_α {dollars/hour/(feet/hour^2)} is accessed which yields a cost of $D_c c_\alpha \alpha(t_5 - t_4 + t_3 - t_2 + t_1)$ dollars/year.

Besides these costs of operation, holding, joining, feeding and acceleration, an additional cost is incurred when the line needs to be stopped due to buffer storage runout. As the joining time (T_J) in reality is probabilistic with mean $\bar{T}_J$ and standard deviation σ_T, a runout duration can be obtained as $\{T_J - (\bar{T}_J + z_\beta \sigma_T)\}$ hours under a normal assumption. Hence, for a known probability density function $f_{T_J}\left(T_J\right)$ of joining time T_J, the total expected runout cost becomes $D_c c_{td} \int_{\bar{T}_J + z_\beta \sigma_T}^{\infty} \left(T_J - \bar{T}_J - z_\beta \sigma_T\right) f_{T_J}\left(t\right) dt$. Here, β denotes the level of significance for buffer runout, and z_β is denoted for the corresponding variable of standard normal distribution. Considering all costs involved in the coil accumulator and the coil joining process, the total expected cost $TC^B(B)$ is formulated for the coil feeding and accumulation system.

In order to avoid the buffer runout with level of confidence $1-\beta$, a minimum buffer capacity in the coil accumulator should be maintained, so that enough strip can be supplied to the main production line during, decelerations, coil joining and starting of the next cycle. The required amount of steel strip during the time-slot $t_4 < t \le t_5$ (decelarartin-2), is obtained as $(\alpha/2)(t_5 - t_4)^2$. The required amount of steel strip is $(\bar{T}_J + z_\beta \sigma_T)\lambda$, while a new coil is joined. In addition to these, an additional amount of steel strip with length $\lambda^2/2\alpha$, is needed to maintain an uninterrupted feeding to the main line when a new feed-cycle is started. Thus, a constraint is formulated as in [Eq. (2a)] that defines the minimum buffer capacity in the coil accumulator. Hence, the coil accumulation problem (CAP) becomes

CAP Problem: $\text{Min. } TC^B(B) = c_B B + h_B \bar{I}_B + D_c \left\{ c_J \bar{T}_J + c_\lambda \lambda_{in} t_4 + c_\alpha \alpha (t_5 - t_4 + t_3 - t_2 + t_1) \right\}$

$$+ D_c c_{td} \int_{\bar{T}_J + z_\beta \sigma_T}^{\infty} \left(T_J - \bar{T}_J - z_\beta \sigma_T \right) f_{T_J}(t)\, dt \tag{2}$$

Subject to $$B \ge \frac{\alpha}{2}(t_5 - t_4)^2 + (\bar{T}_J + z_\beta \sigma_T)\lambda + \frac{\lambda^2}{2\alpha} \tag{2a}$$

and $$0 < \beta < 1,\ B \ge 0. \tag{2b}$$

The buffer inventory level $I_B(t)$ in a specific feed-cycle depends on the joining time during the previous cycle. The upper and lower limits of joining time are $\bar{T}_J + z_\beta \sigma_T$ and $\bar{T}_J - z_\beta \sigma_T$, respectively. Three levels of the joining time T_J are represented with an indicator variable y defined as

$$y = \begin{cases} 0, \text{ if } T_J = \bar{T}_J + z_\beta \sigma_T \\ 1, \text{ if } T_J = \bar{T}_J \\ 2, \text{ if } T_J = \bar{T}_J - z_\beta \sigma_T. \end{cases} \tag{3}$$

Based on the definition of y, the buffer inventory level $I_B(t)$ can be obtained by integrating Eq. (1) over time t. The expressions for $I_B(t)$ can be written as

$$I_B(t)\big|_{y=0,1,2} = \begin{cases} I_1(t) = \dfrac{\lambda^2}{2\alpha} + y z_\beta \sigma_T \lambda + \dfrac{\alpha t^2}{2} - \lambda t, & \text{when } 0 < t \le t_1 \\ I_2(t) = I_1(t_1) + (\lambda_{in} - \lambda)(t - t_1), & \text{when } t_1 < t \le t_2 \\ I_3(t) = I_2(t_2) + (\lambda_{in} - \lambda)(t - t_2) - \dfrac{\alpha}{2}(t - t_2)^2, & \text{when } t_2 < t \le t_3 \\ I_4(t) = I_3(t_3), & \text{when } t_3 < t \le t_4 \\ I_5(t) = I_4(t_4) - \dfrac{\alpha}{2}(t - t_4)^2, & \text{when } t_4 < t \le t_5 \\ I_6(t) = I_5(t_5) - \lambda(t - t_5), & \text{when } t_5 < t \le \bar{T}_J + (1-y) z_\beta \sigma_T + t_5 \end{cases} \tag{4}$$

From this expression the total buffer inventory can be found as

$$\int_0^T I_B(t)dt = \int_0^{t_1}\left(\frac{\lambda^2}{2\alpha}+yz_\beta\sigma_T\lambda+\frac{\alpha t^2}{2}-\lambda t\right)dt+\int_{t_1}^{t_2}\left\{I_1(t_1)+(\lambda_{in}-\lambda)(t-t_1)\right\}dt$$

$$+\int_{t_2}^{t_3}\left\{I_2(t_2)+(\lambda_{in}-\lambda)(t-t_2)-\frac{\alpha}{2}(t-t_2)^2\right\}dt+\int_{t_3}^{t_4}I_3(t_3)dt+\int_{t_4}^{t_5}\left\{I_4(t_4)-\frac{\alpha}{2}(t-t_4)^2\right\}dt$$

$$+\int_{t_5}^{T_J+t_5}\left\{I_5(t_5)-\lambda(t-t_5)\right\}dt, \qquad (5)$$

Now, the average buffer inventory level $\bar{I}_B(t)$ at any time t can be calculated as

$$\bar{I}_B(t)=\frac{I_B(t)\big|_{y=0}+I_B(t)\big|_{y=2}}{2}. \qquad (6)$$

In case of probabilistic joining time, buffer stock is increased by a safety stock $z_\beta\sigma_T\lambda$. If $B_{\min}$ is the minimum buffer capacity without considering the safety stock such that $B_{\min}=\frac{\lambda^2}{\alpha}+\bar{T}_J\lambda$, then the required size of the coil accumulator is $B=B_{\min}+z_\beta\sigma_T\lambda$. Now, from Figure 2 it can be written as

$t_1=\frac{\lambda_{in}}{\alpha}$, $t_2-t_1=\frac{B_{\min}-I_1(t_1)}{\lambda_{in}-\lambda}-\frac{\lambda_{in}-\lambda}{2\alpha}$, $t_3-t_2=\frac{\lambda_{in}-\lambda}{\alpha}$, $t_4=\frac{L_c}{\lambda}-(t_5-t_4)-\bar{T}_J$ and $t_5-t_4=\frac{\lambda}{\alpha}$. Hence, $t_2=\frac{B_{\min}}{\lambda_{in}-\lambda}+\frac{\lambda}{\alpha}$, $t_3=\frac{B_{\min}}{\lambda_{in}-\lambda}+\frac{\lambda_{in}}{\alpha}$, $t_4=\frac{L_c}{\lambda}-\frac{\lambda}{\alpha}-\bar{T}_J$, and $t_5=\frac{L_c}{\lambda}-\bar{T}_J$.

The inventory level $I_1(t_1)$ after the acceleration period $0<t\le t_1$, can be found as $I_1(t_1)=\frac{\lambda^2}{2\alpha}+\frac{\alpha t_1^2}{2}-\lambda t_1=\frac{(\lambda_{in}-\lambda)^2}{2\alpha}$. So, from Eq. (5) and Eq. (6) the expected buffer inventory level $\bar{I}_B$ at the coil accumulator is found as

$$\bar{I}_B=\frac{1}{T}\int_0^T \bar{I}_B(t)dt$$

$$=\left(z_\beta\sigma_T\lambda+\frac{\lambda^2}{\alpha}+T_J\lambda\right)+\frac{\lambda^4}{6L_c\alpha^2}-\frac{\lambda\left(\frac{\lambda^2}{\alpha}+\bar{T}_J\lambda\right)^2}{2L_c(\lambda_{in}-\lambda)}-\frac{\left(\frac{\lambda^2}{\alpha}+\bar{T}_J\lambda\right)\lambda(\lambda_{in}+\lambda)}{2\alpha L_c}-\frac{\lambda^3}{6\alpha L_c}\left[\frac{\lambda}{\alpha}+3\bar{T}_J\right]-\frac{\lambda^2\bar{T}_J^2}{2L_c},$$

which reduces to

$$\bar{I}_B=B-\frac{\lambda^2\lambda_{in}\left(\alpha\bar{T}_J+\lambda\right)\left(\alpha\bar{T}_J+\lambda_{in}\right)}{2L_c\alpha^2(\lambda_{in}-\lambda)}. \qquad (7)$$

Again, $t_5 - t_4 + t_3 - t_2 + t_1 = \frac{\lambda}{\alpha} + \frac{\lambda_{in} - \lambda}{\alpha} + \frac{\lambda_{in}}{\alpha} = \frac{2\lambda_{in}}{\alpha}$, $t_5 - t_4 = \frac{\lambda}{\alpha}$ and $t_4 = \frac{L_c}{\lambda} - \frac{\lambda}{\alpha} - \bar{T}_J$. Hence, from Eq. (2), the final version of the revised RCAP Problem becomes

Problem RCAP:

$$\begin{aligned} \text{Min. } TC^B(B) = {} & c_B B + h_B \left\{ B - \frac{\lambda^2 \lambda_{in} \left(\alpha \bar{T}_J + \lambda\right)\left(\alpha \bar{T}_J + \lambda_{in}\right)}{2 L_c \alpha^2 \left(\lambda_{in} - \lambda\right)} \right\} \\ & + D_c \left\{ c_J \bar{T}_J + c_\lambda \lambda_{in} \left(\frac{L_c}{\lambda} - \frac{\lambda}{\alpha} - \bar{T}_J \right) + 2\lambda_{in} c_\alpha \right\} \\ & + D_c c_{st} \int_{\bar{T}_J + z_\beta \sigma_T}^{\infty} \left(T_J - \bar{T}_J - z_\beta \sigma_T \right) f_{T_J}(t)\, dt \end{aligned} \tag{8}$$

Subject to

$$B \geq \frac{\lambda^2}{\alpha} + (\bar{T}_J + z_\beta \sigma_T)\lambda \tag{8a}$$

$$0 < \beta < 1 \text{ and } B \geq 0. \tag{8b}$$

This is a non-linear programming problem, along with a constraint (8a). The non-linearity exists for the buffer runout probability β, when β is a decision variable. For a given value of β, the objective function $TC^B(B)$ is a linear function of B. In that case, the total cost decreases with buffer capacity B. However, the value of β directly affects the buffer runout cost. In other words, the total cost increases as β increases. As a result, lower value of β is desired. The value of β is also controlled by the constraint (8a). In constraint (8a) the buffer capacity B and the standard normal variable z_β are linearly related, while z_β increases with the decrease of β. So, higher value of B is necessary to reduce the buffer runout cost. These two contradicting scenarios are addressed in the formulation, from where an optimum value of B is needed to be determined.

In addition to the optimization of buffer capacity, some additional results are observed regarding the minimum limit for acceleration of strip feeding (α). These results are presented in Theorem 1.

Theorem 1: $\alpha > \max \left\{ \frac{\lambda^2}{L_c - \bar{T}_J \lambda}, \frac{\lambda_{in}\left(\lambda_{in} - 2\lambda\right)}{\lambda \bar{T}_J} \right\}$.

Proof: For time segment $t_4 > 0$, $\frac{L_c}{\lambda} - \frac{\lambda}{\alpha} - \bar{T}_J > 0$ which reduces to $\alpha > \frac{\lambda^2}{L_c - \bar{T}_J \lambda}$. Similarly,

$t_2 - t_1 = \frac{B_{\min}}{\lambda_{in} - \lambda} - \frac{\lambda_{in} - \lambda}{\alpha} > 0$, where $B_{\min} = \frac{\lambda^2}{\alpha} + \bar{T}_J \lambda$. This yields $\frac{\lambda_{in}\left(\lambda_{in} - 2\lambda\right)}{\lambda \bar{T}_J} < \alpha$. ■

SOLUTION METHODOLOGY

It is observed that, for a given value of buffer runout probability β, the total cost $TC^B(B)$ is a linear function of B for which a zero buffer (B =0) should be the optimum solution to this problem. However, a boundary condition for B is defined by the constraint (8a) which fixes the lower limit for buffer capacity as $B_l = \frac{\lambda^2}{\alpha} + (\bar{T}_J + z_\beta \sigma_T)\lambda$. Now, embedding this boundary condition in Eq. (8) the mathematical formulation of the problem becomes

$$\text{Min. } TC^B(B) = c_B B + h_B \left\{ B - \frac{\lambda^2 \lambda_{in} \left(\alpha \bar{T}_J + \lambda\right)\left(\alpha \bar{T}_J + \lambda_{in}\right)}{2L_c \alpha^2 \left(\lambda_{in} - \lambda\right)} \right\}$$

$$+D_c \left\{ c_J \bar{T}_J + c_\lambda \lambda_{in} \left(\frac{L_c}{\lambda} - \frac{\lambda}{\alpha} - \bar{T}_J \right) + 2\lambda_{in} c_\alpha \right\} + D_c c_{td} \int_{\frac{B}{\lambda} - \frac{\lambda}{\alpha}}^{\infty} \left(T_J - \frac{B}{\lambda} + \frac{\lambda}{\alpha} \right) f_{T_J}(t)\, dt, \quad (9)$$

and $B \geq 0$.

Thus, the objective function in Eq. (8) along with boundary condition in constraint (8a) forms a single variable non-linear programming problem, as shown in Eq. (9). The holding cost and the buffer stock runout cost (that causes line stoppage) can be controlled by changing the buffer capacity B (i.e., the size of the coil accumulator). Higher buffer capacity B increases holding cost whereas this increase of B reduces the buffer runout cost. These conflicting scenarios conclude that there is a optimum point for B that minimizes total cost for the coil feeding system.

The first and second derivatives of $TC^B(B)$ in (9) with respect to B are obtained as

$$\frac{d(TC^B)}{dB} = c_B + h_B - \frac{c_{td} D_c}{\lambda} \int_{\frac{B}{\lambda} - \frac{\lambda}{\alpha}}^{\infty} f_{T_J}(t)\, dt \quad (10)$$

and $\frac{d^2(TC^B)}{dB^2} = \frac{c_{td} D_c}{\lambda^2} f\left(\frac{B}{\lambda} - \frac{\lambda}{\alpha} \right) > 0$, respectively. (11)

The second derivative is positive definite for any $\alpha > 0$. So, $TC^B(B)$ is a convex function in B.

The minimum value of $TC^B(B)$ can be found by setting $\frac{d(TC^B)}{dB}$ = 0 which yields

$$\frac{d(TC^B)}{dB} = c_B + h_B - \frac{c_{td} D_c}{\lambda} \left[1 - F_{T_J}\left(\frac{B}{\lambda} - \frac{\lambda}{\alpha} \right) \right] = 0, \text{ or } F_{T_J}\left(\frac{B}{\lambda} - \frac{\lambda}{\alpha} \right) = 1 - \lambda \frac{c_B + h_B}{c_{td} D_c}.$$

The coil joining time is assumed to be normally distributed as $T_J \sim N(\bar{T}_J, \sigma_T^2)$. So $\left[\left(\frac{B}{\lambda} - \frac{\lambda}{\alpha} \right) - \bar{T}_J \right] \Big/ \sigma_T = F_{T_J}^{-1}\left[1 - \lambda \frac{c_B + h_B}{c_{td} D_c} \right]$, which finally leads to

$$B^* = \lambda\left[\frac{\lambda}{\alpha} + \bar{T}_J + \sigma_T F_{T_J}^{-1}\left(1 - \lambda\frac{c_B + h_B}{c_{td} D_c}\right)\right]. \quad (12)$$

This solution for B^* in Eq. (4.12) is optimum when a specific value of β is not assigned. On the other hand, if a particular value of β is given and the solution of B in Eq. (12) does not satisfy constraint (8a), the optimum solution will be dominated by the constraint. In the latter situation, the optimum solution B^* is thus obtained as

$$B^* = \frac{\lambda^2}{\alpha} + (\bar{T}_J + z_\beta \sigma_T)\lambda \text{ , where } z_\beta = F_{T_J}^{-1}(\beta). \quad (13)$$

The process for obtaining the optimum solution B^* is stated with a flow diagram in Figure 3. It is noticeable from Eq. (12) that the optimum buffer capacity (or the coil accumulator) B^* does not depend on coil joining cost (c_J), strip feeding cost (c_λ) , acceleration cost (c_α), rate of coil feeding to the accumulator (λ_{in}) and the length of a coil (L_c). However, these parameters affect the total cost of the coil joining and buffer accumulation. When other parameters are known, the optimum buffer capacity B^* can be directly obtained from Eq. (12).

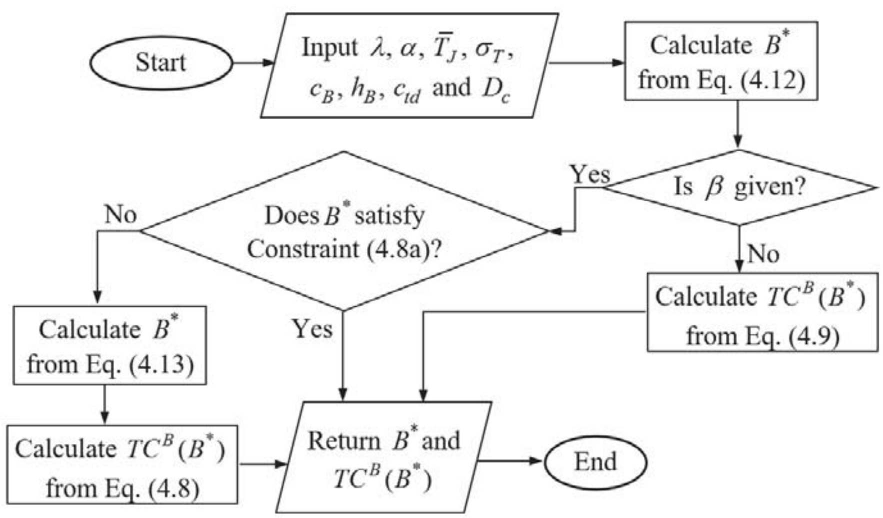


Figure 3. Flow diagram of optimum solution search.

Some numerical instances are presented in the next section for describing the solution and the sensitivity of the solution with respect to some system parameters is presented subsequently.

NUMERICAL EXPERIMENTATION

The continuous steel pipe manufacturing process involves a special type of buffer storage facility known as a coil accumulator. The size of the coil accumulator determines the buffer capacity of the system. The optimum solution for the buffer capacity can be obtained by following the flow diagram in Figure 3. A numerical illustration of this process in presented here.

Example 1: Optimum buffer capacity (size of coil accumulator)

In a steel pipe manufacturing system the coil joining time (T_J) is normally distributed such that $T_J \sim N(10,\ 0.5^2)$ in minutes. All other parameters involved in the feeding stage, are listed in Table 1.

Table 1. Input parameters for Example 1

c_B $/feet/year	c_J $/hour	c_λ $/feet	c_α $-hour/feet	α feet/hour^2	λ feet/hour	λ_{in} feet/hour	c_{td} $/hour	h_B $/feet/year	D_c coils/year	L_c feet/coil
10	600	0.10	1/30	36,000	3,000	6,000	600	20	1000	2500

The probability of buffer runout (β) is not assigned here. Hence, the optimum solution for the buffer capacity B^* can be calculated from Eq. (12). It becomes $B^* = 775.91$ feet which yields $TC^B(B^*)$ = $867,665.79 per year from Eq. (9). This optimum solution indicates that the coil accumulator should have enough space to hold a steel strip of length 775.91 feet. Both higher and lower capacity than 775.91 feet will increase the total cost for the coil joining and buffer storage facility. The optimum solution is graphically represented in Figure 4.

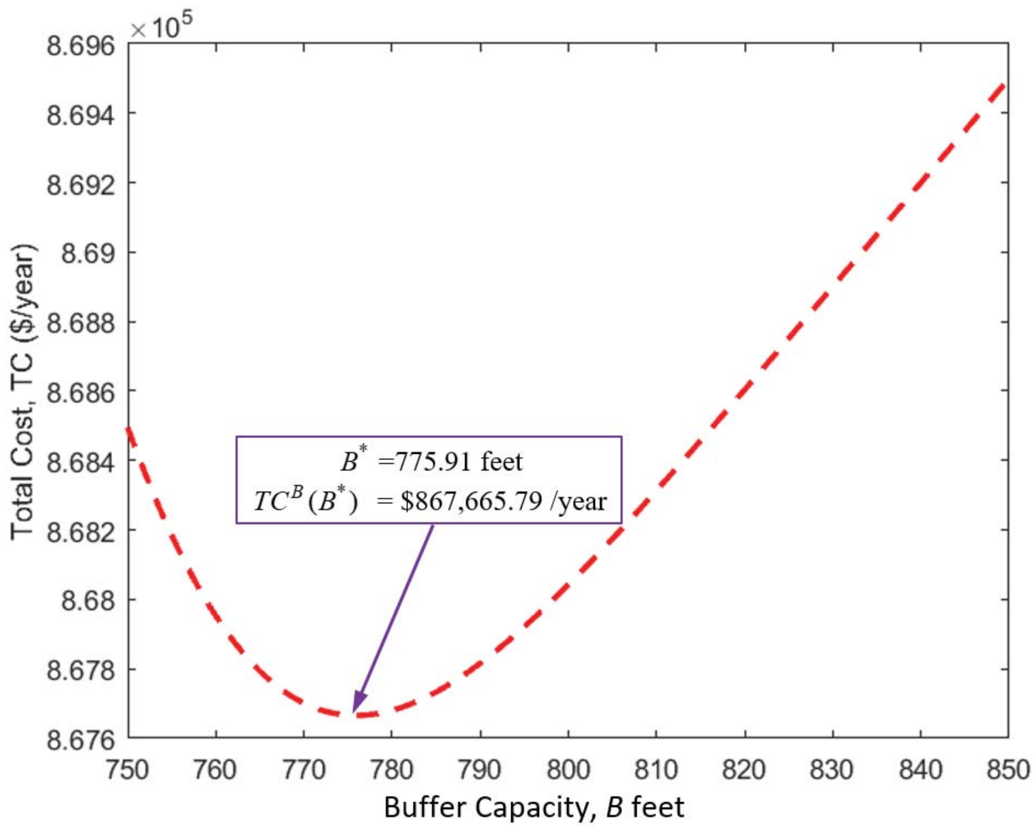


Figure 4. Optimum solution of Example 1.

SENSITIVITY OF THE SOLUTION

From Eq. (12) and Eq.(13) it is observed that, the optimum buffer capacity B^* largely depends on buffer runout probability (β), the output rate of steel strips from the coil accumulator (λ), acceleration rate of strips input (α) and coil joining time (T_J). The effects of these parameters on the optimum solution are analyzed in the following sub-sections.

Effect of β

Now, the probability of buffer runout can be calculated from the optimum solution found. From Eq. (13), $B^* = \left(\lambda^2/\alpha + \bar{T}_J\lambda\right) + \left(z_\beta \sigma_T \lambda\right)$. Here, $B_{\min} = \lambda^2/\alpha + \bar{T}_J\lambda$ is the minimum buffer capacity which is essential for continuing the pipe manufacturing process during the coil joining operation. The minimum buffer capacity is computed as $B_{\min} = 750$ feet. So, $z_\beta \sigma_T \lambda = B^* - B_{\min}$ = 25.91 feet is the amount of safety stock for handling uncertainty in joining time. From this, the standard normal variable z_{β^*} is found as $z_{\beta^*} = \left(B^* - \lambda^2/\alpha - \bar{T}_J\lambda\right)\Big/(\sigma_T\lambda) = 25.91/(\sigma_T\lambda)$ =1.0364. Thus, from the standard normal table the probability of buffer runout is found as β^* = $F_{T_J}(z_{\beta^*})$ = 0.1515. This result indicates that, if the buffer capacity in the coil accumulator is set at $B^* = 775.91$ feet, the buffer storage in the accumulator will runout 15.15% times. Though this probability is noticeably large, yet the total cost for the coil joining, coil accumulator and line stoppage (due to buffer runout) is minimum at $TC^B(775.91)$ = \$867,665.79 per year.

Any deviation from the optimal $B^* = 775.91$ feet will lead to an increase to the total cost until other related system parameters are changed. If the production manager is not happy with the runout probability β^* = 0.1515, he/she sets a different acceptable value of β = 0.10 (for example). With this new value, the solution for the buffer capacity B = 775.91 does not satisfy constraint (8a). Hence, constraint (8a) becomes effective and dominating to the optimum solution. When, under this condition, the optimum solution B^* is evaluated using Eq. (13) [See Figure 3] as B^* =782.04 feet. In this case, the total cost becomes $TC^B(782.04)$ = \$867,697.89/year which is obtained from Eq. (8).

Table 2. Optimum solutions for different limiting values of β.

SL	β	B^* feet	$TC(B^*)$ \$/year	SL	β	B^* feet	$TC(B^*)$ \$/year
1	0.0001	842.97	869,289.38	6	0.0500	791.12	867,838.10
2	0.0005	8.3226	868,968.55	7	0.1000	782.04	867,697.88
3	0.0010	827.25	868,819.06	8	0.1500	775.91	867,665.79
4	0.0100	808.16	868,261.70	9	0.1515	775.75	867,665.79
5	0.0200	801.34	868,077.03	10	0.2000	775.75	867,665.79

The effect of β-value on the solution is presented in Table 2, which is more visualized in Figures 5 and 6. Figure 5 presents a joint effect of B and β on the total cost. When a low value is imposed to the probability of buffer runout β, the optimum buffer capacity B is dominated by constraint (8a). Thus, a certain threshold value β^*, of can be determined for a given set of system parameters. Below β^* the constraint becomes tight and above β^* the constraint becomes relaxed.

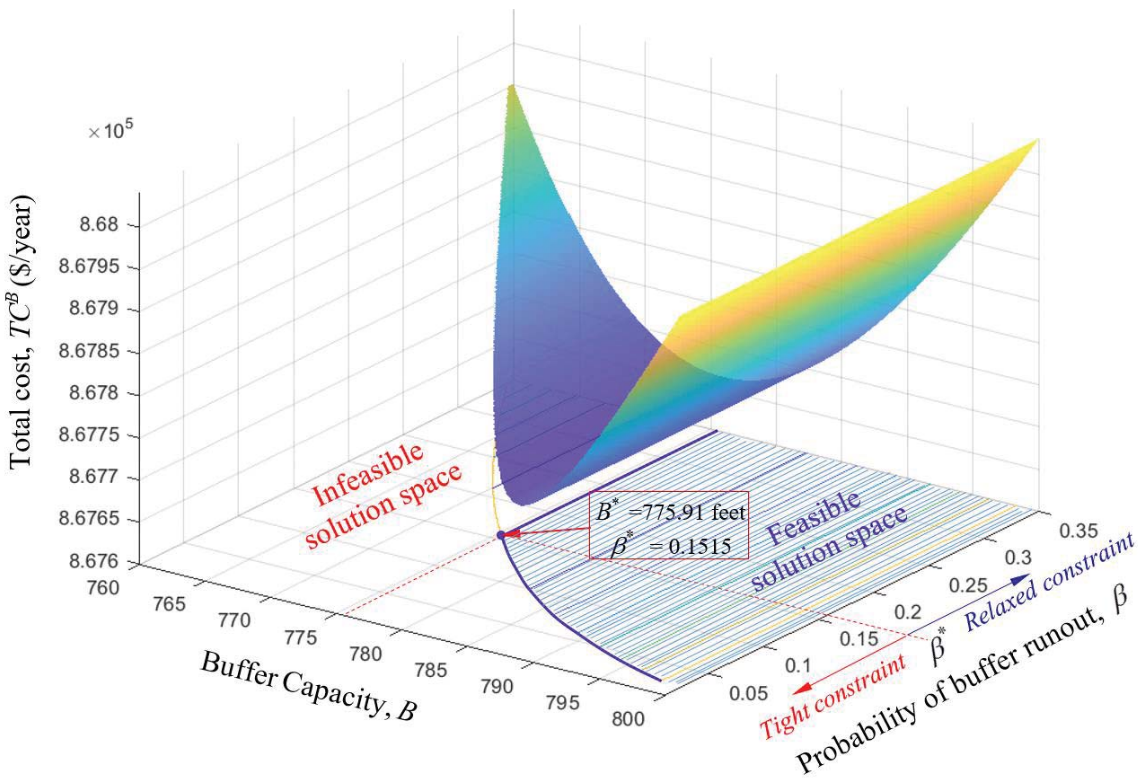


Figure 5. Variation in solution for different β value.

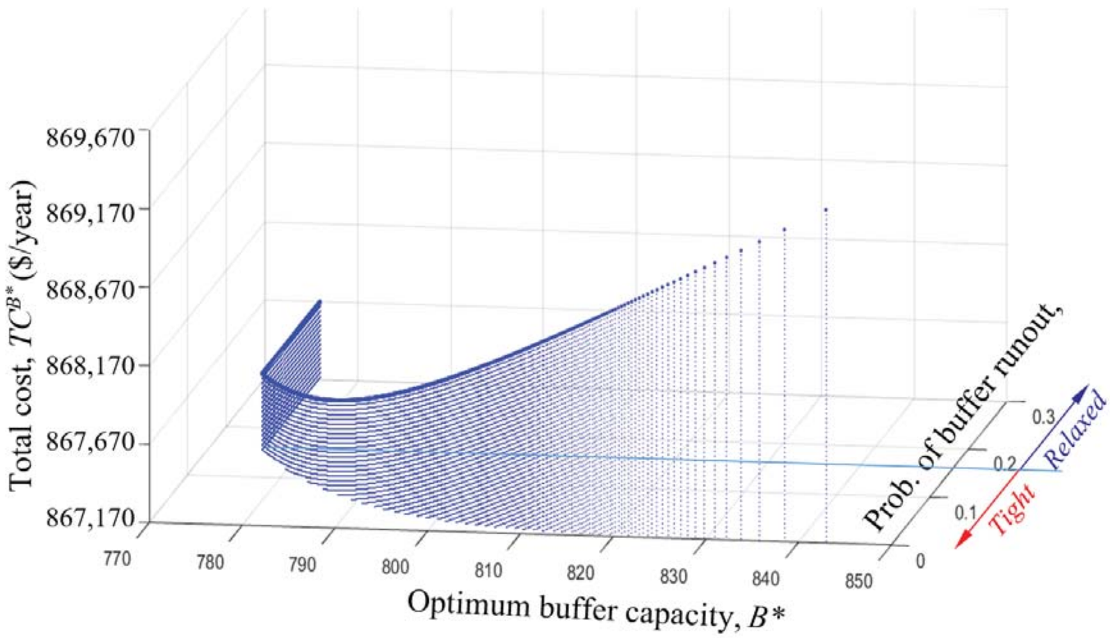


Figure 6. Effect of β value on the optimum solution.

Sensitivity of the Solution with Respect to Joining Time

The mean time for joining a coil (T_J) is a random value with known probability distribution. Both the mean ($\bar{T}_J$) and the standard deviation (σ_T) of joining time have direct effect on the optimum solution. In order to see the effects of joining time on the optimum solution, $\bar{T}_J$ and σ_T in Example 1 are varied individually. Six different values of $\bar{T}_J$ are taken; these are 5, 10, 15, 20, 25 and 30 minutes. For each of the values of $\bar{T}_J$, several solution points of B and the corresponding $TC^B(B)$ are measured. The solution points for each of the six values of $\bar{T}_J$ are plotted in Figure 7 with different colors. It is observed that, if $\bar{T}_J$ increases, the optimum total cost $TC^B(B^*)$ also increases. The increasing trend of $TC^B(B^*)$ with respect to $\bar{T}_J$ seems to be asymptotic [Figure 7].

Similarly, Figure 8 is plotted for six different values of σ_T; these are 0.5, 1.5, 2.5, 3.5, 4.5 and 5.5 minutes. The solutions corresponding to each of the values of σ_T, are represented with different colored lines in Figure 8. It is found that the optimum solutions follow a linear increasing trend with σ_T.

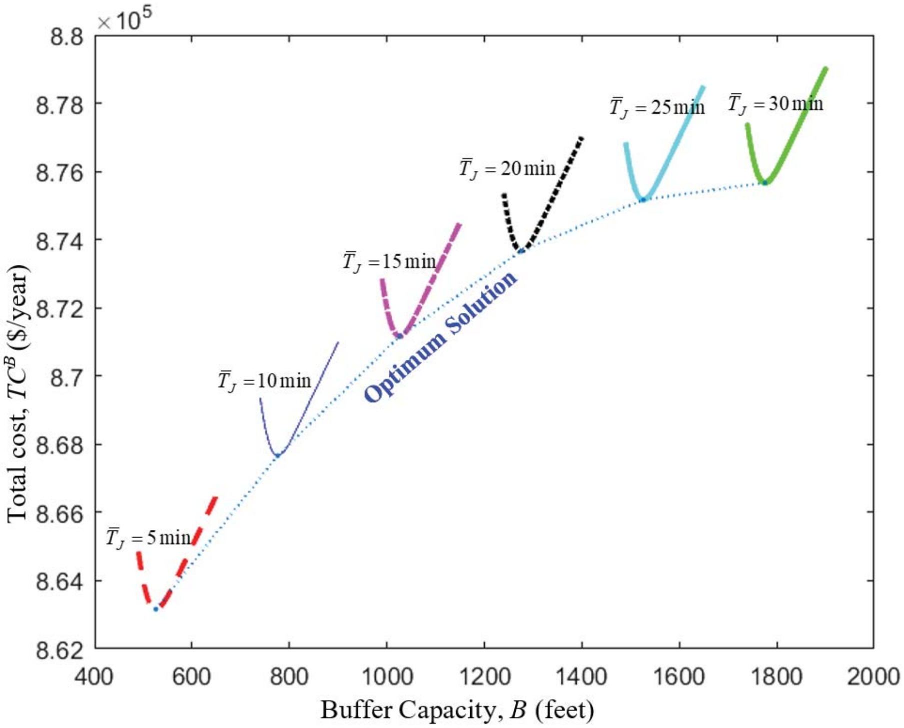


Figure 7. Sensitivity of the solution with respect to $\overline{T}_J$.

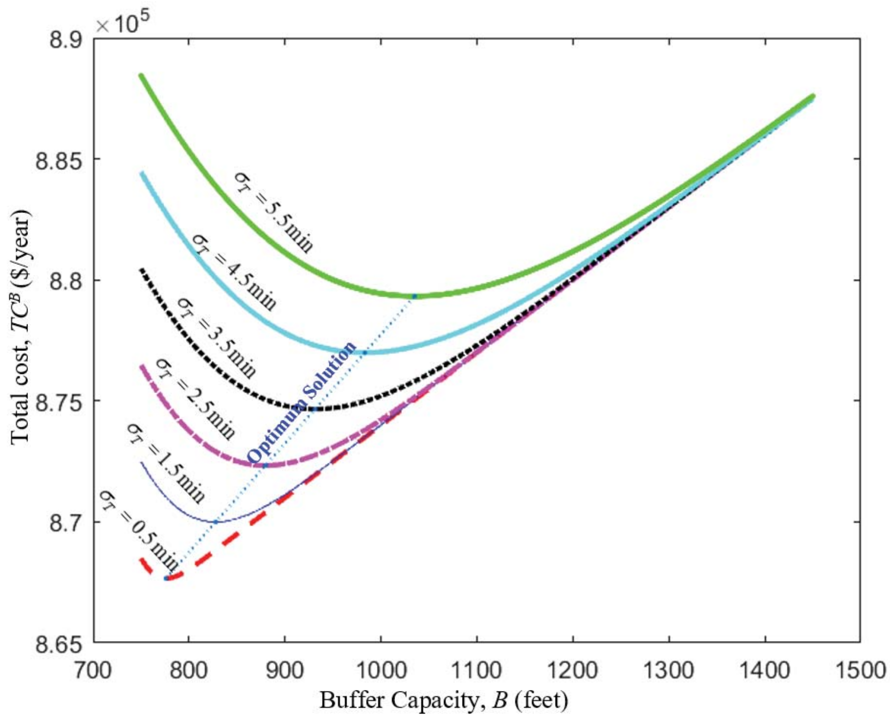


Figure 8. Sensitivity of the solution with respect to σ_T .

Sensitivity of the Solution with Respect to Acceleration

The acceleration and/or deceleration of steel strip input (α) to the coil accumulator should satisfy the condition in Theorem 1. This condition defines the minimum value of α that is needed to be ensured for achieving a desired input rate of strips to the coil accumulator. Now, for the given parameters in Table 1, it is calculated that, $\lambda^2 / \left(L_c - \bar{T}_J \lambda\right) = 1.25$, and $\lambda_{in}\left(\lambda_{in} - 2\lambda\right) / \left(\lambda \bar{T}_J\right)$ =0. Hence, from Theorem 1, the minimum value of α is found as $\alpha = \max(1.25, 0)$ =1.25.

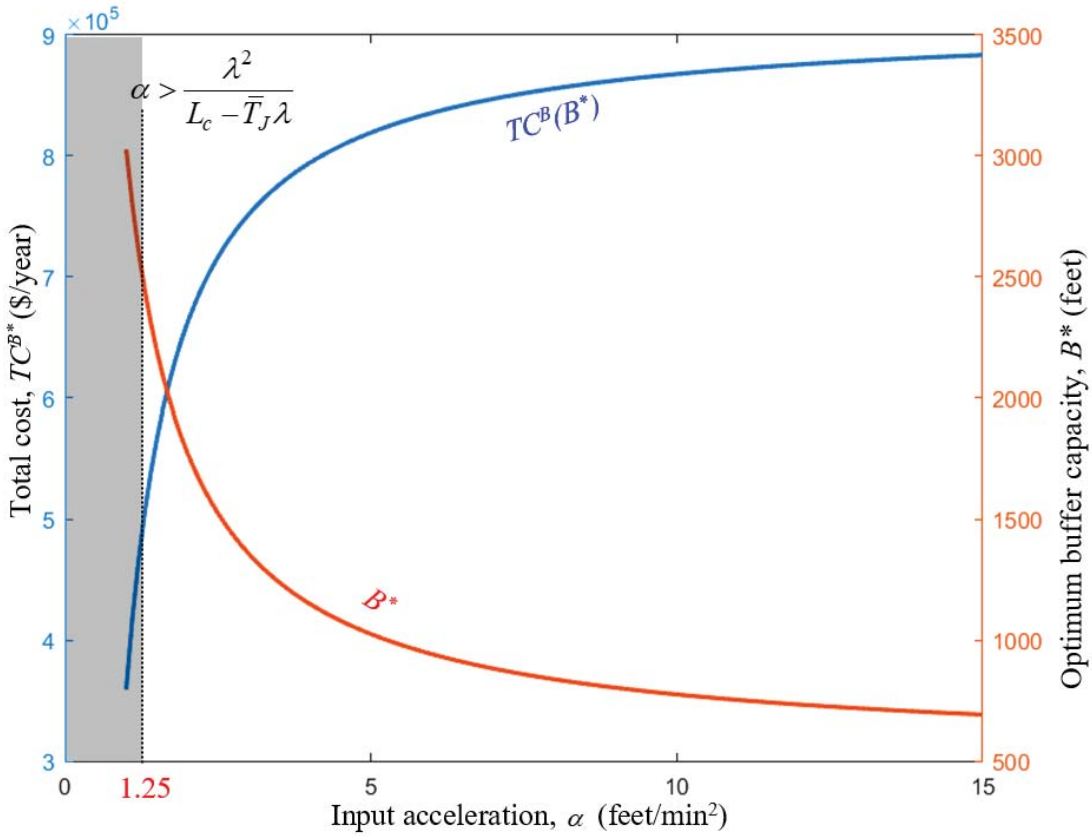
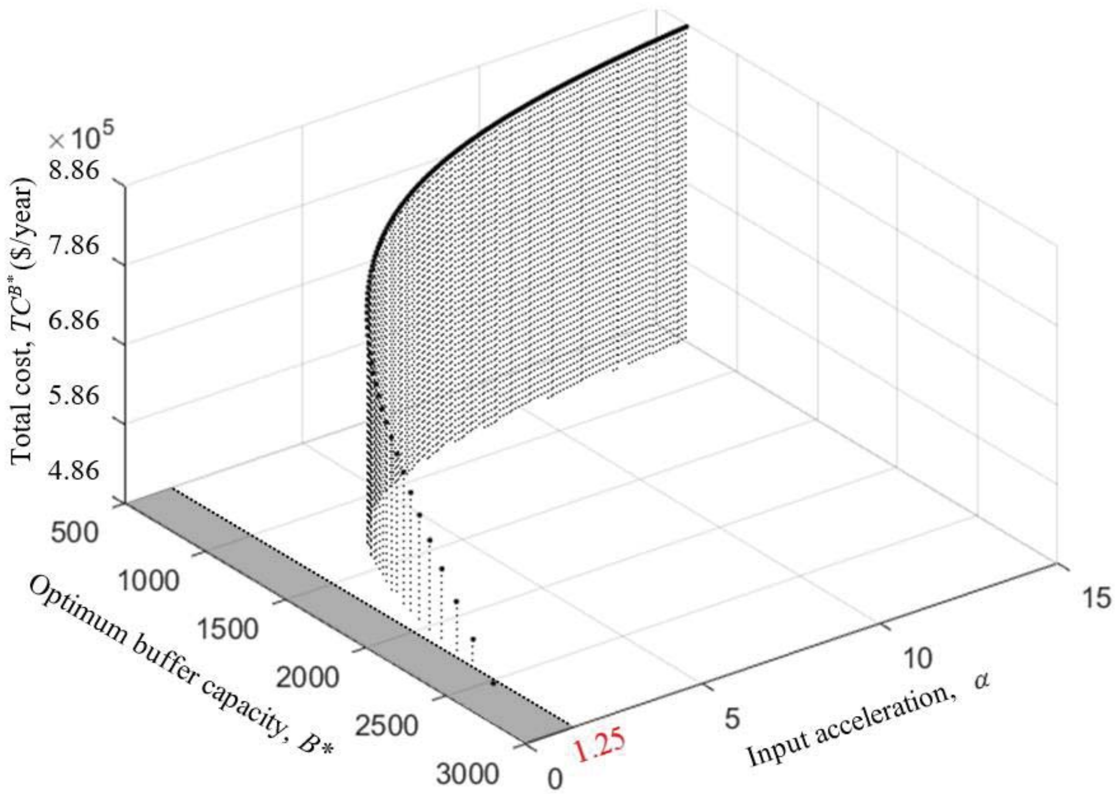


Figure 9. Effect of input acceleration α on optimum solution: 2D represntation.



(b)

Figure 10. Effect of input acceleration α on optimum solution: 3D represntation.

So, in Example 1 the value of α is varied from 1.25 to 15 feet/min^2. For several values of α the optimum solution, B^* and the corresponding total cost, $TC^B(B^*)$ are calculated. The optimum solution points are plotted in Figures 9 and 10. Both Figure 9 and 10 are representing the same results, but in different ways. In Figures 9 and 10, it is shown that, with the increase of α the optimum buffer size decreases; whereas the total cost increases asymptotically. So, in order to obtain a lower value of total cost, the input acceleration should be kept as low as possible if the corresponding B^* is attainable with the available coil accumulator.

CONCLUSION

In continuous steel manufacturing process, steel strips are fed into the line through a coil accumulator. The coil accumulator acts as a buffer storage that helps avoid disruption of the continuous process during joining a new coil with the previous strip in the line. This research problem concerns about the capacity of buffer storage in the coil accumulator. The problem is formulated as a constrained non-linear programming problem, for which an optimum solution is obtained in closed form.

The optimum solution is dominated by the constraint at some pre-determined low value of the buffer runout probability. On the other hand, the optimum solution can be calculated by a closed form expression, if the buffer runout probability is relaxed or set at a higher value. The optimum solution for buffer capacity is not affected by input rate of the strip to the coil accumulator, the length of the steel strips in a coil, and the unit costs for the input velocity, acceleration and coil joining. It is observed that, optimum solution for the total cost increases linearly with joining time variability. The optimum total cost increases with the mean joining time as well. Some other observations of the solution conclude that the buffer runout probability decreases with increase of buffer capacity. Acceleration of strips input should be kept low for the resulting lower total cost.

In a more realistic situation, the company faces a constraint for capital expenditure, and only a few sizes of the coil accumulator are available in the market. So, an extension of this research can be done while considering the budget constraint and discrete size of coil accumulator. However, the outcome of this research develops a basis for reducing the cost for buffer storage in coil accumulator. It also helps reduce the downtime of the pipe mill, which is resulted by the line stoppage due to longer joining time.

REFERENCES

Alvarezvargas, R., Dallery, Y. and David, R. (1994) “A study of the continuous-flow model of production lines with unreliable machines and finite buffers,” *Journal of Manufacturing Systems* **13**(3): 221-234.

Anderson, D. R. and Moodie C. L. (1968) “Optimal buffer storage capacity in production,” *International Journal of Production Research* **7**(3): 233-240.

Basu, R. N. (1977) “The interstage buffer storage capacity of nonpowered assembly lines: A simple mathematical approach,” *International Journal of Production Research* **15**(4): 365-385.

Bierbooms, R., Adan, I. J. B. F. and Vuuren, M. V. (2013) “Approximate Performance Analysis of Production Lines with Continuous Material Flows and Finite Buffers,” *Stochastic Models* **29**(1): 1-30.

Buzacott, J. A. (1967a) “Automatic transfer lines with buffer stocks,” *International Journal of Production Research* **5**(3): 183-200.

Buzacott, J. A. (1967b) “Prediction of the efficiency of production systems without internal storage,” *International Journal of Production Research* **6**(3): 173-188.

Dolgui, A. B., Eremeev, A. V. and Sigaev, V. S. (2017) “Analysis of a Multicriterial Buffer Capacity Optimization Problem for a Production Line,” *Automation and Remote Control* **78**(7): 1276-1289.

Fitouhia, M. C., Nourelfath, M. and Gershwin, S. B. (2017) “Performance evaluation of a two-machine line with a finite buffer and condition-based maintenance,” *Reliability Engineering and System Safety* **166**: 61-72.

Fu, M. and Xie, X. L. (2002) “Derivative estimation for buffer capacity of continuous transfer lines subject to operation-dependent failures,” *Discrete Event Dynamic Systems-Theory and Applications* **12**(4): 447-469.

Gebennini, E., Grassi, A., Fantuzzi, C. (2015) “The two-machine one-buffer continuous time model with restart policy,” *Annals of Operations Research* **231**(1): 33-64.

Hillier, F. S. and Boling, R. W. (1966) “The effect of some design factors on the efficiency of production lines with variable operation times,” *Journal of Industrial Engineering* **17**(12): 651-658.

Hossain, M.S.J. and Sarker, B.R. (2018) “Buffer capacity of coil accumulator in the seam welded pipe manufacturing mill,” 2018 MIE Graduate Student Conference, April 14, Louisiana State University, Baton Rouge, LA 70803.

Hosseini, B. and Tan, B. (2017) “Simulation and optimization of continuous-flow production systems with a finite buffer by using mathematical programming,” *IISE Transactions* **49**(3): 255-267.

Knott, A. D. (1970) “The inefficiency of a series of work stations—A simple formula,” *International Journal of Production Research* **8**(2): 109-119.

Kolb, O. and Gottlich, S. (2015) “A continuous buffer allocation model using stochastic processes,” *European Journal of Operational Research* **242**(3): 865-874.

Ozdogru, U. and Altiok, T. (2015) “Continuous material flow systems: analysis of marine ports handling bulk materials,” *Annals of Operations Research* **231**:79-104.

Panwalkar, S. S. and Smith, M. L. (1979) “A predictive equation for average output of K stage series systems with finite interstage queues,” *AIIE Transactions* **11**(2): 136-139.

Rajaram, K. and Tian, Z. L. (2009) “Buffer location and sizing to optimize cost and quality in semi-continuous manufacturing processes: Methodology and application,” *IIE Transactions* **41**(12): 1035-1048.

Rao, N. P. (1975) “Two-stage production system with intermediate storage,” *AIIE Transactions* **7**(4): 414-421.

Rao, N. P. (1976a) “A generalization of the ‘bowl phenomenon’ in series production system,” *International Journal of Production Research* **14**(4): 437-443.

Rao, N. P. (1976b) “A viable alternative to the ‘method of stages’ solution of series production systems with Erlang service times,” *International Journal of Production Research* **14**(6): 699-702.

Sarker, B. R. (1984) “Some comparative and design aspects of series production systems,” *IIE Transactions* **16**(3) :229-239.

Seong, D., Chang, S. Y. and Hong, Y. (2000) “An algorithm for buffer allocation with linear resource constraints in a continuous-flow unreliable production line,” *Asia-Pacific Journal of Operational Research* **17**(2): 169-180.

Tan, B. and Gershwin, S. B. (2009) “Analysis of a general Markovian two-stage continuous-flow production system with a finite buffer,” *International Journal of Production Economics* **120**(2): 327-339.

Tan, B. and Gershwin, S. B. (2011) “Modelling and analysis of Markovian continuous flow systems with a finite buffer,” *Annals of Operations Research* **182**(1): 5-30.

Turki, S., Hennequin, S. and Sauer, N. (2013) “Perturbation analysis for continuous and discrete flow models: a study of the delivery time impact on the optimal buffer level,” *International Journal of Production Research* **51**(13): 4011-4044.

DECISION SCIENCES INSTITUTE

Examining the impact of contract design on supplier performance: A theoretically-informed, multi-criteria, decision analysis

Ilias Vlachos
La Rochelle Business School, Excelia-Group
Email: vlachosil@excelia-group.com

ABSTRACT

Contracts has been studied extensively, yet how contract design influences supplier performance is largely unexplored. This study reviewed 18 contract design factors according to contract, relational, and contextual theoretical views. Multi-criteria methods (DEMATEL/ANP) were used to uncover the impact of the contract design on 3 supplier groups: average-, over-, and under-performers.
Results show that in the case of over- performers, contracts have a dual, yet discrete, transaction and relational role: at transaction level, they emphasise formality, protect from opportunism and include both liquidated damages and legal action clauses. At relational level, they focus on relational learning and incentivizing suppliers.

KEYWORDS: Contract Design, Supplier performance, Multi-criteria decision making (MCDM) methods, DEMATEL, ANP

INTRODUCTION

Interfirm contracts have been extensively researched in economics (Argyres et al. 2007; Ryall and Sampson 2009), yet few empirical studies on the actual process of contract design exist in the management literature (Ariño et al. 2014; Selviaridis and van der Valk 2019). Interfirm contracts (or simply contracts hereafter) have four characteristics: (i) they are idiosyncratically incomplete and, often, are too incomplete (Williamson 1987), which exposes companies to opportunism and conflicts with trading partners (Cabral 2017); (ii) Contracts are often ambiguous, which makes its interpretation depending on contextual factors (Gilson et al. 2014); (iii) Contract design is considered as a complex process; a wide number of factors play a role such as performance standards, firm strategy, business environment, the type of service or goods contracted, and the decision-making process itself (Weber and Mayer 2011); (iv) Contracts may evolve over time which makes their design a path-dependent, dynamic process (Argyres, Bercovitz, and Mayer 2007).
To deal with inherent incompleteness, complexity, ambiguity, and dynamism, one research view, based on efficiency theories, argues that contracts are interfirm governing mechanisms that aim to manage and reduce transaction costs (Benaroch et al. 2016; Williamson 1987): according to this view, contracts should be as less incomplete as possible which also minimises the supplier opportunism, yet at the expense of complexity. However, this view overlooks the potential use of contracts as mechanisms that frame and develop interfirm relationships: another view, the relational view, argues that by building trust and incentive partners, contracts can frame an interfirm relationship which allows companies to build relationship capital and generate mutual benefits (Möhring and Finch 2015; Ariño et al. 2014). A third view argues that the context defines contracts (Gilson, Sabel, and Scott 2014; Goorha 2018):

competition, force factors beyond firm control, such as force majeure and competitive forces, dictates what to include or not in a contract.

Few studies have attempted to synthesise these views and provide empirical support in order to improve procurement and supplier performance via effective contract design (Wang 2010; Vlachos 2012). This study empirically integrates these three views to answer the following research question: How contract design can improve supplier performance? To answer this question, this study examined a global firm with hundreds of supplier and procurement contracts annually. It compared 3 groups of suppliers: under-performers, average, and over-performers to uncover differences and similarities in the contract factors between them. Improving supplier performance by theoretically informed, yet affluent and easily applicable, contracting can have significant practical implications. Furthermore, this study provides empirical evidence to help bridge the gap between theory and praxis: First, it reviews 18 contract factors under three theoretical views: efficiency, relational, and contextual factors. Then, analysis shows that contract and relational factors influence more than contextual factors the contract design with over-performing suppliers.

The remainder of the study is organised as follows: The next section reviews the literature on theoretical views on contract design. Then, the research methods section presents the research design and case study. The research findings follow and the study concludes with discussion of the findings, presentation of managerial implications, and research limitations and recommendations for further research.

LITERATURE REVIEW

Contract Design Factors

In order to uncover the factors influencing contract design, prior literature was reviewed under three research views: contract, relational, contextual. In total, 18 factors were uncovered.

Efficiency factors

According to efficiency theories such as transaction cost economics, contracts act as mechanisms that minimise costs in interfirm transactions (Benaroch, Lichtenstein, and Fink 2016). Transaction costs arise because of limited information, industry dynamics, market uncertainty, and the prospect of opportunistic behaviour stemming from incomplete contracts (Williamson 1987). To minimise costs, firms design contracts that are formal, law-binding texts that include clauses regarding liquidated damages and legal actions. However, to minimize post-contractual opportunism, firms must incur higher transaction costs by designing more complex contacts to cover all aspects and contingencies in the interfirm interactions (Wacker et al. 2016). However, increased contract complexity may reduce incompleteness, yet it might increase transaction costs, particularly for contracts with long duration, which indicates that designing contracts based on transaction factors may result in suboptimal contracts and poor supplier performance.

Formality

Contracts as governing mechanisms of formal interfirm arrangements are formal, lawful texts (Carson et al. 2006). However, formality in contracts could undermine interfirm trust and therefore inspire the same opportunistic actions they are formally designed to depress (White et al. 2018; Watanabe et al. 2017). Mellewigt et al. (2007) propose that formality and trust can substitute or complement each other depending on contractual complexity. A formal contract still can be incomplete. For example, Cavusgil et al. (2004) suggest that cultural and geographical distances can reduce the effectiveness of contracts, thus those clauses should be excluded in a formal contract. Poppo and Zenger (2002) propose that formality and relational governance can be complements when contracting parties can improve their exchange performance and use contracts as a key mechanism in the governance of the improving interfirm relations.

Contract duration

Prior literature has examined the impact of duration on contract characteristics such as complexity and completeness, yet results are inconclusive (Iyer and Sautner 2018). Longer-duration contracts can be more complex since more contingencies can emerge over a longer period of time (Schepker et al. 2014). Further, longer duration implies higher costs in negotiating and agreeing on terms including exchange hazards that induce partners to engage in longer-term contractual commitments (Weber et al. 2011). Despite the higher complexity and the derived processing and transaction costs, the economic value of long-term collaborations is typically greater than that of short-term relationships, making investments in complex contracts economically more feasible (Brown et al. 2016; Wang et al. 2016).

Post-contractual opportunism

Post-contractual opportunism is considered one of the most concerning threats in interfirm relationships (Tahawy and Ibrahim 2015). Opportunistic behaviour is hard to detect in the pre-contractual stage, thus, when parties foresee a possibility of opportunism, they mitigate it by designing interfirm contract accordingly. For example, contracts with price adjustment clauses, which decrease the gap between contract and market prices compared to fixed-price agreements, can mitigate ex-post motivations to behave opportunistically (Casas-Arce et al. 2018). Grant (2007)'s study on vulnerability to post contractual opportunism in supply chain partnerships shows that, although long time horizons can negate the propensity for opportunistic behaviour, yet, without commitment to partnership, suppliers become sceptical and lose credibility of their partners.

Studies on IT outsourcing arrangements also provide insights on post-contractual opportunism since asset transfers to vendors incentivizes them to continue to invest in the transaction-specific assets to improve service (Dhillon et al. 2017). However, following transaction cost economics (Williamson 1987), the asset transfer increases bilateral dependence and would elevate the risk of post-contractual opportunistic behaviour (Bhattacharya et al. 2014). Therefore, contracts should specify terms that mitigate post-contractual opportunistic risks and at the same time provide the required incentives to make partnership work for both parties beneficially. Chang et al. (2017) study on IT outsourcing find that asset transfer does significantly affect contract design, manifested in the inclusion of clauses that protect both clients and vendors.

Liquidated damages

Most contracts include clauses of Liquidated Damages, which refer to a fixed or determined sum agreed by the contracting parties to be payable on breach by the other party. Lee et al. (2018) compare bonus contracts with penalty contracts in a cross-country experiment and find that they affect contract effectiveness while national culture moderates these effects. Wilkinson-Ryan (2009) suggest that a contract with a liquidated damages clause is more likely to default than one without it since firms will try to comply with moral and social norms to meet their obligations. Morgan et al. (2018) examine strong contracts as a tool to control supplier opportunism in outsourcing settings and found that threats of liquated damages became an accepted tool of behavioural control across all actors in a case study of the defence sector. Wacker, Yang, and Sheu (2016) argue that performance ambiguity can open the door for both parties to sue for liquidated damages, therefore companies should complement relational governance with contractual governance. However, compliance may impede the potential success of outsourcing contracts when collaboration is required to deal with future demand uncertainty (Delbufalo and Bastl 2018).

Legal actions
Similar to liquidated damages, legal actions refer to un-liquidated damages aim to fairly compensate the buyer for their loss as a direct result of the supplier not meeting their contractual obligations (Yang et al. 2018). Legal actions follow a contractual breach and they are calculated by the court based on: (i) Loss of Expectation within reasonable *contemplation. This is usually the most financially beneficial and popular method to calculate the compensation and provides the supplier incentive to meet their contractual obligations because there is no financial gain from breach; (ii)* Reliance loss: similar to liquidated damages, it incurs the value that spent or lost as a direct result of the supplier's default. (iii) Restitution: the value of any benefit a customer may have forsaken as a direct result of the supplier's default.
Legal actions can be costly, time-consuming and damaging to trustful supplier relationships: Ring and van de Ven (1994) point out that the reliance on trust at the interpersonal level may be conditioned by legal systems or organizational role responsibilities. When trust is low, contracts enhance legal obligations of the contracting parties. Formal contracts lead to greater levels of trust, out of which, over time, and by abiding legal actions, stem informal psychological contracts between buyers and suppliers which can supersede formal legal contracts (Handfield and Bechtel 2002). Legal actions may affect not only trust but also trading terms. Ranjan and Lee (2007) examine contracts in the context of global trade and found that contract enforcement directly affects trade volume and product differentiation. Weber and Mayer (2011) highlight that contracts may induce certain behaviours; thus indirectly affecting supplier performance and categorising contracts into two groupings: (i) promotional contracts that endorse trust and take a partnership approach; and (ii) prevention contracts that dictate high conformance to avoid contract default; this type of contracts are expected to rely more on legal actions than trust.

Relational factors
A growing number of studies recognise the pivotal role of contracts in framing and developing interfirm relationships (Möhring and Finch 2015). In sharp contrast to penalty clauses and legal actions, the relational view of the contracts proposes supplier incentives and building trust to create a successful interfirm relationship (Poppo and Zhou 2014; Ariño et al. 2014; Selviaridis and van der Valk 2019). As such, firms should consider relationship-building factors during contract design such as trust, incentives, relationship learning, partnership type, and supplier re-organisation.
Trust
A significant number of studies have examined the role of trust in formal contracts, yet results are inconclusive (Poppo and Zhou 2014). Three factors appear to mediate the effect of trust on contract effectiveness: flexibility, cost, and complexity. Contracts may vary in their degree of flexibility, which is the extent to which the contract is, deliberately or unintentionally, subject to different interpretations and adaptations (Arve and Martimort 2016; Furlotti 2007). Contract flexibility can be particularly valuable when partners face a high degree of (demand or supply) uncertainty, or unanticipated contingencies which allows them to adopt without incurring expensive costs of renegotiation and revision of agreed contract terms (Schepker et al. 2014; Boeh 2011). In general, negotiating and drafting contracts can be costly especially in complex contracts; when partners trust each other they may agree on flexible contracts that reduces the transaction and relationship cost and contract complexity (Ariño et al. 2014).

Incentives
Narayanan and Raman (2004) argue that incentives in contracts are an essential part of supply chain alignment such as Selviaridis and van der Valk (2019) who find that the framing of contractual performance incentives influences supplier behavioural and relational responses. Gunasekaran et al. (2015) review studies of green supply chain

collaboration and incentives and found various types of incentives while Iossa and Martimort (2016) suggest that contracts containing few incentives are open to corruption in the context of Public–Private Partnerships. Several studies have modelled the impact on contractual incentives on supply chains i.e. Wang and Shin (2015) examined the impact of contracts and competition on upstream innovation in a supply chain.

Partnership type

Contracts as a governing mechanism reflect the type of partnership between the contracting parties. Supply chain literature has extensively researched the types of collaboration and integration among chain partners: partnerships may vary from spot markets to vertical integration and strategic alliances (Yeung et al. 2013). Partnership-specific contracts and relational governance can be both substitutes and complements depending upon contingencies posed by environmental and behavioural (Abdi and Aulakh 2017). Hagedoorn and Hesen (2007) present how different modes of inter-firm technology partnering (i.e. equity joint ventures, non-equity partnerships, and licensing contracts) align with their contractual implications. Reuer and Arino (2007) investigate the contractual features of strategic alliances and show that contractual provisions depend on asset specificity and alliance's duration type (pre-specified or open-ended). Furlotti (2007) contractual incompleteness is associated with the partnership type and the control parties agree to exert.

Supplier reorganisation

Interfirm relationships are dynamic and evolve over time, which contracts should also reflect. Coltman et al. (2009) reviewed the evolution of supply chain contracts and suggested that contract design is a process that initially produces inefficient and inequitable contracts, yet, over time, by building trust and accumulating knowledge, firms improve contracts. This process evolution may affect the contracting parties directly by forcing them to re-organise their processes and functions to meet contractual agreements, especially when small and medium enterprises deal with larger, powerful partners (Ghosal 2015). This process re-organisation and resource re-configuration, which can take the form of business process re-engineering, continuous improvement and total quality management, can be a source of competitive advantage and result in contract renewal, yet it can also backfire with increased contract complexity and ambiguity and subsequently higher transaction costs (Arndt and Pierce 2018).

Relationship learning

Relationships offer valuable learning opportunities (Myers 2018). An organization can access knowledge from a partner to learn new skills (Preston et al. 2017). For example, a supplier may possess market and consumer knowledge (forecasts and trends, consumer performances); this knowledge can be valuable to a byer- its trading and contracting partner-in order to adopt its processes, produce new products etc (Lui 2009). A contract can serve as a mechanism to acquire this knowledge by embedding relationship learning into contract clauses (De Luca and Cano Rubio 2019).
Relationship learning is not always guaranteed: partners may refuse to share knowledge or protect against knowledge spill-overs which can allow an opportunistic partner to become a competitor (Yuan et al. 2018). Relationship learning also depends on the duration of the contract, the length and history of past relations and the expectation of continued exchange in future (Lui 2009; Myers 2018). As a result, how contracts deal with relationship learning may vary significantly (Dekker and Abbeele 2010). Ryall and Sampson Ryall and Sampson (2009) review the key terms from 52 technology development contracts in telecommunication companies and found high variance on contract purposes and relational governance.

Contextual factors
The context that companies operate, including the industry, market, and competitive landscape, is unavoidable complex, unpredictable and dynamic (Goorha 2018). By reflecting these realities, contact design becomes more incomplete and complex in the need to incorporate these contextual factors which are beyond the control of the trading parties (Gilson, Sabel, and Scott 2014).

Force majeure
An event is characterised as 'Force majeure' when it and its effects cannot be anticipated nor controlled, and it can be a result of either an act of nature or people (Tanenbaum 2006). Force majeure is the most common contractual term and refers to circumstances beyond the control and without the fault or negligence of the non-performing party (Chopra and Sodhi 2014). A force majeure clause aims to protect the supplier from being reprimanded in situations where they cannot meet their contractual obligations due to an event beyond their control. Despite their popularity, Tanenbaum (2006) stresses that force majeure clauses could increase the risk of default when suppliers are not adequately prepared to mitigate disruptions and suggested combining force majeure clauses with disaster recovery and business continuity provisions to ensure the supplier anticipates force majeure events.

Change in policies
Raj et al. (2018) examine how companies can design their supply contracts to meet external pressures for more sustainable operations. Liu and Cetinkaya (2009) focus on how leadership, contract flexibility, and information asymmetry affect the designing of supply contracts in supplier vs buyer-driven channels and find that, contrary to common wisdom, leadership is not necessarily beneficial for either party under conditions of information asymmetry. Since procurement is increasingly becoming international, contracting with foreign parties can expose to information asymmetry on government policies of that foreign country (Fang et al. 2017), the case of banning Huawei devices and restricting their access to software platforms being a recent example.
Dekker et al. (2018) examine contract design in cross-border business relationships, which are characterised by significant risk due to information asymmetry and complexity, and find that host country contracts have a shorter duration, more renewal provisions, less flexibility, and relatively greater contracting costs. Esmaeili et al. (2016) investigate the short- and long-term behaviour of agents in a two-echelon CLSC and demonstrate that anticipating policy changes affect pricing decisions.

Supplier technology
Changes in supplier's technology e.g. subcontracting, changes in sub-tier supplier' technologies and standards cannot be anticipated in a contract yet they may adversely affect the supplier performance (Casas-Arce, Kittsteiner, and Martínez-Jerez 2018).
Knowledge misappropriation (Colombo and Piva 2018) and usage of non-proprietary technology (Kauremaa and Tanskanen 2016) can expose contractual partners with conflicting interests to disputes and hazards. Contracts should appropriate risks and benefits in order to mitigating risks deriving from changes in supplier technology (Yeung et al. 2013). Ryall and Sampson (2009) review the key terms from 52 technology development contracts from 36 telecommunication companies and find great variance in the contract structure depending on contract purpose and type of relational governance.

Competitive factors
Contracts have to reflect and protect the competitive position of a firm within an industry. One of the most applied models of competition, Porter (2008) five forces model determines the competitiveness of a firm against five forces: buyer power,

supplier power, the competitive threat posed by current rivals, the availability of substitutes, and the threat of new entrants (Porter 2008). One characteristic of five forces model that makes it applicable to contract design is its completeness in considering all potential forces that would threaten the competition position of a firm.

Clauses in a formal contract oblige suppliers and customers not to exploit their powers, which make contracts a mechanism to control competitive forces. A monopolistic or oligopolistic supplier holds a lot of power and may demand higher prices or have little incentive to excel if the customer cannot easily switch suppliers. On the other hand, a powerful customer may have a significant influence over a supplier's operations. Reliance on a powerful customer make impact contract characteristics such as type of partnership, incentives, trust and duration.

A previously poor performing supplier may react to the threat of rivals, new entrants or substitutions by investing in supply chain improvements in order to meet their contractual obligations. They may also become suddenly interested in developing long term customer supplier relationships in an attempt to minimise the risk of contract termination.

Figure 1 presents the conceptual framework of this study which also used to guide the network modelling using multi-criteria decision analysis.

METHODOLOGY

Research design

The research design employed a single case study with a mixed methods data collection and analysis. Case studies are appropriate to explore emergent organizational and inter-organizational phenomena to develop an in-depth understanding of their nature and complexity (Yin 1989), allowing researchers to intertwine theory and empirical data to test existing theories in the light of empirical data.

The case study was an international manufacturing company with its headquarters in France. A multi-billion-dollar, global corporation with more than 1,800 suppliers over 30 countries, the company designs and runs thousands of contracts every year. Contract design involves different departments and managers including procurement leaders, legal advisers, contract managers, and financial officers. An experienced researcher conducted informal interviews with experts followed by structured data collection.

Data collection and analysis

Initially, semi-structured interviews were conducted with key personnel. Procurement leaders provided their support and input on supplier history and relationship status. Other procurement personnel were also interviewed for their general views on supplier performance including On-Time-Delivery (OTD), which was a critical performance indicator for this company. In total, 17 employees participated. The majority of participants have 10+ years direct experience in procurement. Most of the procurement leaders had been responsible for their supplier for the entire period that was studied. The interviews had the objective to confirm the relevancy to this company of contract factors found in the literature, provide insights of all suppliers in a specific product category, and give input for the quantitative analysis.

All suppliers were classified as minor or major. Suppliers were major if their annual spend was deemed significant based on an undisclosed formula. Minor suppliers were omitted in this study. The company classed twelve suppliers as major in a specific product category from about forty in total.

After initial interviews, six suppliers were selected, and their performance was evaluated based on their monthly OTD for three consequentially years starting from 2010. Other performance metrics were the same i.e. product quality. Two of them were grouped together as over-performed suppliers and the other two as under-performed

suppliers, based on the actual data. The selection of more than one supplier reduced the single-case bias in subsequent analysis. OTD analysis was triangulated with interviews and company reports.
The next step was the collection of qualitative data. The procurement leaders of the over- and under-performed suppliers were provided inputs for the suppliers they overlook, and nine other procurement leaders and managers were asked to provide input for a typical, average supplier they deal with.

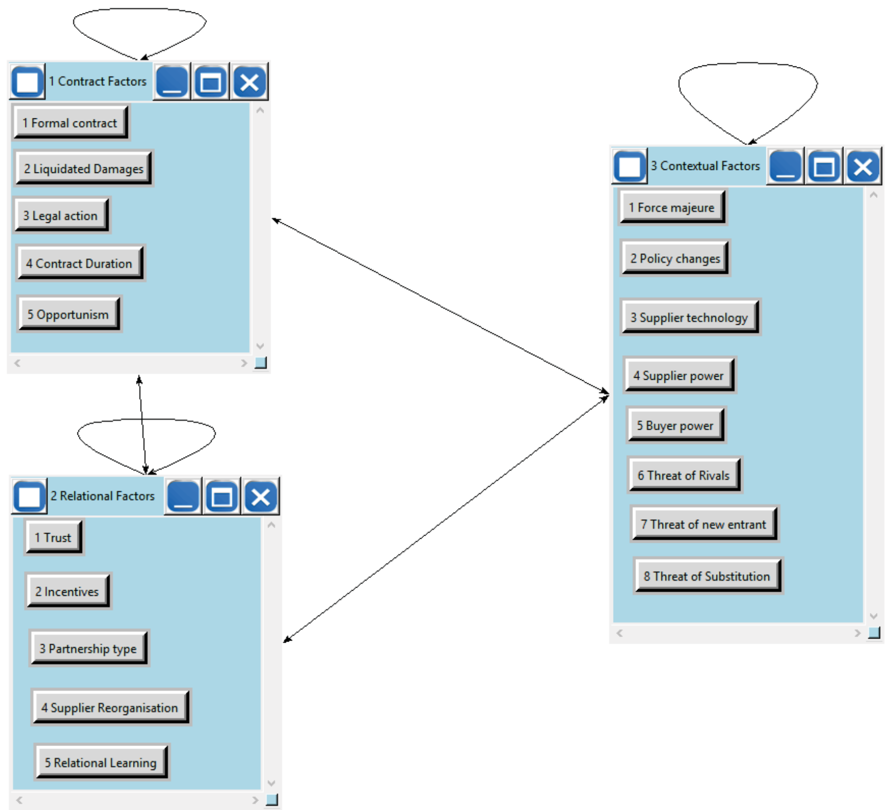


Figure 1 Conceptual Framework and Network model.

Justification of DEMATEL/ANP method
Contract design is a multi-criterion, decision-making problem (MCDM), per se, since contract clauses need to incorporate contract, relational, contextual, and completive factors. An advantage of MCDM methods is that they rely on a relatively small number of experts' opinions rather than a large-scale (Mangla et al. 2018). This advantage makes them suitable when empirical data derive from a single case study. MCDM methods include Data Envelopment Analysis (DEA), linear programming, goal programming (GP), DEMATEL, analytic hierarchy process (AHP), analytic network process (ANP), simple multi-attribute rating technique, and fuzzy set theory (Lee et al. 2011), yet no prior study has applied MCDM in contract design process.
The DEMATEL method takes into account the interrelations between attributes and divides the relevant attributes into cause and effect groups in a visual structural map. The method has been widely applied in a range of studies usually in combination with other MCDM methods, such as ANP. The ANP handles dependence within a criterion (inner dependence) and among different criteria (outer dependence). In a hybrid model

of DEMATEL and ANP, the key interdependences of variable clusters are obtained via DEMATEL, and the ANP algorithm determines the interdependences between the clusters of variables (Wu et al. 2017).

DEMATEL analysis
The following steps were applied for each of the three supplier groups.
DEMATEL Step 1: Generating the direct-relation matrix.
The comparison scale among the criteria has four levels: 0 (no influence), 1 (low influence), 2 (high influence), and 3 (very high influence). Experts are given pairs of factors and make wise comparisons in terms of influence and direction between criteria. The expert evaluations are the initial data obtained as the direct-relation matrix that is a n×n matrix A, in which a_{ij} is denoted as the degree to which the criterion i affects the criterion j (equation 1).

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad (1)$$

DEMATEL Step 2: Normalizing the direct-relation matrix.
The normalisation of the direct-relation *matrix* A produces the normalized direct-relation matrix X obtained through formulas (2), (3) and (4).

$$X = A/k \quad (2)$$

$$k = Max\left[\max_{1\le i\le n}\sum_{i=1}^{n} a_{ij}\,,\ \max_{1\le j\le n}\sum_{j=1}^{n} a_{ij}\right] \quad (3)$$

$$X = \begin{bmatrix} a_{11}/k & \cdots & a_{1j}/k & \cdots & a_{1n}/k \\ \vdots & & \vdots & & \vdots \\ a_{i1}/k & & a_{ij}/k & & a_{in}/k \\ \vdots & & \vdots & & \vdots \\ a_{n1}/k & & a_{nj}/k & & a_{nn}/k \end{bmatrix} \quad (4)$$

DEMATEL Step 3: Compute the total-relation matrix.
Having calculated the normalized direct-relation matrix X, the total relation matrix T can be acquired by using formula (5), in which I denotes the identity matrix (6).

$$T = X + X^2 + X^3 + \cdots + X^p = X(I - X)^{-1}, p \to \infty \quad (5)$$

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (6)$$

The totals for each row and each column in formula (4) can be obtained as follows:

$$r_i = \sum_{i=1}^{n} t_{ij}\,, i = 1,2, \dots, n \quad (6)$$

$$c_j = \sum_{j=1}^{n} t_{ij}\,, j = 1,2, \dots, n \quad (7)$$

where r_i represents the direct influence value which is given by the factor a_i; c_j represents the indirect influence value which is given by the factor a_j. Vector D and vector R, respectively denote the sum of rows and the sum of columns from total- relation matrix **T** respectively.
DEMATEL Step 4: Set a threshold value and obtain the impact-relation map.
The total relation matrix contains the values of impact between the factors. However, the structural relations in the matrix should not take into account unsuitable effects

between the factors. Based on the matrix ***T***, each aspect *tij* of matrix ***T*** provides information about how aspect *i* influences aspect *j*. If all the information from matrix ***T*** converts to the network relation map (NRM) the map will be too complex to show the necessary information for decision-making. A threshold value (P) is necessary to remove those effects from consideration in matrix T. Only those aspects, whose influence level in matrix T is higher than the threshold value, can be chosen and converted into the impact-digraph-map. Typically, experts discuss how to decide each factor's threshold to make the rational decisions.

In this study, the frequency of t_{ij} was decided by the experts, yet the T value was also decided to cut off less than 30% of values. To do so, the frequencies of t_{ij} were calculated and the T threshold was found. If the threshold value is too low, the map will be too complex to reveal the necessary information for decision-making. If the threshold value is too high, many aspects will be presented as independent aspects without revealing the relationships with other aspects. Therefore, a number of trial-and-error attempts were pursued to justify the correct T value. Each time the threshold value increases, some aspects or relationships will be removed from the map. After the threshold value and relative impact-digraph-map are decided upon, the final influence result can be illustrated.

The sum of rows and the sum of columns are respectively denoted as vector D and vector R through formulas (6)– (7). The horizontal axis vector (D+R) named as ''*Prominence*'' is calculated as the sum of D and R. Prominence signifies the importance of each factor. The vertical axis (D-R) named as "*Relation*" is calculated by subtracting D from R. *Relation* classifies factors into two groups: a cause group and an effect group. When a factor has positive in the vertical D-R axis, then it belongs to the cause group, otherwise it belongs to the effect group. Therefore, the causal diagram can be acquired by mapping the dataset of the (P=D + R, L=D - R) as a dispersion graph, providing valuable insight for decision making.

Analytic Network Process (ANP)

The purpose of the Analytic Network Process (ANP) approach is to solve problems involving interdependence and feedback among criteria or alternative solutions. ANP is the general form of the AHP, which has been used in multi-criteria decision-making (MCDM) in order to consider non-hierarchical structures.

The ANP handles dependence within a criterion (inner dependence) and among different criteria (outer dependence). DEMATEL method is not going to be used only to calculate the level of impacts among different groups of factors, but the normalised total-influence matrix will be incorporated into un-weighted supermatrix W in the ANP to calculate the level of interdependences of different factors.

Although key interdependences of clusters can be obtained via DEMATEL, the ANP algorithm determines interdependences between clusters. The total-influence matrix acquired by DEMATEL is similar to the concept of ANP, which confirms the importance and influence of criteria through questionnaires. The ANP algorithm runs in four steps and it was repeated for each supplier group:

ANP Step 1: Construct the structure of the network and establish its objectives. Then, the network is decomposed into network hierarchical structure.

<u>ANP Step 2: Calculate the unweighted supermatrix W.</u>

Since DEMATEL produced the total-influence matrix, the unweighted supermatrix W can be calculated by normalizing the sum of influence for each criterion in each hierarchy under the criteria of total-influence matrix. To normalize the total-influence matrix produced by DEMATEL, the criteria total-influence matrix T_c (8) yields T_c^a as shown in equation (9) where T_c^{a11} is obtained by equations (10) and (11). T_c^{aij} to T_c^{ann} are calculated with the same equations (10,11). In the normalized criteria total-influence matrix, the interdependence relationship among clustering is incorporated into the un-weighted supermatrix W as shown in equation (12). The equation (9) shows the calculation of W_{11} and the calculation of element W_{ij} to W_{nn} are based on the same way.

$$T_c = \begin{array}{cc} & \begin{array}{cccc} D_1 & D_2 & \cdots & D_n \\ c_{1_1}\ \cdots\ c_{1_{m1}} & c_{2_{m2}}\ \cdots\ c_{2_{m2}} & \cdots & c_{n_1}\ \cdots\ c_{n_{mn}} \end{array} \\ \begin{array}{cc} D_1 & \begin{array}{c} c_{1_1} \\ c_{1_{21}} \\ \vdots \\ c_{1_{m1}} \end{array} \\ \begin{array}{c} D_2 \\ \vdots \\ \vdots \end{array} & \begin{array}{c} c_{2_1} \\ c_{2_2} \\ \vdots \\ c_{2_{m2}} \\ \vdots \end{array} \\ D_n & \begin{array}{c} c_{n1} \\ c_{n2} \\ \vdots \\ c_{nm_n} \end{array} \end{array} & \begin{bmatrix} T_c^{11} & T_c^{12} & \cdots & T_c^{1n} \\ T_c^{21} & T_c^{22} & \cdots & T_c^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_c^{n1} & T_c^{n2} & \cdots & T_c^{nn} \end{bmatrix} \end{array} \tag{8}$$

$$T_c^{\alpha} = \begin{array}{cc} & \begin{array}{cccc} D_1 & D_2 & \cdots & D_n \\ c_{1_1}\ \cdots\ c_{1_{m1}} & c_{2_{m2}}\ \cdots\ c_{2_{m2}} & \cdots & c_{n_1}\ \cdots\ c_{n_{mn}} \end{array} \\ \begin{array}{cc} D_1 & \begin{array}{c} c_{1_1} \\ c_{1_{21}} \\ \vdots \\ c_{1_{m1}} \end{array} \\ \begin{array}{c} D_2 \\ \vdots \\ \vdots \end{array} & \begin{array}{c} c_{2_1} \\ c_{2_2} \\ \vdots \\ c_{2_{m2}} \\ \vdots \end{array} \\ D_n & \begin{array}{c} c_{n1} \\ c_{n2} \\ \vdots \\ c_{nm_n} \end{array} \end{array} & \begin{bmatrix} T_c^{a11} & T_c^{a12} & \cdots & T_c^{a1n} \\ T_c^{a21} & T_c^{a22} & \cdots & T_c^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_c^{an1} & T_c^{an2} & \cdots & T_c^{ann} \end{bmatrix} \end{array} \tag{9}$$

$$d_j = \sum_{j=1}^{n} t^{ij}\ , j = 1,2,\ldots,n \tag{10}$$

$$T_c^{a11} = \begin{bmatrix} t_{c^{11}}^{11}/d_1^{11} & \cdots & t_{c^{1j}}^{11}/d_1^{11} & \cdots & t_{c^{1n}}^{11}/d_1^{11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{i1}}^{11}/d_2^{11} & & t_{c^{ij}}^{11}/d_2^{11} & & t_{c^{in}}^{11}/d_2^{11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{n1}}^{11}/d_n^{11} & & t_{c^{nj}}^{11}/d_n^{11} & & t_{c^{nn}}^{11}/d_n^{11} \end{bmatrix} = \begin{bmatrix} t_{c^{11}}^{a11} & \cdots & t_{c^{1j}}^{a11} & \cdots & t_{c^{1n}}^{a11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{i1}}^{a11} & & t_{c^{ij}}^{a11} & & t_{c^{in}}^{a11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{n1}}^{a11} & & t_{c^{nj}}^{a11} & & t_{c^{nn}}^{a11} \end{bmatrix} \tag{11}$$

$$W = \begin{array}{cc} & \begin{array}{ccccccccc} & D_1 & & & D_2 & \cdots & D_n & & \\ c_{1_1} & \cdots & c_{1_{m1}} & c_{2_{m2}} & \cdots & c_{2_{m2}} & \cdots & c_{n_1} \cdots & c_{n_{mn}} \end{array} \\ \begin{array}{cc} & c_{1_1} \\ D_1 & c_{1_{21}} \\ & \vdots \\ & c_{1_{m1}} \\ D_2 & c_{2_1} \\ \vdots & c_{2_2} \\ & \vdots \\ \vdots & c_{2_{m2}} \\ & \vdots \\ & c_{n1} \\ D_n & c_{n2} \\ & \vdots \\ & c_{nm_n} \end{array} & \begin{bmatrix} W^{11} & W^{12} & \cdots & W^{1n} \\ W^{21} & W^{22} & \cdots & W^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W^{n1} & W^{n2} & \cdots & W^{nn} \end{bmatrix} \end{array} \tag{12}$$

$$W_{11} = \begin{array}{cc} & \begin{array}{cccc} c_{11} & c_{12} & \cdots & c_{1m_1} \end{array} \\ \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \end{array} & \begin{bmatrix} t_{c^{11}}^{a11} & \cdots & t_{c^{21}}^{a11} & \cdots & t_{c^{m1}}^{a11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{12}}^{a11} & & t_{c^{22}}^{a11} & & t_{c^{m2}}^{a11} \\ \vdots & & \vdots & & \vdots \\ t_{c^{1m_1}}^{a11} & & t_{c^{2m_1}}^{a11} & & t_{c^{mm_1}}^{a11} \end{bmatrix} \end{array} \tag{13}$$

<u>ANP Step 3: Obtain the weighted supermatrix</u>

by normalizing the sum of impact for each hierarchy and each dimension in the dimensions total-influence matrix as illustrated in equation (14). Normalizing the total influence matrix T_D yields T_D^a (15). The weighted supermatrix is obtained by incorporating the unweighted supermatrix into the normalized dimensions total-influence matrix (16).

$$T_D = \begin{bmatrix} t_D^{11} & \cdots & t_D^{1j} & \cdots & t_D^{1n} \\ \vdots & & \vdots & & \vdots \\ t_D^{i1} & & t_D^{ij} & & t_D^{in} \\ \vdots & & \vdots & & \vdots \\ t_D^{n1} & & t_D^{nj} & & t_D^{nn} \end{bmatrix} \tag{14}$$

$$T_D^a = \begin{bmatrix} t_D^{11}/d_1 & \cdots & t_D^{1j}/d_1 & \cdots & t_D^{1n}/d_1 \\ \vdots & & \vdots & & \vdots \\ t_D^{i1}/d_2 & & t_D^{ij}/d_2 & & t_D^{in}/d_2 \\ \vdots & & \vdots & & \vdots \\ t_D^{n1}/d_n & & t_D^{nj}/d_n & & t_D^{nn}/d_n \end{bmatrix} = \begin{bmatrix} t_D^{a11} & \cdots & t_D^{a1j} & \cdots & t_D^{a1n} \\ \vdots & & \vdots & & \vdots \\ t_D^{ai1} & & t_D^{aij} & & t_D^{ain} \\ \vdots & & \vdots & & \vdots \\ t_D^{an1} & & t_D^{anj} & & t_D^{ann} \end{bmatrix} \tag{15}$$

$$W = \begin{bmatrix} t_D^{a11} \times W_{11} & t_D^{a21} \times W_{12} & \cdots & \cdots & t_D^{an1} \times W_{1n} \\ & & & & \vdots \\ t_D^{a12} \times W_{21} & t_D^{a22} \times W_{22} & & & \vdots \\ \vdots & \vdots & & & \\ \vdots & \cdots & t_D^{aij} \times W_{ij} & \cdots & t_D^{ani} \times W_{ni} \\ \vdots & \vdots & & & \vdots \\ t_D^{a1n} \times W_{n1} & t_D^{a2n} \times W_{n2} & & & t_D^{ann} \times W_{nn} \end{bmatrix} \quad (16)$$

ANP Step 4: Obtain the limited supermatrix,
by multiple productions of the weighted supermatrix until the vector values in the limited supermatrix become stable (equation 17, with W being the limited supermatrix and z tending to infinity). The vectors of the limited supermatrix represent the relative weights of each factor in relation to the defined objective. Sorting the limited supermatrix W according to the relative weights of each factor gives insights on the significance and contribution of each factor as well as each cluster to the objective of network.

$$\lim_{z \to \infty} = W_w^z \quad (17)$$

FINDINGS

Contract Factor Inter-relationships

Table 1 presents the sums of influences of contract factors by supplier group on DEMATEL analysis. Figure 2 presents the Relation-Prominence Map for average-, over- and under- performing suppliers.
Prominence (p, D+R in Table 1) illustrates the strength of influences that are given and received of a given factor indicating the degree of central role that the factor plays in the system. For over-performers, the highest prominence scores were: Formality (p_o=8.54), Contract Duration (p_o=8.03), Partnership type (p_o=7.96) and Relationship Learning (p_o=7.74), while the lowest scores were: Force majeure (p_o=8.03), Supplier Reorganisation (p_o=8.03), and Supplier technology (p_o=8.03). On the other hand, for under-performers, the highest prominence scores were: Formality (p_u=5.39), Contract Duration (p_u=5.18), Trust (p_u=4.96), and Partnership type (p_u=4.84) while the lowest were: R Supplier Reorganisation (p_u=0.45), T Force majeure (p_u=0.57), T Supplier technology (p_u=1.53), and C Opportunism (p_u=1.75). Firms put more emphasis on Relationship Learning when contracting over-performers and on trust for under-performers respectively.
Relation (r, D-R in Table 1) shows the net effect that a given factor contributes to the system; a factor with positive relation scores has a net influence on the other factors and it should be grouped into *cause* group while negative effect shows that other factors have a stronger effect of the factor causing it to be grouped into the *effect* group. The highest cause factors for over-performers were: Opportunism (r_o=1.73), Supplier power (r_o= 0.91), Policy changes (r_o= 0.85), and Buyer power (r_o= 0.81) while for under-performers were: Supplier power (r_u=1.07), Opportunism (r_u=0.98), New entrant (r_u=0.80), Buyer power (r_u=0.62), which indicates that although both groups had similar cause factors, under-performers focused on contextual factors including the risk of new entrants, while for over-performers, post-contractual Opportunism had the highest effect on other factors. Another striking difference was in the effect groups: for over-performers Liquidated Damages (r_o= -1.29) was derived from other contract factors while for under-performers, Relational Learning (r_u=-0.79) was a low-score effect factor. Therefore, Relational Learning was a high-score Prominence factor for companies contracting over performers but a low-score Relation factor for under-performers, who also relied more on trust than over-performers.

Vlachos **Contract Design impact on Supplier Performance**

Table 1 Sum of influences of contract factors by supplier groups

		Average-Performance				Over-Performers				Under-Performers			
Factors	**Group**	**D**	**R**	**D+R**	**D-R**	**D**	**R**	**D+R**	**D-R**	**D**	**R**	**D+R**	**D-R**
Formality	Contract	4.26	4.34	8.60	-0.08	4.11	4.44	8.54	-0.33	2.56	2.83	5.39	-0.27
Liquidated Damages	Contract	2.27	3.62	5.88	-1.35	3.44	4.30	7.74	-0.86	1.80	2.24	4.04	-0.43
Legal Action	Contract	2.41	3.19	5.60	-0.78	2.40	3.69	6.09	-1.29	1.86	2.47	4.33	-0.61
Contract Duration	Contract	3.13	4.49	7.63	-1.36	2.90	3.38	6.28	-0.47	2.18	3.00	5.18	-0.82
Post-contractual opportunism	Contract	2.15	0.98	3.12	1.17	3.43	4.13	7.55	-0.70	1.37	0.38	1.75	0.98
Trust	Relational	3.36	4.36	7.72	-1.00	2.08	2.98	5.06	-0.90	2.24	2.72	4.96	-0.48
Incentives	Relational	2.20	3.01	5.21	-0.81	3.54	4.48	8.03	-0.94	1.09	1.55	2.64	-0.45
Partnership Type	Relational	4.00	4.50	8.51	-0.50	3.61	4.36	7.96	-0.75	2.09	2.76	4.84	-0.67
Supplier Reorganisation	Relational	1.44	0.78	2.22	0.65	2.61	0.88	3.49	1.73	0.10	0.35	0.45	-0.25
Relationship Learning	Relational	3.59	4.40	7.99	-0.81	1.30	0.53	1.83	0.77	1.63	2.42	4.04	-0.79
Force Majeure	Contextual	-	-	-	-	0.67	-	0.67	0.67	0.48	0.09	0.57	0.38
Policy Changes	Contextual	2.47	0.97	3.44	1.50	2.11	1.26	3.37	0.85	1.08	0.88	1.97	0.20
Supplier Technology	Contextual	0.99	1.40	2.39	-0.41	0.86	1.27	2.13	-0.41	0.63	0.90	1.53	-0.26
Supplier Power	Contextual	4.23	3.21	7.44	1.02	3.95	3.04	6.99	0.91	2.68	1.61	4.29	1.07
Buyer Power	Contextual	4.51	3.58	8.09	0.93	4.22	3.41	7.64	0.81	2.61	1.99	4.60	0.62
Threat of Rivals	Contextual	2.59	1.40	3.99	1.18	2.41	2.05	4.46	0.36	1.51	1.03	2.54	0.48
Threat of New Entrant	Contextual	2.71	1.38	4.09	1.34	2.62	1.99	4.61	0.63	1.81	1.01	2.82	0.80
Threat of Substitution	Contextual	2.22	2.92	5.14	-0.70	2.52	2.60	5.12	-0.07	1.77	1.26	3.03	0.51

D+R are denoted as p in text; D-R are denoted as l in text.

D+R ***Prominence,*** D-R ***Relation***

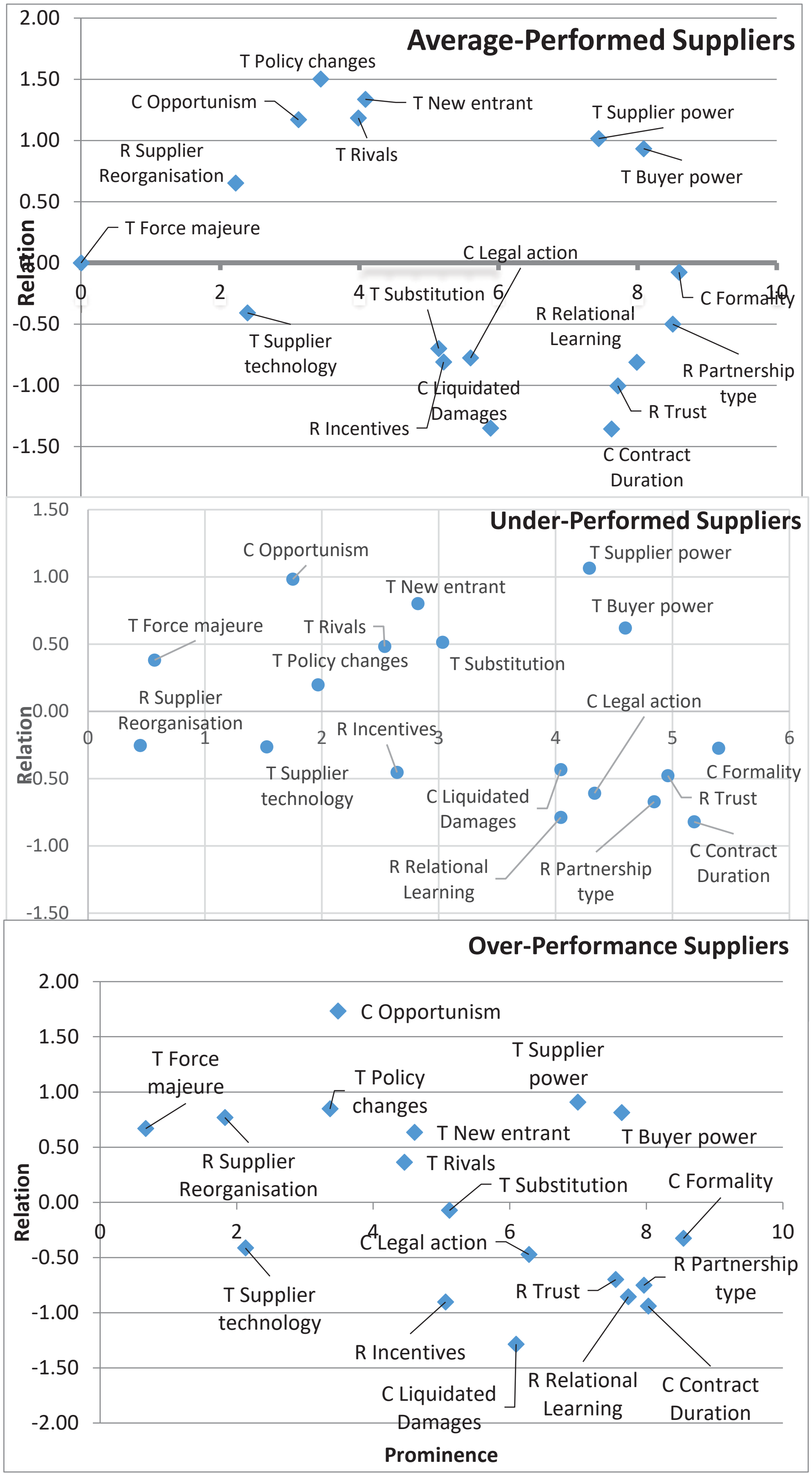


Figure 2 Relation-Prominence Map for 3 groups of suppliers

Contract Factor Rankings
Table 2 presents the rankings of contract design factors by supplier group derived from ANP analysis. Average-performing suppliers and over-performing suppliers has similar rankings while there were differences between over- and under-performing suppliers. Specifically, according to the limiting (l) values, the most important factors were for over performers: Formality (l_o= 0.077), Partnership type (l_o=0. 074286), and Relational Learning (l_o=07411) while for under-performers were: Trust (l_u=07411), Partnership type (l_u =07411), and Formality (l_u =07411). Therefore, trusting their suppliers produced less than expected performance. Contracts with under-performing suppliers showed also lower scores and emphasis on Incentives, and Relational Learning, and higher scores in most competitive factors (rivals, substitution, buyer and supplier power), which indicates that focusing on the context and not on the relationship itself results in poor performing suppliers.
Contracts that produced exceptional performance had a dual role: they put more emphasis on contractual and relational factors than contextual ones. This supports the theoretical views that contacts are governing mechanisms that reduce transaction costs but at the same time they can build and frame interfirm relationships. Specifically, the top contract factors with over-performing suppliers were contact and relational factors and the factors with les influence were contextual (Table 2). On the contrary, contracts with under-performing suppliers ranked contextual factors higher than over-performing and average-performing suppliers: i.e. Buyer power 9th vs 10th , Supplier power 8th vs 11th, rivals 14th vs 15th and Substitution 12th vs 14th respectively.
Interviews with key personnel confirmed the results of DEMATEL/ANP analysis. Procurement leaders emphasised the ineffectiveness of contractual tools such as liquidated damages, ranging from them being difficult to adequately calculate and cover damages sufficiently to discourage default, to the potential damage implementation can have on the customer- supplier relationship. Procurement managers reported that suppliers may prefer to pay liquidated damages rather than fixing the root cause of problems until it reaches a critical level. It was also highlighted that incentives were effective but not consistent especially when procurement personnel feel it unfair to reward the supplier in addition to payment for meeting their contractual obligations. Building interfirm relationships was desirable, yet beyond a point it could be less feasible or costly to tie-up in a specific supplier in the long run.

DISCUSSION
There is a rich literature on contracts (Liu and Cetinkaya 2009) but relatively few studies attempt to understand the contract design process and synthesise contract factors from different theoretical perspectives (Lumineau et al. 2011; Williamson 1987). Results show that even within the same company and same department, there are differences among the factors that influence contract design and these differences are related to supplier performance. Findings indicate that despite many commonalities, the contracting with under- and over- performers presents significant differences: In the case of over- performers, contracts have a dual, yet discrete, transaction and relational role: at transaction level, they reduce cost and complexity, and ambiguity by emphasising formality, protecting from opportunism and include both liquidated damages and legal action clauses. At relational level, they focus on relational learning and incentivizing suppliers. However, in the case of under-performers, contracts appear to focus on contextual factors, which can be a source of ambiguity particularly in complex environments. Furthermore, under-performing contracts emphasise trust, in contract to relationship learning, which appears to have a negative impact on supplier performance. These results have important theoretical and managerial implications.

Table 2 ANP Analysis: Rankings of contract design factors by supplier group

	Over-Performed Suppliers			**Under-Performed Suppliers**			**Average-Performed Suppliers**		
Name	**Normalized**	**Limiting**	**rank**	**Normalized**	**Limiting**	**rank**	**Normalized**	**Limiting**	**rank**
C Contract Duration	0.20848	0.069493	5	0.21568	0.071894	5	0.21015	0.070051	5
C Formality	0.23166	0.077221	1	0.23857	0.079522	3	0.23319	0.07773	1
C Legal action	0.19774	0.065912	6	0.20243	0.067477	6	0.19524	0.06508	6
C Liquidated Damages	0.18456	0.06152	7	0.18486	0.06162	7	0.18408	0.061361	7
C Opportunism	0.17756	0.059187	9	0.15846	0.052821	11	0.17734	0.059112	9
R Incentives	0.17979	0.05993	8	0.17097	0.056989	10	0.18133	0.060443	8
R Partnership type	0.22286	0.074286	2	0.24609	0.08203	2	0.22204	0.074014	2
R Relational Learning	0.22233	0.07411	3	0.22466	0.074885	4	0.21888	0.07296	3
R Supplier Reorganisation	0.15391	0.051304	12	0.10028	0.033428	15	0.16156	0.053853	11
R Trust	0.22111	0.073704	4	0.258	0.086001	1	0.21619	0.072062	4
T Buyer power	0.16558	0.055193	10	0.17535	0.058451	9	0.16429	0.054762	10
T Force majeure	0.08953	0.029844	18	0.07331	0.024437	18	0.08854	0.029514	18
T Policy changes	0.12012	0.040041	16	0.1	0.033335	16	0.12364	0.041212	14
T Supplier power	0.15481	0.051604	11	0.18141	0.06047	8	0.15391	0.051302	12
T Supplier technology	0.09254	0.030845	17	0.07768	0.025894	17	0.09956	0.033186	17
T Threat of new entrant	0.13016	0.043386	13	0.13367	0.044556	13	0.1264	0.042132	13
T Threat of Rivals	0.12104	0.040348	15	0.12195	0.04065	14	0.12208	0.040693	15
T Threat of Substitution	0.12622	0.042072	14	0.13662	0.045541	12	0.1216	0.040532	16

Notes: Normalized: Normalized By Cluster, rank in supplier group (over-, under-, average-performance), C: Contract, R: Relational, T: ConTextual Factors

Theoretical Implications

Due to inherent contract complexity and incompleteness (Kalkanci et al. 2014), findings support that organisational learning improves contract design by overcoming contract inefficiencies due to bounded rationality, complex technical clauses, cultural distances and unforeseen contingencies (Ariño et al. 2014; Lui 2009).

Following transaction cost economics, asset-specific investments reduce transaction costs (Williamson 1987), therefore, formality and clauses to avoid post-contractual opportunism would reduce costs. Both factors were found to relate to overperformers in this study. Trust can be also considered as an asset- specific investment (Chiou and Droge 2006). However, one fundamental principle of relationship marketing, which is that trust improves interfirm relationships and can create loyalty (Ravald and Grönroos 1996; Ariño et al. 2014), is not confirmed. Findings do not support that trust improves supplier performance, rather the opposite: trusting suppliers relates to supplier under-performance. Previous studies have warned of the dark side of interfirm relationships (Oliveira and Lumineau 2019) including the fact that trusting suppliers may backfire by motivating them to post-contractual opportunism or make them assume that poor performance will be tolerated (Lumineau 2017).

On the other hand, incentives and relationship learning did relate to over-performance, which can be explained by the fact that contracts evolve over time (Argyres and Mayer 2007) by reflecting the relationship dynamics in which business partners are engaged and the business environment they interfirm relationships are embedded. To this respect, learning reflects the contract dynamism and frames the contract respectively, in the way formality reduces complexity and ambiguity.

Managerial Implications

Contract design is a complex process that involves managers and experts from different departments and backgrounds. Contract design is a process that it is affected by many factors that are related to each other. For example, there are trade-offs between trust and formality, trust and opportunism, contextual factors and complexity and so on. One way to deal with this inherent contract complexity and incompleteness, it to understand the interrelationships between the contract factors that affect the supplier performance. By applying multi-criteria decision making tools, like DEMATEL/ANP demonstrated in this study, companies can reveal the underlying relationships and assumptions when they design contracts and then link them to supplier performance. In this study, one key performance metrics was used, but companies may evaluate suppliers across different sets of key performance indicators that better reflect their strategic priorities.

This method allows to: (i) synthesise expert opinions from different backgrounds and expertise (ii) remove bounded rationality often present in human decision making (iii) uncover cause and group effect factors as well as prominence factors that lead the contract design (iv) rank contract factors in order of significance and (v) associate contract design factors to supplier performance.

This study also shows that companies should also keep regular face-to-face supplier management workshops on both management and working levels to align goals and expectations and to develop a good working relationship. In this way, companies will be able to crystalize what type of expectations they have from each supplier and adopt their contract accordingly.

Findings show that contracts have a dual role: reduce transaction costs and foster relationship learning. Managers should to educate staff and develop policies and appropriate techniques that enhance relational learning. Contextual factors appear to make contracts more complex and ambiguous.

Limitations and Recommendations

This study employed a single case study, which can present methodological limitations in terms of generalisation of results and reliability of data collected. Multi-criteria decision-making methods rely on evidence provided by experts in their field. When experts come from the same company, this may create a common method bias. Information from external experts would reduce the common bias effect. Future studies should apply the research design to more companies and industries, for example a multi-country, multi-industry of contract design factors and their effects on supplier performance. Although this study examined 18 contract factors, other factors may be relevant to specific industries, sectors or countries. Future studies should examine contract factors in more detail.

This study chose DEMATEL/ANP to analyse the empirical data. Multi-criteria methods are useful in codifying and analysing explicit knowledge, yet they have limitations in tacit knowledge. Future research should replicate the study with other multi-criterial methods such as fuzzy DEMATEL and compare the results across different methodologies. Although not illustrated in this case study, this method can be used as a diagnostic tool in case of disputes or service failures to uncover topics that require improvement.

REFERENCES

Abdi, M., and P. S. Aulakh. 2017. "Locus of Uncertainty and the Relationship Between Contractual and Relational Governance in Cross-Border Interfirm Relationships." *Journal of Management* 43(3):771-803.

Argyres, N., and K. Mayer, J. . 2007. "Contract Design as a Firm Capability: An Integration of Learning and Transaction Cost Perspectives." *The Academy of Management Review* 32(4):1060-77.

Argyres, N. S., J. Bercovitz, and K. J. Mayer. 2007. "Complementarity and Evolution of Contractual Provisions: An Empirical Study of IT Services Contracts." *Organization Science* 18(1):3-19.

Ariño, A., J. J. Reuer, K. J. Mayer, and J. Jané. 2014. "Contracts, Negotiation, and Learning: An Examination of Termination Provisions." *Journal of Management Studies* 51(3):379-405.

Arndt, F., and L. Pierce. 2018. "The behavioral and evolutionary roots of dynamic capabilities." *Industrial and Corporate Change* 27(2):413-24.

Arve, M., and D. Martimort. 2016. "Dynamic procurement under uncertainty: Optimal design and implications for incomplete contracts." *American Economic Review* 106(11):3238-74.

Benaroch, M., Y. Lichtenstein, and L. Fink. 2016. "Contract design choices and the balance of ex-ante and ex-post transaction costs in software development outsourcing." *MIS Quarterly* 40(1):57-82.

Bhattacharya, S., S. Hasija, and L. N. Van Wassenhove. 2014. "Designing efficient infrastructural investment and asset transfer mechanisms in humanitarian supply chains." *Production and Operations Management* 23(9):1511-21.

Boeh, K. K. 2011. "Contracting Costs and Information Asymmetry Reduction in Cross-Border M&A." *Journal of Management Studies* 48(3):568-90.

Brown, T. L., M. Potoski, and D. V. Slyke. 2016. "Managing Complex Contracts: A Theoretical Approach." *Journal of Public Administration Research and Theory* 26(2):294-308.

Cabral, S. 2017. "Reconciling Conflicting Policy Objectives in Public Contracting: The Enabling Role of Capabilities." *Journal of Management Studies* 54(6):823-53.

Carson, S. J., A. Madhok, and T. Wu. 2006. "Uncertainty, opportunism, and governance: The effects of volatility and ambiguity on formal and relational contracting." *Academy of Management Journal* 49(5):1058-77.

Casas-Arce, P., T. Kittsteiner, and F. A. Martínez-Jerez. 2018. "Contracting with Opportunistic Partners: Theory and Application to Technology Development and Innovation." *Management Science* 65(2):842-58.

Cavusgil, S. T., S. Deligonul, and C. Zhang. 2004. "Curbing foreign distributor opportunism: An examination of trust, contracts, and the legal environment in international channel relationships." *Journal of International Marketing* 12(2):7-27.

Chang, Y. B., V. Gurbaxani, and K. Ravindran. 2017. "Information technology outsourcing: Asset transfer and the role of contract." *MIS Quarterly: Management Information Systems* 41(3):959-73.

Chiou, J.-S., and C. Droge. 2006. "Service quality, trust, specific asset investment, and expertise: Direct and indirect effects in a satisfaction-loyalty framework." *Journal of the Academy of Marketing Science* 34(4):613-27.

Chopra, S., and M. S. Sodhi. 2014. "Reducing the Risk of Supply Chain Disruptions." *MIT Sloan Management Review* 55(3):73-80.

Colombo, M. G., and E. Piva. 2018. "Knowledge misappropriation risks and contractual complexity in entrepreneurial ventures' non-equity alliances." *Small Business Economics*:1-21.

Coltman, T., K. Bru, N. Perm-Ajchariyawong, T. M. Devinney, and G. R. Benito. 2009. "Supply chain contract evolution." *European Management Journal* 27(6):388-401.

De Luca, P., and M. Cano Rubio. 2019. "The curve of knowledge transfer: a theoretical model." *Business Process Management Journal.* 25(1):10-26.

Dekker, H., C. , and A. V. d. Abbeele. 2010. "Organizational Learning and Interfirm Control: The Effects of Partner Search and Prior Exchange Experiences." *Organization Science* 21(6):1233-50.

Dekker, H. C., T. Kawai, and J. Sakaguchi. 2018. "The Interfirm Contracting Value of Management Accounting Information." *Journal of Management Accounting Research* In Press(https://doi.org/10.2308/jmar-52058).

Delbufalo, E., and M. Bastl. 2018. "Multi-principal collaboration and supplier's compliance with codes-of-conduct." *The International Journal of Logistics Management* 29(4):1237-54.

Dhillon, G., R. Syed, and F. de Sá-Soares. 2017. "Information security concerns in IT outsourcing: Identifying (in) congruence between clients and vendors." *Information & Management* 54(4):452-64.

Esmaeili, M., G. Allameh, and T. Tajvidi. 2016. "Using game theory for analysing pricing models in closed-loop supply chain from short- and long-term perspectives." *International Journal of Production Research* 54(7):2152-69.

Fang, H. C., E. Memili, J. J. Chrisman, and C. Penney. 2017. "Industry and Information Asymmetry: The Case of the Employment of Non-Family Managers in Small and Medium-Sized Family Firms." *Journal of Small Business Management* 55(4):632-48.

Furlotti, M. 2007. "There is more to contracts than incompleteness: A review and assessment of empirical research on inter-firm contract design." *Journal of Management & Governance* 11(1):61-99.

Ghosal, V. 2015. "Business Strategy and Firm Reorganization: Role of Changing Environmental Standards, Sustainable Business Initiatives and Global Market Conditions." *Business Strategy and the Environment*(2):123-44.

Gilson, R. J., C. F. Sabel, and R. E. Scott. 2014. "Text and context: contract interpretation as contract design." *Cornell L. Rev.* 100:23.

Goorha, P. 2018. "Contextual Contracts: On a Context-sensitive Approach to Contract Theory." *Journal of Interdisciplinary Economics* 30(2):191-209.

Grant, S. B. 2007. "Vulnerability to Post Contractural Opportunism in Supply Chain Co-Operative Exchange." *Marketing Management*:169.

Gunasekaran, A., N. Subramanian, and S. Rahman. 2015. "Green supply chain collaboration and incentives: Current trends and future directions." *Transportation Research Part E: Logistics and Transportation Review* 74:1-10.

Hagedoorn, J., and G. Hesen. 2007. "Contract Law and the Governance of Inter-Firm Technology Partnerships - An Analysis of Different Modes of Partnering and Their Contractual Implications." *Journal of Management Studies* 44(3):342-66.

Handfield, R. B., and C. Bechtel. 2002. "The role of trust and relationship structure in improving supply chain responsiveness." *Industrial Marketing Management* 31(4):367-82.

Iossa, E., and D. Martimort. 2016. "Corruption in PPPs, incentives and contract incompleteness." *International Journal of Industrial Organization* 44:85-100.

Iyer, R., and Z. Sautner. 2018. "Contracting between Firms: Empirical Evidence." *Review of Economics and Statistics* 100(1):92-104.

Kalkanci, B., K. Y. Chen, and F. Erhun. 2014. "Complexity as a contract design factor: A human-to-human experimental study." *Production and Operations Management* 23(2):269-84.

Kauremaa, J., and K. Tanskanen. 2016. "Designing interorganizational information systems for supply chain integration: a framework." *The International Journal of Logistics Management* 27(1):71-94.

Lee, W.-S., A. Y. Huang, Y.-Y. Chang, and C.-M. Cheng. 2011. "Analysis of decision making factors for equity investment by DEMATEL and Analytic Network Process." *Expert Systems With Applications* 38:8375-83.

Lee, Y. S., D. Ribbink, and S. Eckerd. 2018. "Effectiveness of bonus and penalty incentive contracts in supply chain exchanges: Does national culture matter?" *Journal of Operations Management* 62:59-74.

Liu, X., and S. Cetinkaya. 2009. "Designing supply contracts in supplier vs buyer-driven channels: The impact of leadership, contract flexibility and information asymmetry." *IIE Transactions* 41(8):687-701.

Lui, S. S. 2009. "The Roles of Competence Trust, Formal Contract, and Time Horizon in Interorganizational Learning." *Organization Studies*(4):333.

Lumineau, F. 2017. "How contracts influence trust and distrust." *Journal of Management* 43(5):1553-77.

Lumineau, F., M. Fréchet, and D. Puthod. 2011. "An organizational learning perspective on the contracting process." *Strategic Organization* 9(1):8-32.

Mangla, S. K., S. Luthra, N. Rich, D. Kumar, N. P. Rana, and Y. K. Dwivedi. 2018. "Enablers to implement sustainable initiatives in agri-food supply chains." *International Journal of Production Economics* 203:379-93.

Möhring, M. M., and J. Finch. 2015. "Contracts, relationships and innovation in business-to-business exchanges." *Journal of Business & Industrial Marketing* 30(3/4):405-13.

Morgan, R., D. Doran, and S. J. Morgan. 2018. "Strong contracts: the relationship between power and action." *International Journal of Operations and Production Management* 38(1):272-94.

Myers, C. G. 2018. "Coactive vicarious learning: Toward a relational theory of vicarious learning in organizations." *Academy of Management Review* 43(4):610-34.

Narayanan, V. G., and A. Raman. 2004. "Aligning Incentives in Supply Chains." 82(11):94-102.

Oliveira, N., and F. Lumineau. 2019. "The dark side of interorganizational relationships: An integrative review and research agenda." *Journal of Management* 45(1):231-61.

Poppo, L., and T. Zenger. 2002. "Do Formal Contracts and Relational Governance Function as Substitutes or Complements?" *Strategic Management Journal* 23(8):707-25.

Poppo, L., and K. Z. Zhou. 2014. "Managing contracts for fairness in buyer-supplier exchanges." *Strategic Management Journal* 35(10):1508-27.

Porter, M. E. 2008. "The Five Competitive Forces That Shape Strategy." *Harvard Business Review* 86(1):78-93.
Preston, D. S., D. Q. Chen, M. Swink, and L. Meade. 2017. "Generating Supplier Benefits through Buyer-Enabled Knowledge Enrichment: A Social Capital Perspective." *Decision Sciences* 48(2):248-87.
Raj, A., I. Biswas, and S. K. Srivastava. 2018. "Designing supply contracts for the sustainable supply chain using game theory." *Journal of Cleaner Production* 185:275-84.
Ranjan, R., and J. Y. Lee. 2007. "Contract Enforcement and International Trade." *Economics and Politics* 19(2):191-218.
Ravald, A., and C. Grönroos. 1996. "The value concept and relationship marketing." *European Journal of Marketing* 30(2):19-30.
Reuer, J. J., and A. Arino. 2007. "Strategic alliance contracts: dimensions and determinants of contractual complexity." *Strategic Management Journal* 28(3):313-30.
Ring, P. S., and A. H. van de Ven. 1994. "Developmental Processes of Cooperative Interorganizational Relationships." *The Academy of Management Review* 19(1):90-118.
Ryall, M., D. , and R. Sampson, C. . 2009. "Formal Contracts in the Presence of Relational Enforcement Mechanisms: Evidence from Technology Development Projects." *Management Science* 55(6):906-25.
Schepker, D. J., W. Y. Oh, L. Poppo, and A. Martynov. 2014. "The Many Futures of Contracts: Moving Beyond Structure and Safeguarding to Coordination and Adaptation." *Journal of Management* 40(1):193-225.
Selviaridis, K., and W. van der Valk. 2019. "Framing contractual performance incentives: Effects on supplier behaviour." *International Journal of Operations and Production Management* 39(2):190-213.
Tahawy, K. E., and S. E. Ibrahim. 2015. "Factors Affecting Pre-Contractual and Post-Contractual Outsourcing Relationships." *Academy of Management Annual Meeting Proceedings* 2015(1):1-.
Tanenbaum, W. A. 2006. "Why Combining Force Majeure and Disaster Recovery Provisions Makes Sense." *Intellectual Property & Technology Law Journal* 18(10):17-20.
Thomas, M., M. Anoop, and W. Antoinette. 2007. "Trust and Formal Contracts in Interorganizational Relationships: Substitutes and Complements." *Managerial and Decision Economics*(8):833.
Vlachos, I. 2012. "The impact of transaction cost economics and competitive factors upon the design of contracts: a study using DEMATEL." *International Journal of Decision Sciences, Risk and Management* 4(3/4):217.
Wacker, J. G., C. Yang, and C. Sheu. 2016. "A transaction cost economics model for estimating performance effectiveness of relational and contractual governance: Theory and statistical results." *International Journal of Operations & Production Management* 36(11):1551-75.
Wang, J., and H. Shin. 2015. "The impact of contracts and competition on upstream innovation in a supply chain." *Production and Operations Management* 24(1):134-46.
Wang, M., Q. Zhang, Y. Wang, and S. Sheng. 2016. "Governing local supplier opportunism in China: Moderating role of institutional forces." *Journal of Operations Management* 46(1):84-94.
Wang, W. 2010. "A model for maintenance service contract design, negotiation and optimization." *European Journal of Operational Research* 201(1):239-46.
Watanabe, K., S. S. Bánkuti, and A. E. Lourenzani. 2017. "“Pingado Dilemma”: Is formal contract sweet enough?" *Journal of Rural Studies* 54:126-37.
Weber, L., and K. J. Mayer. 2011. "Designing effective contracts: Exploring the influence of framing and expectations." *The Academy of Management Review* 36(1):53-75.

Weber, L., K. J. Mayer, and J. T. Macher. 2011. "An analysis of extendibility and early termination provisions: The importance of framing duration safeguards." *The Academy of Management Journal* 54(1):182-202.

White, G. O., T. Hemphill, T. Weber, and K. Moghaddam. 2018. "Institutional origins of WOFS formal contracting: A judicial arbitrariness perspective." *International Business Review* 27(3):654-68.

Wilkinson-Ryan, T. 2009. "Do Liquidated Damages Encourage Breach - A Psychological Experiment." *Michigan Law Review* 108:633.

Williamson, O. E. 1987. *The economic institutions of capitalism : firms, markets, relational contracting*: Free Press: New York.

Wu, K.-J., M.-L. Tseng, A. S. F. Chiu, and M. K. Lim. 2017. "Achieving competitive advantage through supply chain agility under uncertainty: A novel multi-criteria decision-making structure." *International Journal of Production Economics* 190:96-107.

Yang, D., S. Sheng, S. Wu, and K. Z. Zhou. 2018. "Suppressing partner opportunism in emerging markets: Contextualizing institutional forces in supply chain management." *Journal of Business Research* 90:1-13.

Yeung, K., P. K. C. Lee, A. C. L. Yeung, and T. C. E. Cheng. 2013. "Supplier partnership and cost performance: The moderating roles of specific investments and environmental uncertainty." *International Journal of Production Economics* 144:546-59.

Yin, R. K. 1989. *Case study research: design and methods*, *Working papers in applied social research: 5*: Sage.

Yuan, Y., B. Feng, F. Lai, and B. J. Collins. 2018. "The role of trust, commitment, and learning orientation on logistic service effectiveness." *Journal of Business Research* 93:37-50.

DECISION SCIENCES INSTITUTE

The Effects of Reliability Growth in Performance-based Logistics Contracts

Hasan Uvet
Georgia Gwinnet College
Email: huvet@ggc.edu

Sedat Cevikparmak
University of North Texas
Email: sedat.cevikparmak@unt.edu

Hasan Celik
University of North Texas
Email: hasan.celik@unt.edu

Saban Adana
University of North Texas
Email: saban.adana@unt.edu

ABSTRACT

This paper investigates the motivations behind investment for reliability on availability of the systems and total service cost in performance-based logistics contracts. The effects of interactions between reliability, spare parts, and the length of the contract on availability and service costs were analyzed with a simulation using BlockSim.

KEYWORDS: Performance-based logistics, performance-based contracts, reliability improvement, spare parts

INTRODUCTION

In the late 1990s, DoD came up with a new logistics model based on paying for performance rather than a transactional model to overcome these problems (Geary & Vitasek, 2005). The DoD called this model a performance-based logistics (PBL), which is also known as "power by the hour" and “outcome-based service contracts” in the private sector, is a paradigm shift in the DoD’s management (Ng and Nudurupati, 2010; Ong et al., 2005). The DoD defined PBL as a “… the purchase of support as an integrated, affordable, performance package designed to optimize system readiness and meet performance goals for a weapons system through long-term support arrangements with clear lines of authority and responsibility” (DAU, 2005).

The problems started with the overuse of the systems which bring about high life cycle costs in DoD, because of extended maintenance, hardware extinction, and structural tiredness (Berkowitz, Gupta, Simpson, & McWilliams, 2005; Devries, 2005; Kratz & Buckingham, 2010). Additionally, considering projected life of the systems, this longevity causes more maintenance time that reduces readiness of systems. These problems result with spending 72% of the total funds of the life cycle costs for sustainment of the systems (Berkowitz, Gupta, Simpson, & McWilliams, 2005). Furthermore, considering the continued budgetary pressure in addition to this high operation and maintenance costs, DoD required looking for new approaches to decrease life cycle costs while maintaining readiness of systems (Kobren, 2009; Kratz & Buckingham, 2010).

A transaction-based logistics approach, which is the traditional approach of DoD, is getting more inefficient, because of the 31% increase in the cost of sustainment in a transaction-based logistics approach in one year (Randal, Pohlen, and Hanna, 2010; GAO, 2005). Labor and material costs are the main components of these high expenses in the transaction-based logistics. (Gansler & Lucyshyn, 2006). Due to the high profit in after sales support for providers, the transaction-based approach could not align the customers' requirements with contractors' preferences (Keating & Huff, 2005). Thus, the transaction-based approach places the burden of risk on customers (Quick, 2011). Because of opportunism of contractors whose business model depends on the spare parts, contractors are not concerned with the life cycle cost of systems, thereby the reliability growth. (Hunter; 2006b; Gansler & Lucyshyn, 2006; GAO, 2005; Quick, 2011).

PBL contracts differ from transaction-based contracts in that the suppliers are rewarded when they achieve the desired performance objectives (Berkowitz, Gupta, Simpson, & McWilliams, 2005; Geary and Vitasek, 2005). This approach makes increasing system performance and reducing life cycle cost of systems possible (Berkowitz, Gupta, Simpson, & McWilliams, 2005; Geary and Vitasek, 2005; Jin, Tian, and Xie, 2014; Kim, Cohen, and Netessine, 2007; Randal, Pohlen, and Hanna, 2010).

The DoD is able to share the risk with a supplier with the advantage of this new logistics support strategy based on the pay for performance strategy (Gansler & Lucyshyn, 2006). PBL contracts motivate suppliers to make an investment to increase the reliability of systems that result in high operational readiness of systems (Gansler & Lucyshyn, 2006; Vitasek et al., 2007). This upfront investment by suppliers improving affordability over the remaining life of the supported system is affecting the spare and repair strategy (Kim, Cohen, and Netessine, 2007; Randal, Pohlen, and Hanna, 2010).

However, besides these advantages, PBL consists of some potential risks, such as supplier opportunism, selection of the wrong supplier (Gardner, 2008). In PBL, the contract identifies what is required such as target availability of systems, but the supplier determines how to attain the objectives (Kim, Cohen, and Netessine, 2007). How to get the required outcome is dependent on provider's decisions about strategies, such as making an upfront investment to increase reliability, holding spare parts to increase readiness, and increasing service capacity to reduce mean time to repair. In PBL, customers expected a high level of product reliability which is the one of the three fundamental features of enabling mission capability, along with availability and maintainability (DoD 2005). However, considering to supplier opportunism and uncertainties in long-term contracts, suppliers reluctant to increase the reliability of components (Stauss et al. 2010).

The primary purpose of this study is to understand the motivations behind upfront investment for reliability growth and the effects of reliability improvement on availability of the systems and total service cost in PBL. The other goal of this study is to understand the effects of interactions between reliability, spare parts, and the length of the contract.

The primary research question of this study is 'What is the consequences of reliability growth in PBL?' As secondary research questions, we will investigate these issues: How can we minimize design cost of systems while increasing the reliability of systems in PBL? What are the consequences of reliability growth on the availability of systems and service cost in PBL?

The remainder of this paper is organized as follows. The literature review will be presented in Section 2. In Section 3, research methodology and mathematical models are represented. In Section 4 and 5, a numerical example is described for optimum reliability allocation, and the effects of reliability growth on the availability of systems and service cost in PBL, respectively. The paper will be concluded in Section 6 with a discussion of limitations, future research, and conclusion.

LITERATURE REVIEW

PBL offers an exceptional model, win-win-situation, for customers and suppliers (Stauss, Hypko, et al. 2010). All efforts in PBL, such as maintenance programs and inventory management, are designed to reduce the cost of ownership for sustaining systems (Gansler & Lucyshyn; 2006; Vitasek et al., 2007). Fewer maintenance actions in PBL create more high-profit margins for suppliers, because of the high availability rates (Ong et al., 2005). Additionally, the supplier is motivated by incentives to invest the reliability of systems that result in high operational readiness of systems (Gansler & Lucyshyn, 2006; Vitasek et al., 2007). This investment for reliability will increase availability and readiness of systems while reducing total service cost, spare parts acquiring, holding costs, and logistic footprint (Cohen, M. A., & Netessine, S., 2007; Devries, 2005; Gansler, 2006).

Many studies are conducted about a trade-off between spare parts and reliability. Kumar et al. (2007) developed a multi-objective optimization model (goal programming model) to optimize reliability, maintainability and supportability that are important in PBL support contracts under PBL (Kumar, Nowicki et al. 2007). Sols et al. (2007) proposed a framework for PBL contracts and describes the fundamental features of a successful PBL initiative. In their study, which is based on the literature survey and the interviews, they come up with three essential characteristics - will, purpose, and reward - for efficient and effective PBL contractual agreements. They also proposed that successful implementation of PBL contracts depend on the appropriate selection of the metrics that show the effectiveness of systems and the adoption of a fair reward scheme (Sols, Nowicki et al. 2007). Nowicki et al. (2008) developed an optimization model for the spare asset allocation problem under three different revenue functions: step, exponential and linear revenue functions. They found that an optimal spares asset allocation can not be sustained without considering the associated profit stream (Nowicki, Kumar et al. 2008). Kim et al. (2010) analyzed the efficiencies of two contract types used in PBL, based on sample-average downtime and based on cumulative downtime, using a principal-agent contracting framework. They found that when a component is highly reliable, implementation of performance-based contracts (PBCs) may create high agency cost (Kim, Cohen et al. 2010). Mirzahosseinian and Piplani (2011) investigated how do component reliability, maintenance facility, and inventory management affect the availability of systems under the PBL. They found that the base stock level of the spare parts had an insignificant impact on the system availability. Thus, to attain a minimum target availability level, the supplier has to enhance the component reliability and the repair time, rather than increase the stock of spares (Mirzahosseinian and Piplani 2011). Jin et al. (2012) investigated the main factors that affect the system performance, and how they interact with each other to achieve the operational availability under PBC. They found that there is strong relationship amongst system cost, reliability, and spare parts stocking (Jin et al., 2012). They showed that under a longer service agreement in PBC, suppliers as OEMs are eager to invest in reliability improvement. Also, they revealed that increasing the spare parts inventory has less impact on the availability of highly reliable systems (Jin et al., 2012). Guajardo et al. (2012) showed in his experimental study, which was based on the real data of the Rolls-Royce company, how product reliability was impacted by the use of two various support strategies: time and material based (T&MC) strategy and performance-based strategy. They found that the reliability of products is much higher (25%–40%) under PBC than under T&MC (Guajardo, Cohen, et al., 2012). Jin and Tian (2012) investigated how operational availability can be achieved while minimizing the logistics footprint through increasing reliability under PBC. In their study, trade-offs between the reliability design and the spare part level are examined and compared in different scenarios. They showed that the multi-phase replenishment policy is effective to allocate spare parts to a changing installed base (Jin and Tian 2012). Kim et al. (2015) developed a game-theoretical model to investigate the interaction investment in spare assets and product reliability under the traditional resource-based contract (RBC) and PBC. They found that PBC motivates suppliers to make the upfront investment for reliability improvement by powerful incentives, which creates savings in acquiring and holding spare assets (Kim, Cohen & Netessine, 2015). On the contrary, under the RBC, the supplier invests more in inventory and less in reliability (Kim, Cohen & Netessine, 2015). Bakshi and Kim et al. (2015) developed the game-theoretical model, a principal-agent model, to investigate the interaction between reliability signaling and the vendor's unrestricted investment in spares inventory. They found that customers are eager to accept PBC when mature technologies are available for acquired products rather than products with newly developed technology (Bakshi, Kim, et al. 2015).

Although there are many studies about a trade-off between spare parts and reliability, there is no study that directly investigates the effects of the reliability of systems on availability and service cost in PBL. Also, unlike the previous studies, target reliability allocation of systems was attained considering a minimum and a maximum reliability of subsystems under the various feasibility rates. Supplier's reservation utility and supplier's utility function were conducted under the agency theory perspective based on the two different contract types: Fixed Firm Price (FFP) and Cost Plus Award Fee (CPAF). The 7other contribution of this study is the investigation of the direct impact of reliability growth to availability rates and service costs.

THEORETICAL DEVELOPMENT and MATHEMATICAL MODEL

Nomenclature

A	System or subsystem availability
$A\min$	Availability target
$MTBF$	Mean time between failure
MDT	Mean downtime
$MTTR$	Mean time to repair
$MLDT$	Mean logistics delay time
C_T	Total service cost
$c_i\ (sc_i)$	Spare cost of component/subsystem i
$c_i\ (mc_i)$	Maintenance cost of component/subsystem i
$c_i\ (shc_i)$	Holding cost of component/subsystem i
C_D	Total system design cost
$c_i\ (R_i)$	Design cost of component/subsystem i
R_i	Reliability of component/system i,
n	Number of components within the system considered in the optimization
$R_{i,min}$	Minimum reliability of component/subsystem i
$R_{i,max}$	Maximum achievable reliability of component/subsystem i
R_S	System reliability
R_T	System reliability target
f_i	Feasibility of increasing the reliability of subsystem I
I_r	Reliability importance

β_i Weibull shape parameter of component/system i,
α_i Weibull scale parameter of component/system i,
a, b_1, b_2 Fixed price, reward, and the penalty for suppliers, respectively.

In this study, we conducted a mixed methodology of quantitative research. The effects of reliability growth, spare parts, and the length of the contract in PBL will be analyzed with simulation (BlockSim).

In PBL, the customer outsources the task of delivering the system's performance to the supplier (Helander and Moller, 2008). This relationship between the supplier (agent) and customer (principal) creates the agency problem that emerges when the preferences and goals of the principal and agent are in conflict (Eisenhardt, 1989). Although customers have information about supplier's capability, such as service capacity, technological capability, and technical ability, they cannot forecast the supplier's action in the performing of the contract (Bergen et al., 1992). Because of the imperfect information about the agent's performance, it is getting hard to detect whether the agent performs according to the principal's interest or not (Eisenhardt, 1989).

The performance-based payment tied to achievement of the outcomes transfers risks from the principal (customer) to the agent (supplier) (Eisenhardt, 1989; Firchau, 1987). Even if the incentives inherent to performance-based payment in PBL transfers risks from the customer to the supplier, it still contains risks, which mentioned above, for customers. Considering customers' inability to observe suppliers' actions in PBL, because of the imperfect information, customers need to mitigate this risk in PBL, especially in long-term contracts. To reduce the risks that emerge from agency problems, in PBL, the payment based on the outcome enable alignment of the customers' goals with the supplier's preferences. Customers conduct incentives in the various contract types such as fixed price incentive fee (FPIF), cost plus incentive fee (CPIF), and cost plus award fee (CPAF) (Sols, Nowicki, and Verma, 2007). With these contract types, customers want to motivate the suppliers to act in agreement with the customers' preferences (Firchau, 1987).

One of the fundamental characteristics of PBL tied to arrangements, which is called as tenets of PBL, to get desired objectives is meaningful and significant incentives to supplier tied to achievement of the outcomes (DoD, 2014; UT, 2012). Therefore, it is significant for the customer to propose reasonable incentives and to make effective contracts align the preferences of supplier and his/hers. In our study, we show the relationship between reliability growth and motivation in different contract types.

Mathematical Model

In PBL, customer's incentives to motivate the supplier to make the upfront investment should be meaningful and significant for the supplier. The supplier will seek for a design that will sustain the target reliability of the system at a minimum cost. From the perspective of agency theory, considering the constraints involving the supplier, the customer must ensure that the agent gets at least a reservation level of utility. In that point, the customer can only influence the provider by his choice of the incentive payment.

We modeled the performance-based contract as a principal-agent game-theoretical model. Two players involved in the game are the supplier and the customer. The supplier behaves as an agent whose goal is to minimize the total design cost of the system while achieving the reliability target. Furthermore, the provider's ultimate goal with the result of this reliability growth is to maximize his service profit. The customer is the principal who wants to achieve reliability growth criteria by stipulating the reward function.

With the following models, denoted as Problem P1 and Problem P2, we will show the optimum reliability allocation to reach desired availability target while minimizing total design cost. For the purpose of the first research question, the minimization function will be used to find the objective reliability of systems subject to minimum and maximum reliability of subsystems, agent's reservation utility function, and agent's utility function. These functions depend on the agency theory and reliability theory.

Problem 1:

To find the optimum reliability allocation for supplier with a minimum cost; considering to achieve reliability target with implementation of Firm-Fixed-Price (FFP) contract is shown below:

The customer (principal) has utility function $\mathrm{B(y)}$ for reliability.

- If the outcome is R_T and the supplier (agent) is paid a,
- The customer's utility is $\mathrm{B}\,(R_T\ - R_S)$ (1)

In this equation, R_T is the target reliability of the system, after reliability enhancement; R_S is the current reliability if the system, without reliability improvement.

The supplier's (agent's) utility function is given by (adapted from Jin et al., 2012 and Mettas A., 2000),

$$Min\ C_D = \sum_{i=1}^{n} B_i * c_i(R_i) \ , \tag{2}$$

where

$$c_i\left(R_i; f_i;\ R_{i,min};\ R_{i,max} = e^{\left[(1-f_i)\frac{R_i - R_{i,min}}{R_{i,max}-R_i}\right]}\right) \tag{3}$$

In this equation, $R_{i,max}$ is the maximum achievable reliability of the subsystems or components while $R_{i,min}$ is the minimum inherent reliability of the subsystems or components. R_i represents the current reliability of the subsystems or components. f_i describes the feasibility of increasing the reliability of subsystem or component i and takes values between 0 and 1. It represents the difficulties, such as design complexity, technological limitations, etc., in increasing the subsystem/component's reliability relative to the others. The small feasibility rate indicates the reliability growth of the subsystems or components are difficult to improve. Additionally, this small feasibility rate results in a high design cost to increase the reliability of the subsystems or components (Mettas A., 2000). B_i represents the basic design cost of the subsystems or components.

Reliability of the system

$$R_s(t) = \prod_{i=1}^{n} R_i\ (t) = R_1(t) * \ R_2(t) * \ R_3(t) * \dots * R_n(t) \tag{4}$$

where

$$R(t) = e^{-(\alpha t)^{\beta}} \tag{5}$$

β_i and α_i Weibull shape and scale parameters of component/system i, respectively.

$$R_{i,min}(t)\ \le\ R_i\ (t) \le R_{i,max}(t),\ \ i = 1,2 \dots, n. \tag{6}$$

From equation 2 to 6;

$$Min\ C_D = \sum_{i=1}^{n} B_i\ *\ e^{\left[(1-f_i)\frac{R_i - R_{i,min}}{R_{i,max}-R_i}\right]} \tag{7}$$

Subject to

$$R_s(t) = R_1(t) * R_2(t) * R_3(t) \geq R_T(t) \quad (8)$$

$$R_{1,min}(t) \leq R_1(t) \leq R_{1,max}(t), \quad (9)$$

$$R_{2,min}(t) \leq R_2(t) \leq R_{2,max}(t), \quad (10)$$

$$R_{3,min}(t) \leq R_3(t) \leq R_{3,max}(t), \quad (11)$$

Problem 2:

To find the optimum reliability allocation for supplier with a minimum cost; considering to achieve reliability target with implementation of Cost Plus Incentive Fee (CPIF) contract is shown below:

The customer (principal) has utility function $\mathrm{B(y)}$ for reliability.

- If the outcome is R_T and the supplier (agent) is paid b_1,
- The principal's utility is $\mathrm{B}\,(R_T - R_S)$

Where the linear revenue functions are (adopted from Nowicki et al. (2008))

$$r(R_S) = \begin{cases} a + b_1(R_S - R_T) \ for\ R_S \geq R_T \\ a + b_2(R_T - R_S) \ for\ R_S < R_T \end{cases} \quad (12)$$

Here R_T is the target system reliability designated by the customer. a, b_1 and b_2 are positive parameters determined by the customer. In particular, a is the fixed payment; and b_1 and b_2 are the reward and penalty rate, respectively (or the piece rate in game theory). A larger value of $b1$ (or $b2$) implies that the OEM receives more compensation (or penalty) given the same level of R_S depending on whether $R_S \geq R_T$ or $R_S < R_T$.

The supplier's (agent's) utility function is given by

$$u\,(b_1, C_D) = b_1(R_S - R_T) - C_D \quad (13)$$

where C_D is the level of effort the supplier chooses to increase reliability and b_1 is the award in PBL that the supplier receives.

The agent has reservation utility $u0$ is given by

$$u0\,(b_2, C_D) = C_D - b_2\,(R_T - R_S) \quad (14)$$

where b_2 is the penalty in PBL that the supplier receives.

From equation 2 to 6 and 12-14;

$$Min\ C_D = \sum_{i=1}^{n} B_i * e^{\left[(1-f_i)\frac{R_i - R_{i,min}}{R_{i,max} - R_i}\right]} \quad (7)$$

Subject to

$$R_s(t) = R_1(t) * R_2(t) * R_3(t) \geq R_T(t) \quad (8)$$

$$R_{1,min}(t) \leq R_1(t) \leq R_{1,max}(t), \quad (9)$$

$$R_{2,min}(t) \leq R_2(t) \leq R_{2,max}(t), \quad (10)$$

$$R_{3,min}(t) \leq R_3(t) \leq R_{3,max}(t), \quad (11)$$

$$u\ (b_1, C_D) = b_1(R_S\ - R_T) - C_D \geq 0 \tag{15}$$

$$u0\ (b_2, C_D) = C_D - \ b_2\ (R_S\ - R_T)\ \leq 0 \tag{16}$$

Problem 3:

To minimize the total service cost is essential for the supplier. All of the suppliers' efforts concentrate on reducing this cost in PBL while achieving the desired outcomes. However, how this cost is affected by suppliers' action still not clear, especially in long-term contracts. With the third problem, we will investigate the effects of reliability growth on the availability of systems and service cost in PBL. For this purpose, simulation (BlockSim) will be used to find the consequences of reliability growth on the availability of systems and service cost. Our assumptions are; normal distribution will be utilized for maintenance time and logistic delay time.

$$C_T(s, R_S, t) = \sum_{i=1}^{n} c_i(R_i) + \sum_{i=1}^{n} c_i(sc_i) + \sum_{i=1}^{n} c_i(shc_i) + \sum_{i=1}^{n} c_i(mc_i) \tag{17}$$

$$A = \frac{MTBF}{MTBF+MDT} \tag{18}$$

where

$$MDT = MTTR + MLDT \tag{19}$$

The comparison of the effects of system reliability on system availability and total service cost will be made as a result of this problem. In this issue, C_T is the total service cost for the supplier and it is affected by the reliability of the system, number of spare parts, and length of contract. In this equation, $c_i\ (R_i)$ is the design cost of component/subsystem i, $c_i\ (sc_i)$ is the spare cost of component/subsystem i, $c_i\ (mc_i)$ is the maintenance cost of component/subsystem i, and $c_i\ (shc_i)$ is the holding cost of component/subsystem i. Availability is the primary performance metric and defined as where MTBF is the mean time between failures. MDT is the mean downtime which is the sum of mean time to replacement (MTTR) of a part and mean logistics delay time (MLDT) (Jin et al. 2014).

SIMULATION AND NUMERICAL EXAMPLE

Reliability Allocation

The system consists of three components connected in a series configuration. The system will fail if any of its subsystems fails (i.e., series configuration). Each system has different reliability and different feasibility rates. Each system has different reliability and different feasibility rates. The target reliability of the system is 90% within the time consideration of 200 hours. The Weibull shape, scale parameters, feasibility, and basic design cost of component/system i are shown in Table 1:

Table 1: Reliability and Cost Parameters of Components

Parameter		Value
β_i	β_1 (HardDriveBlock_1) β_2 (Main BoardBlock_2) β_3 (PowerSupplyBlock_3)	1,3 1,35 2,1
α_i (hours)	α_1 , α_2 , α_3	420 – 650 - 590
f_i	f_1 , f_2 , f_3	0.2 - 0.8 - 0.5
B_i (dollars)	B_1, B_2 , B_3	240.000,60.000,150.000

Analysis of Current Reliability of System/Subsystems:

The reliabilities, probability of failure rates, and importance of each component and reliability of system for 200 hours are shown in Table 2:

Table 2: Current Reliability of System and Components			
Mission End Time (Hr)	200		
	Block Outputs		
Block Name	*Reliability*	*Prob. of Failure*	*Importance*
HardDriveBlock_1	0.683064	0.316936	0.735789
Main BoardBlock_2	0.81572	0.18428	0.616132
PowerSupplyBlock_3	0.902012	0.097988	0.557189
	Reliability	*Prob. of Failure*	
System	0.502591	0.497409	
Failure Rate (Hr) =	0,004935/Hr		
Mean Life (Hr) =	231,947982 Hr		

Analysis of Target Reliability of System/Subsystems:

According to results of the first problem; the reliability allocation of the system considering minimum design cost is shown in Table 3-4. New total reliability of the system is presented in Table 5. Block and system reliability versus time shown in Figure 1-2.

Table 3: Reliability Allocation of System						
Block Name	$R_{i,max}$	$f_{i,}$	Reliability importance of component i (200 hr)	$R_{i,C}$ (200 hr)	$R_{i,T}$ (200 hr)	$B_{i,DC}$ (200 hr) (dollars)
PowerSupplyBlock_3	0,999	0,5	0,557189	0,902012	0,973752	35.457
Main Board Block_2	0,999	0,8	0,616132	0,81572	0,953352	12.075
HardDrive Block_1	0,999	0,2	0,735789	0,683064	0,969485	621.072
System	0,999	-	-	0.502591	0,90	668.604

$R_{i,max}$: Maximum achievable reliability of component i.
f_i: Feasibility of component i.
$R_{i,C}$: Current reliability of component I (200 hr).
$R_{i,T}$: Target Reliability of component i (200 hr).
$B_{i,DC}$: Design cost of component of component i (200 hr) (dollars).

Minimum total design cost of the system for R_T (0.90) is $668.604. Therefore, customers should have to consider this minimum total design cost of the system for award and penalty amount in the contract to motivate them to make the upfront investment to increase the reliability of systems. Additionally, this total design cost can reduce the fleet size. Thus, fleet size of the systems directly affects these amounts, design cost, award, and penalty. With this solution, meaningful and significant incentives (award and penalty), which are the one of the tenets of PBL tied to arrangements (DoD, 2014; UT, 2012), become more predictable for the customer.

The long-term effects of this reliability growth on system availability and total service cost will be analyzed below with Problem 3 and Problem 4.

Table 4: Target Reliability of System and Components

Block Name	Reliability	Prob. of Failure	Importance
HardDriveBlock_1	0,969487	0,030513	0,928324
Main BoardBlock_2	0,953579	0,046421	0,943811
PowerSupplyBlock_3	0,973516	0,026484	0,924482
	Reliability	Probability of Failure	
System	0,90	0,10	

Figure 1: Block Reliability versus Time

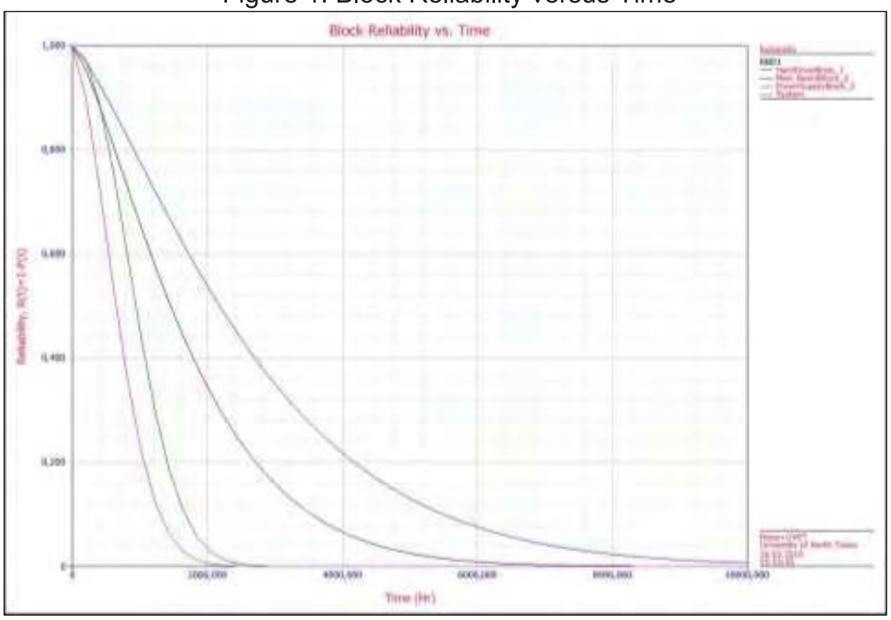
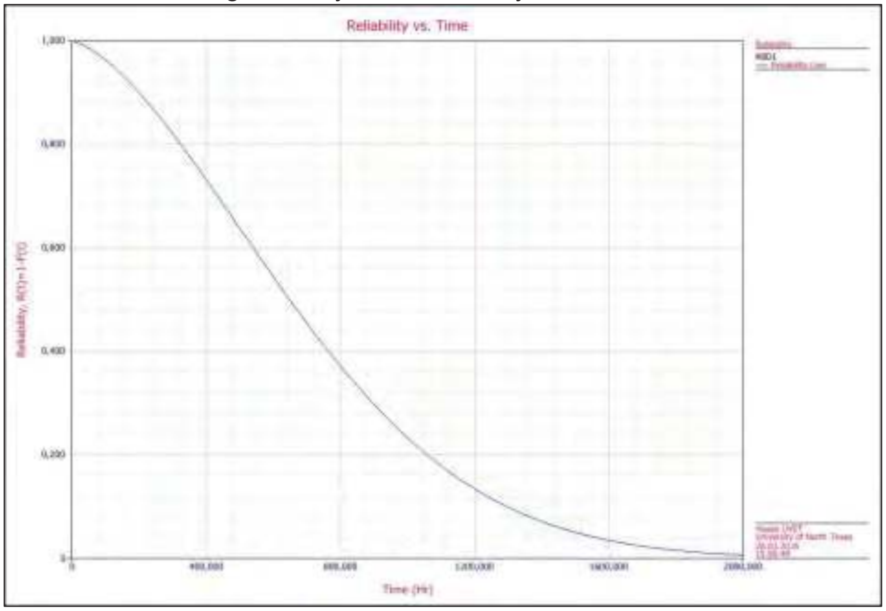


Figure 2: System Reliability versus Time



The Effects of Reliability Growth on Availability and Service Costs

In this study, we adopt a single-echelon service supply chain to support the operation of repairable systems. There are many studies have been conducted in the area of service parts inventory systems. The multi-echelon technique for a recoverable item control (METRIC) model, which was introduced by Sherbrooke (1968), is one of the first studies of repairable inventory model, and it is necessary for controlling expensive critical parts that are replaced correctively (Driessen et al., 2010). The goal of the METRIC model is to minimize inventory investment and to lower the holding or backorder cost subject to service level requirement (Sherbrooke, 1968). The stock holds various levels of spare parts per type in multi-item systems in which consists of several component types. In the event of failure, parts are sent to the repair center for reconditioning or repair. After maintenance, they are sent back to the stockroom under the one-for-one replenishment policy. The spare parts cost, maintenance cost, holding cost, and the other variables of each component/subsystem used in the simulation are shown in Table 5.

Table 5: Cost and Maintenance Data of component/subsystem i

	HardDriveBlock_1	MainBoardBlock_2	PowerSupplyBlock_3
$c_i\ (sc_i)$ (\$)	25000	27000	20640
$c_i\ (shc_i)$ (\$)	2000	2000	2000
Downtime Rate	Weibull Distribution		
Task Duratıon	Exponential Distribution		
Logistic Delay	Normal Distribution (Mean: 11; Std.: 4; Unit: Hour)		
Crew Task	Exponential Distribution (limit of tasks this crew can perform: 1)		
Cost Per Task	Normal Distribution (Mean: 250; Std.: 50; Unit: Dolar)		

$c_i\ (sc_i)$: Spare parts cost of component/subsystem i (\$)
$c_i\ (shc_i)$: Holding cost of component/subsystem i (\$)

Simulation (BlockSim) results with current reliability of the system

This simulation was conducted with current reliability, without enhancement of reliability of components, under the four various time considerations - 200 hr, 500 hr., 1000 hr., one year- with one spare capacity for the each part of the system. The results of this simulation were presented in Table 6.

Table 6: Simulation results for current reliability (spare parts =1)

System Overview	t=200 hr, s=1, current R_S	t=500 hr, s=1, current R_S	t=1000 hr, s=1, current R_S	t=1 year, s=1, current R_S
General				
Mean Availability (All Events):	**0.929974**	**0.897894**	**0.875977**	**0.859233**
Reliability(200):	0.529	0.074	0.001	0
MTBF (Total Time) (Hr):	347.222	253,678	211.729	190.942
Number of Failures:	0.576	1.971	4.723	45.244
System Cost Summary				
Misc. Corrective Costs:	$3,649	$13,263	$32,198	$315,524
Costs for Parts (CM):	$14,400	$49,275	$118,075	$1,131,100
Total CM Costs:	$18,049	$62,538	$150,274	$1,446,624
Block Consequential Costs:	$16,777	$55,891	$131,679	$1,248,137
Pool Holding Costs:	$1,930	$5,551	$11,196	$97,484
Total Cost (Service and Spare Costs):	**$36,757**	**$123,981**	**$293,149**	**$2,792,245**

The results of simulation show that the availability of the systems decreases with the increase of the length of the contract. While the availability rate was 93% under the time consideration of 200 hr., this availability rate drops almost 7% to 86% under the one-year contract. On the other hand, the total cost of supplier increases from $36.757 to $2.792.245 in a year. This cost is not utility function of the supplier since it does not include incentives (penalty and reward) for the provider based on to attain the target objectivity of contract in PBL. Considering to 90% target availability was determined by the customer in PBL contracts, then depending on the results of the simulation, the supplier will be penalized for the low availability rates under the 500 hr-1000 hr- 1-year length of contracts. Therefore, provider’s total cost will increase with the penalty cost under the PBL. We can conclude that, besides the incentives (award and penalty) for the supplier, also, the length of contract and target availability rates are significant for the supplier’s decision to make the upfront investment to increase the reliability of the system.

According to results of simulation under the time consideration of current reliability, 500 hr., and one spare capacity the graphs of availability and reliability versus time, total service cost versus time, and system failures are shown in Figure 3, 4, and 5.

Figure 3: Availability and Reliability versus Time (R_C, 500 hr.,1)

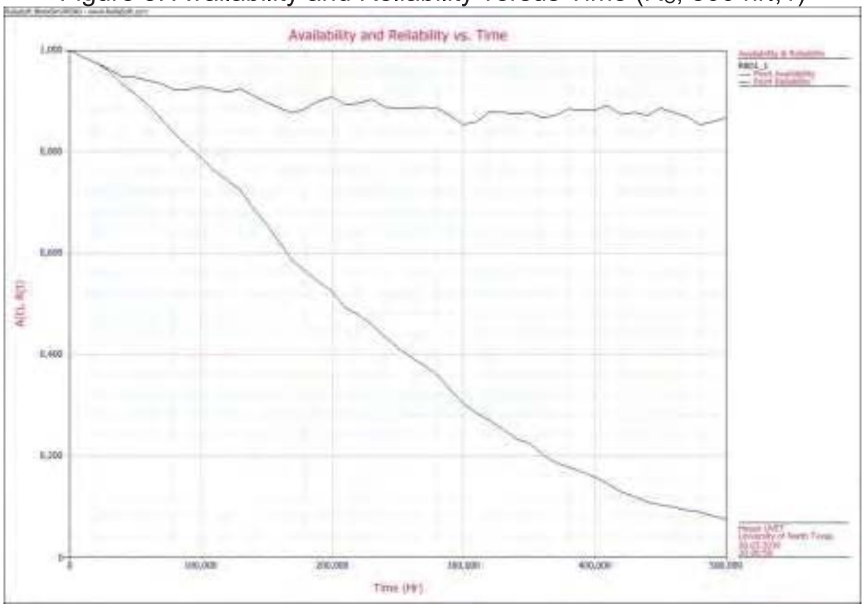
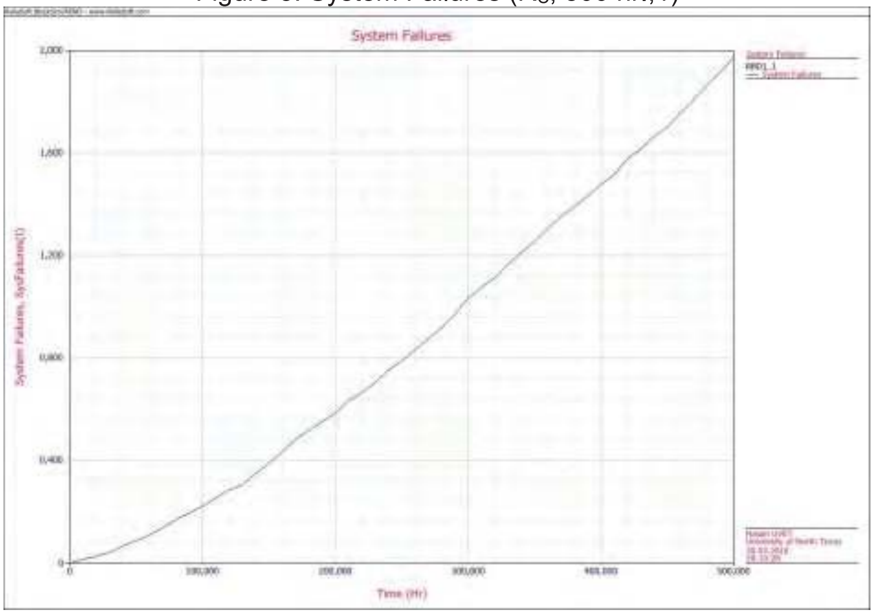


Figure 4: Total Service Cost versus Time (R_C, 500 hr.,1)

Figure 5: System Failures (R_C, 500 hr.,1)



The second simulation was conducted with current reliability, without enhancement of reliability of components, under the four various time considerations - 200 hr, 500hr., 1000 hr., one year- with spare acquisition type, fixed probability (0.1) of stock out for the each component of the system. The results of this simulation were presented in Table 7.

Table 7: Simulation results for current reliability (spare acquisition type: fixed probability (0.1) of stockout)

System Overview	t=200 hr, s=0, R_S (200hr)	t=500 hr, s=0, R_S(200hr)	t=1000 hr, s=0, R_S(200hr)	t=1 year, s=0, R_S(200hr)
Mean Availability (All Events):	**0.923873**	**0.888037**	**0.869557**	**0.853627**
Reliability(200):	0.519	0.066	0.001	0
MTBF (Total Time) (Hr):	336.134	243.664	213.949	192.384
Number of Failures:	0,595	2,052	4,674	44,905
System Cost Summary				
Misc. Corrective Costs:	$3,958	$14,515	$33,794	$327,478
Costs for Parts (CM):	$1,300	$4,825	$11,350	$111,375
Total CM Costs:	$5,258	$19,340	$45,144	$438,853
Block Consequential Costs:	$17,419	$58,112	$131,227	$1,247,450
Total Costs (Service and Spare Costs)	**$22,677**	**$77,452**	**$176,371**	**$1,686,303**

The results of simulation show that the availability of the systems decreases with the increase of the length of the contract. While the availability rate was 92.4% under the time consideration of 200 hr., this availability rate decreases to 85.3% under the one-year contract. On the other hand, the total cost of supplier increases from $ 22.678 to $1.686.303 in a year. The results of the simulation show that considering to 90% target availability was determined by the customer in PBL contracts; the supplier will be penalized for the low availability rates under the 500 hr-1000 hr-1 year length of contracts.

Comparison of two simulations under the different spare parts (1 and 0) shows that the availability rates did not change significantly. There is almost 0.6 % decrease in this availability rates. On the other hand total cost of the supplier, severely affected with the spare parts. Under the one-year contract length, the total costs of provider reduced from $2.792.245 to $1.686.303. To be able to show this comparison clearly, we put the results of two simulations in one table which are presented in Table 8.

Table 8: Simulation results for current reliability under the two different spare parts

System Overview	t=200 hr	t=500 hr	t=1000 hr	t=1 year
s=1, current R_S				
Mean Availability (All Events):	0.929974	0.897894	0.875977	0.859233
Total Costs:	$36,756	$123,981	$293,149	$2,792,244
s=0, R_S (200hr)				
Mean Availability (All Events):	0.923873	0.888037	0.869557	0.853627
Total Costs:	$22,.677	$77,452	$176,371	$1,686,303

<u>Simulation (BlockSim) results with target reliability of the system:</u>

This simulation was conducted with target reliability, after enhancement of reliability of components, under the four various time considerations - 200 hr, 500hr., 1000 hr., one year- with one spare capacity for the each part of the system. The results of this simulation were presented in Table 9.

Table 9: Simulation results for target reliability (spare parts =1)

System Overview	t=200 hr, s=1, R_T(200hr)	t=500 hr, s=1, R_T(200hr)	t=1000 hr, s=1, R_T(200hr)	t=1 year, s=1, R_T(200hr)
Mean Availability (All Events):	**0.987722**	**0.978038**	**0.967664**	**0.951943**
Reliability(200):	0.901	0.636	0.231	0
MTBF (Total Time) (Hr):	1904.76	1168.22	818.33	561.85
Number of Failures:	0.105	0.43	1.22	15.38
System Cost Summary				
Misc. Corrective Costs:	$641	$2,854	$8,391	$107,671
Costs for Parts	$2,625	$10,700	$30,550	$384,400
Total CM Costs:	$3,266	$13,554	$38,941	$492,071
Block Consequential Costs:	$2,765	$10,952	$30,712	$382,553
Pool Holding Costs:	$1,766	$5,315	$11,192	$97,942
Total Service and Spare Costs:	$7,797	$29,821	$80,845	$972,566
Design Cost	*$668,604*	*$668,604*	*$668,604*	*$668,604*
Total Costs:	**$676,401**	**$698,425**	**$749,449**	**$1,641,170**

The results of simulation show that there is an apparent growth in the availability rates. The availability of the systems decreases with the increase of the length of the contract. While the availability rate was 98.7% under the time consideration of 200 hr., this availability rate drops only 3.5% to 95.2% under the one-year contract. This decreased rate is the almost half of the reduce under the R_S (200hr). On the other hand, the total cost of supplier increases from $676.401 to $1.641.170 in a year. This cost is not utility function of provider since it does not include incentives (penalty and reward) for the supplier based on to attain the target objectivity of contract in PBL. According to results of the simulation, the supplier will be awarded for the high availability rates under the 200 hr.-500 hr-1000 hr-1 year length of contracts, considering to 90% target availability was determined by the customer in PBL contracts.

Because of the design cost, the total cost of the length of 200 hr.-contract is pretty high while comparing with the cost $36.757 under the R_S (200hr). However, this total cost under the one-year contract decreases from $2.792.245 to $1.641.170. Considering incentives, penalty and reward costs for these contracts, the decline in the total cost will be much higher than this amount. Therefore, provider's total cost will drop with the incentives under the PBL. We can conclude that the length of the contract is the most significant factor for the suppliers to increase the reliability of the systems.

According to results of simulation under the time consideration of target reliability, 500 hr., and one spare capacity (R_T, 500 hr.,1), the graphs of availability and reliability versus time, total service cost versus time are shown in Figure 6 and 7.

Figure 6: Availability and Reliability versus Time (R_T, 500 hr.,1)

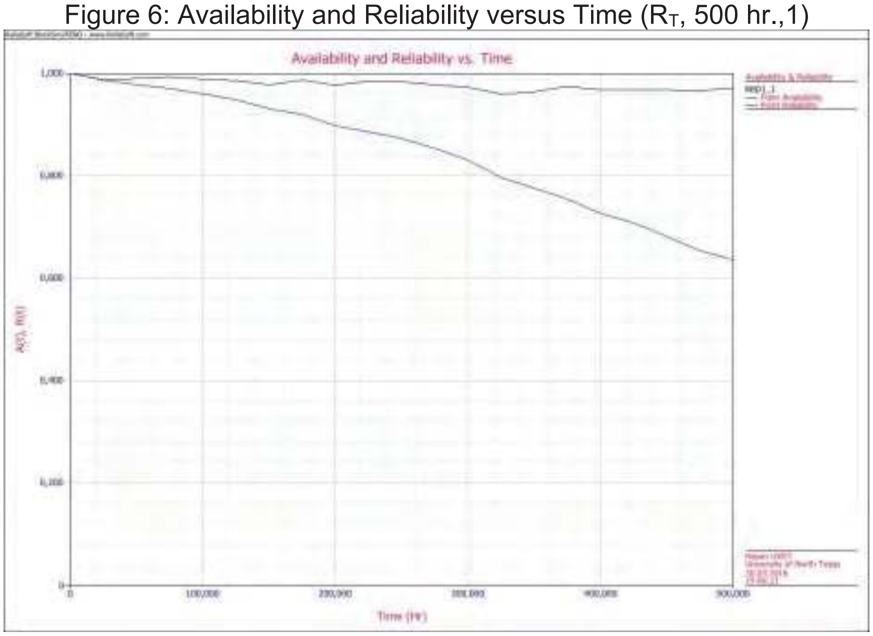
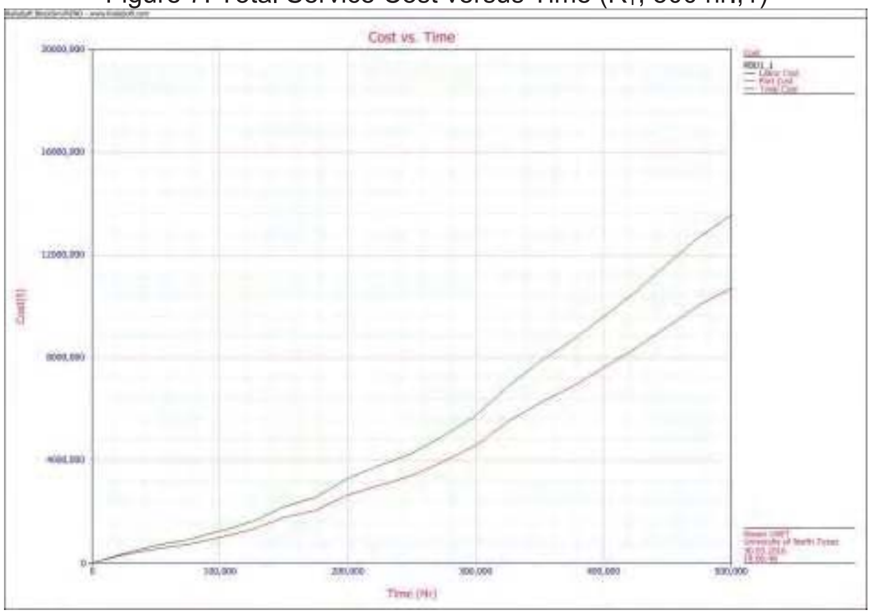


Figure 7: Total Service Cost versus Time (R_T, 500 hr.,1)



The second simulation was conducted with target reliability, after enhancement of reliability of components, under the four various time considerations - 200 hr, 500hr., 1000 hr., one year- with spare acquisition type, fixed probability (0.1) of stock out for the each component of the system. The results of this simulation were presented in Table 10.

Table 10: Simulation results for target reliability (spare acquisition type: fixed probability (0.1) of stock out)

System Overview	t=200 hr, R_T(200hr)	t=500 hr, R_T(200hr)	t=1000 hr, R_T(200hr)	t=1 year, R_T(200hr)
Mean Availability (All Events):	**0.986716**	**0.978143**	**0.966223**	**0.949386**
Reliability(200):	0.898	0.653	0.235	0
Expected Number of Failures:	0.103	0.409	1.221	15.369
MTBF (Total Time) (Hr):	1941.7	1222.5	819	562.1
Number of Failures:	0.103	0.409	1.221	15.369
System Cost Summary				
Misc. Corrective Costs:	$690	$2,836	$8,754	$113,219
Costs for Parts (CM):	$275	$875	$2,650	$37,475
Total CM Costs:	$965	$3,711	$11,404	$150,694
Block Consequential Costs:	$2,837	$10,767	$30,786	$383,302
Pool Holding Costs:	$0	$0	$0	$0
Total Service and Spare Costs:	$3,802	$14,477	$42,190	$533,996
Design Cost	*$668,604*	*$668,604*	*$668,604*	*$668,604*
Total Costs:	**$672,406**	**$683,081**	**$710,794**	**$1,222,599**

The results of simulation show that there is an apparent growth in the availability rates. The availability of the systems decreases with the increase of the length of the contract. While the availability rate was 98.7% under the time consideration of 200 hr., to 94.9% under the one-year contract. From this results, we can conclude that increasing the spare parts inventory has less impact on the availability of highly reliable systems. On the other hand, the total cost of the supplier increases from $672.406 to $1.22.599 in a year. Also, this price does not include incentives (penalty and reward) for the provider based on to attain the target objectivity of contract in PBL. Considering to 90% target availability was determined by the customer in PBL contracts, then we can see the results of the simulation, the supplier will be awarded for the high availability rates under the 200 hr.-500 hr-1000 hr-1 year length of contracts.

CONCLUSION, LIMITATIONS, AND FUTURE RESEARCH

In our study, we investigated the effects of reliability growth in PBL. We analyzed how reliability decision of the supplier was affected with incentives of customers. Additionally, we examined the relationship between spare parts, the length of the contract and reliability growth which are essential decision factors in PBL.

We found that there was strong relationship amongst system design cost, reliability, spare parts stocking, and incentives. The main contribution of this study is the effects of upfront investment for reliability growth in PBL become more predictable for suppliers. In our study, we managed to build the game-theoretical model to analyze the relationship between reliability growth and design cost. We found that customers should have to consider the minimum total design cost of the system for award and penalty amount in contracts to motivate them to make the upfront investment to increase the reliability of systems. Our study shows that meaningful and significant incentives; which is the one of the tenets of PBL tied to arrangements (DoD, 2014; UT,2012); has become more predictable for the customer in our game-theoretical model.

Also, we revealed that increasing the spare parts inventory has less impact on the availability of highly reliable systems. This result also consistent with the Jin et al., (2012)'s and Mirzahosseinian and Piplani (2011)'s findings.

Additionally, the results of simulation show that the availability of the systems reduces with the increase of the length of the contract. Additionally, we found that this decreased rate of the availability in longer contracts was reduced with the high reliability of the systems.

We showed that under a longer agreement in PBL, suppliers are eager to invest in reliability enhancement. While high design cost of the systems increases the total cost of supplier under the short-term length contract, this total cost of provider significantly decreases the whole cost with the long-term contracts.

In our study, not only we put forward the impacts of reliability enhancement on the availability and total costs suppliers but also we showed that how suppliers' decisions about reliability improvement are affected by the contract features, such as length of the contract and incentives.

One of the limitations of this research is the effects of system fleet size did not consider in this study. Suppliers' motivation for reliability enhancement will be affected directly with the fleet size of the system. In the future studies, this impacts of the fleet size to design cost, reliability, spare parts, and availability rates can be analyzed. In future research, besides the linear reward scheme which is applied to the second problem, also, exponential and serial award rate can be used in the game-theoretical model (Nowicki et al. (2008). One of the other limitations of this study is we ignore the manufacturing cost of the new systems with increased reliability. Since the production cost will be affected with reliability enhancement, the impacts of this cost can be examined in the future studies. Also, future research can be done with real data.

Finally, we can conclude that PBL contracts motivate suppliers to make the upfront investment for reliability improvement by powerful incentives, which able to make savings in acquiring and holding spare assets.

References

Agripino, M., Cathcart, T., & Mathaisel, D. (2012). A lean sustainment enterprise model for military systems. Acquisition Review Quarterly, 9.

Bakshi, N., Kim, S. H. (2015). Signaling new product reliability with after-sales service contracts. Management Science, 61(8), 1812-1829.

Berkowitz, D., Gupta, J. N., Simpson, J. T., & McWilliams, J. B. (2005). Defining and implementing performance-based logistics in government. Defense Acquisition Review Journal, 11(3), 255-267.

Bergen, M., Dutta, S. & Walker Jr, O. C. (1992). Agency relationships in marketing: A review of the implications and applications of agency and related theories. The Journal of Marketing, 56(3), 1-24.

Boudreau, M. W., & Naegle, B. R. (2003). Reduction of total ownership cost. Doctoral dissertation, Naval Postgraduate School, Retrieved from http://stinet.dtic.mil/cgibin/GetTRDoc?AD=A419826&Location=U2&doc=GetTRDoc.pdf, February 18.

Buyukgural, F. (2012). A 4-step process evaluation model to assess the success of performance-based logistics contracts. Master's thesis, Air Force Institute of Technology Air University Air Education and Training Command. Retrieved from http://www.dtic.mil/cgibin/GetTRDoc?AD=ADA500475&Location=U2&doc=GetTRDoc.pdf, January 18.

Cohen, M. A. & Netessine, S. (2012). Power by the hour: Can paying only for performance redefine how products are sold and serviced? Working paper, Knowledge@Wharton, 2012.

DAU (2005). Performance Based Logistics - Defense Acquisition University. (n.d.). Retrieved from https://acc.dau.mil/adl/en-US/32536/file/6154/PBL%20Guide%20-%20March%202005.pdf, January 16.

Department of Defense, (2014). Performance Based Logistics (PBL) Guidebook. Retrieved from https://acc.dau.mil/psa, March 18.

Devries, H. J. (2005). Performance-based logistics: Barriers and enablers to effective implementation. Defense Acquisition Review Journal, 11(3), 241-253.

Driessen, M.A., Arts, J.J., Houtum, G.J., Rustenburg, W.D., Huisman, B. (2010). Maintenance spare parts planning and control: A framework for control and agenda for future research. Beta Working Paper series, 325.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. Academy of management review, 14(1), 57-74.

Firchau, V. (1987). Information systems for principal-agent problems. In Agency Theory, Information, and Incentives. Springer Berlin Heidelberg, 81-92.

Gansler, J. S. & Lucyshyn, W. (2006). Evaluation of performance based logistics. Doctoral dissertation, Naval Postgraduate School, Retrieved from http://www.acquistionresearch.org/_files/fy2006/umd-lm-06-040.pdf, March 18.

Gardner, C.P. (2008). Balancing government risks with contractor incentives in performance-based logistics contracts. Master Thesis, Retrieved from http://thecenter.utk.edu/images/Users/1/PBL/WhitePapers/BalancingGovernmentRiskWithContractorIncentivesinPerformance-BasedLogisticsContracts.pdf, March 18.

Geary, S. & Vitasek, K. (2005). It's a war of ideas. Aviation Week & Space Technology, 163(17), 57-62.

Government Accounting Office (GAO) (2005). Army depot maintenance: Ineffective oversight of depot maintenance operations and system, Retrieved from https://www.gao.gov/products/GAO-05-441, March 18.

Guajardo, J. A., Cohen, M. A., Kim, S. H., & Netessine, S. (2012). Impact of performance-based contracting on product reliability: An empirical analysis. Management Science, 58(5), 961-979.

Helander, A., & Möller, K. (2008). System supplier's roles from equipment supplier to performance provider. Journal of Business & Industrial Marketing, 23(8), 577-585.

Hunter, C. (2006). Public and private partnerships and the future of the Army industrial enterprise. Retrieved from http://www.lexingtoninstitute.org/library/resources/documents/ Defense/public privatepartnerships.pdf, March 18.

Hypko, P., Tilebein, M., & Gleich, R. (2010). Benefits and uncertainties of performance-based contracting in manufacturing industries: An agency theory perspective. Journal of Service Management, 21(4), 460-489.

Jin, T., Ding, Y., Guo, H., & Nalajala, N. (2012, July). Managing wind turbine reliability and maintenance via performance-based contract. In 2012 IEEE Power and Energy Society General Meeting, 1-6.

Jin, T., Tian, Z., & Xie, M. (2015). A game-theoretical approach for optimizing maintenance, spares and service capacity in performance contracting. International Journal of Production Economics, 161, 31-43.

Jones, S. R., & Zsidisin, G. A. (2008). Performance implications of product life cycle extension: the case of the a-10 Aircraft. Journal of Business Logistics, 29(2), 189-214.

Keating, S. & Huff, K. (2005). Managing in the new risk supply chain. Engineering Management, 15(1), 24-27.

Kim, S. H., Cohen, M. A., & Netessine, S. (2007). Performance contracting in after-sales service supply chains. Management Science, 53(12), 1843-1858.

Kim, S. H., Cohen, M. A., Netessine, S., & Veeraraghavan, S. (2010). Contracting for infrequent restoration and recovery of mission-critical systems. Management Science, 56(9), 1551-1567.

Kim, S. H., Cohen, M. A., & Netessine, S. (2017). Reliability or inventory? An analysis of performance-based contracts for product support services. In Handbook of Information Exchange in Supply Chain Management, Springer, Cham, 65-88.

Kobren, B. (2009). What Performance-based logistics is and what it is not - and what it can and cannot do. Defense Acquisition Review Journal, 16(3), 254-267.

Kratz, L., & Buckingham, B. A. (2010). Achieving outcomes-based life cycle management. Defense Acquisition Review Journal, 77(1), 45-66.

Kumar, U. D., Nowicki, D., Ramirez-Marquez, J. E., & Verma, D. (2007). A goal programming model for optimizing reliability, maintainability and supportability under performance based logistics. International Journal of Reliability, Quality and Safety Engineering, 14(03), 251-261.

Mahon, D. (2007). Performance-based logistics: Transforming sustainment. Journal of Contract Management, 5, 53-71.

Mettas, A. (2000, January). Reliability allocation and optimization for complex systems. In Annual Reliability and Maintainability Symposium. 2000 Proceedings. International Symposium on Product Quality and Integrity (Cat. No. 00CH37055), 216-221, IEEE.

Mirzahosseinian, H. & Piplani, R. (2011). A study of repairable parts inventory system operating under performance-based contract. European Journal of Operational Research, 214(2), 256-261.

Nowicki, D., Kumar, U. D., Steudel, H. J., & Verma, D. (2008). Spares provisioning under performance-based logistics contract: profit-centric approach. Journal of the Operational Research Society, 59(3), 342-352.

Ng, I. L., & Nudurupati, S. S. (2010). Outcome-based service contracts in the defence industry - Mitigating the challenges. Journal of Service Management, 2(5), 656-674.

Ong, M., Ren, X., Allan, G., Kadirkamanathan, V., Thompson, H. A., & Fleming, P. J. (2005). Decision support system on the grid. International Journal of Knowledge- Based and Intelligent Engineering Systems, 9(4), 315-326.

Pettingill, K., & Knipper, M. A. (2004). Performance-based logistics (PBL) partnerships: Assessment of implementation methodologies for selected ACAT 1 & 2 systems. Retrieved from http://thecenter.utk.edu/images/Users/1/PBL/WhitePapers/AssessmentofImplementation MethodologiesforSelectedACAT1and2Systems.pdf, March 18.

Randall, W. S., Pohlen, T. L., & Hanna, J. B. (2010). Evolving a theory of Performance based Logistics using insights from service dominant logic. Journal of Business Logistics, 31(2), 35-61.

Russell, S. H. (2007). Supply chain management more than integrated logistics. Air Force Journal of Logistics, 29(2), 56-63.

Quick, Robert M. (2011). Performance-Based Logistics: A Military Logistics Model (Doctoral dissertation). Retrieved from ProQuest digital dissertation database. (UMI No. 3467482), March 17.

Sherbrooke, C.C. (1968). METRIC: a multi-echelon technique for recoverable item control. Oper. Res.. 16 (1), 122–141.

Sols, A., Nowick, D., & Verma, D. (2007). Defining the fundamental framework of an effective Performance-based logistics (PBL) contract. Engineering Management Journal, 19(2), 40-50.

The University of Tennessee, and Supply Chain Visions (2012). The Tenets of PBL Second Edition, A Guidebook to the Best Practices Elements in Performance-Based Life Cycle Product Support Management. Retrieved from https://acc.dau.mil/adl/en-US/550412/file/68356/Learning%20Asset%20PBL%20Tenets%20Guidebook%20Second%20Edition%20June%202012%20Final.pdf, February 18.

Vitasek, K., Cothran, J., & Rutner, S. (2007). Best practices of public private partnering: Developing a win-win strategy for working with depots (White paper). Knoxville, TN: University of Tennessee and Supply Chain Visions.

DECISION SCIENCES INSTITUTE

Understanding Individuals' Motivations for Knowledge Sharing in Virtual Communities

Hanieh Javadi Khasraghi
University of Delaware
hanieh@udel.edu

Xuan Wang
University of Texas Rio Grande Valley
Xuan.Wang@utrgv.edu

Yaojie Li
Columbus State University
Li_yaojie@columbusstate.edu

Xinchen Li
Hengyang Normal University
Xinchenli2015@163.com

ABSTRACT

Over the years, Web 2.0 and its successor – Web 3.0 promote and prosper user-generated content, ease of use, interoperability, and virtual communities. Consequently, a growing number of companies start to leverage virtual communities for sharing knowledge, solving problems, and creating value. To that end, understanding members' motivations for participation in virtual communities is the primary purpose of the current research. We found that, "Trust" and "Shared Vision" can extensively affect individuals' knowledge sharing in virtual communities. Also, men and women's knowledge sharing tends to be divergent, in terms of "Norm of Reciprocity" and "Identification."

KEYWORDS: Knowledge Sharing, Virtual Communities, Motivations

INTRODUCTION

Advances in Internet and social network technologies have enabled people to access to information and knowledge easily, in different places, and at different times. Through virtual communities, one can simply pursue knowledge beyond traditional information sources, such as colleagues at work and peers in school. Chiu and colleagues (2006) indicated that virtual communities are providing numerous opportunities for knowledge sharing and social interactions, in particular, for individuals with common interests, goals, or practices. Similarly, knowledge is crucial for companies, who attempt to convert knowledge into competitive advantages. Slow and sluggish companies, often, are deficient in knowledge generation and diffusion among their employees. Many myopic managers immersed themselves and subordinates in relatively static and isolated legacy systems. In contrast, astute executives are likely to build and/or upgrade knowledge-based virtual communities, which allow their companies to tap collective wisdom while addressing enterprise-wide problems (Chiu et al., 2006). One famous example is Enterprise 2.0, a strategic integration of Web 2.0 into a company's networks and business processes, aiming at deep and wide organizational

collaboration and knowledge sharing (cf. McAfee 2006). Community activities can be oriented toward marketing, commerce and education (cf. Teo et al. 2003). Regardless of the theme and fashion of virtual communities, the viability and proliferation of a virtual community depend mainly on the sustained individual participation and knowledge sharing (Nov et al., 2009). Without adequate knowledge exchange, communities are likely to decline and demise (Chiu et al., 2006). As such, it is imperative to understand individuals' motivation for sharing knowledge in order to design and develop an efficient virtual community. One main category of virtual communities is termed *transactional virtual communities* (Sun et al., 2012), where people post questions and compensate other(s) who are able to answer. Also, another type of virtual communities are operated for public good, without financial incentives. We argue that the assumptions that have been used for transactional virtual communities cannot simply be applied in research on free virtual communities. Therefore, there is a need to understand individuals' motivation for participation in free virtual communities. The question of why people share their knowledge in virtual communities has been the subject of many studies (Burke et al., 2009; Butler, 2001; Chiu et al., 2006). They draw on different theories, such as social cognitive theory, social capital theory, social exchange theory, expectancy-value theory, and self-efficacy theory, to investigate the effect of different factors on the knowledge sharing in virtual communities (Chen & Hung, 2010; Chiu et al., 2006; Sun et al. 2012). Our study, hence, aims to integrate the fragmented constructs from prior research in order to get a better understanding of the knowledge sharing process in virtual communities.
The rest of this paper is organized as follows. First, theoretical background on influential factors on virtual community activities are reviewed and hypotheses are thus developed. Second, research methodology and data analysis are illustrated. Third, the implications for theory and practice, and limitations are discussed. Finally, the conclusion and future research directions are discussed.

THEORETICAL BACKGROUND

Motivation and self-determination theories distinguish intrinsic motivation from extrinsic motivation based on different reasons or goals that give rise to an action (Ryan & Deci, 2000). Intrinsic motivation refers to "doing something because it is inherently interesting or enjoyable," whereas extrinsic motivation refers to "doing something because it leads to separable outcome" (Ryan & Deci, 2000). Prior research has illustrated that extrinsic and intrinsic motivations can result in different performance (Ryan & Deci, 2000). Intrinsic motivation contains two distinct components: enjoyment-based intrinsic motivation (i.e., fun) and obligation/community-based intrinsic motivation (i.e., love of community) (Lindenberg, 2001). Most of the times, people perform tasks that they are extrinsically motivated, even if they are not intrinsically motivated (Kaufmann et al. 2011). In crowdsourcing area, extrinsic motivation can be categorized into three components: immediate payoffs (i.e., earning money), delayed payoffs (i.e., improving ones skills), and social motivation (such as building a network with other creative people) (Kaufmann et al., 2011). Researchers who work on participation in online communities and open source software projects also have applied these motivations in their studies (e.g. Lakhani & Wolf, 2003). Most of the researchers have used social cognitive theory, social capital theory, and social exchange theory to construct a model for investigating individuals' motivations for knowledge sharing in virtual communities. Social cognitive theory has been widely used for validating individuals' knowledge sharing behaviors (Compeau & Higgins, 1995a). According to social cognitive theory, contextual factors, personal factors, and behavior have interacting relationships (Wood & Bandura, 1989). Social capital theory refers to "the sum of the actual and potential resources embedded within, available through, and derived from the network of relationship by an individual or social unit" (Nahapiet & Ghoshal, 1998). Nahapiet and Ghoshal (1998) also argued that "the fundamental proposition of the social capital theory is that network ties provide access to resources." Tsai and Gholshal (1998) found that social capital facilitates knowledge sharing in organizations. Researchers in virtual communities also found that social capital stimulates members to share their knowledge (Chiu et al., 2006). Social exchange theory explains human behavior in social exchange and it demands that both producer and user benefit from their interaction (Anderson et al., 1999). According to social exchange theory, the norm of reciprocity motivates individuals to share their knowledge in virtual communities (Hall, 2001). In the following, the factors from the literature are discussed and hypotheses are developed.

Norm of Reciprocity

One of the important motivational factors for individuals' knowledge-sharing in virtual communities is the norm of reciprocity. That means people share their knowledge since they believe that they can receive benefits from others (Chen & Hung, 2010; Chiu et al., 2006). Although obligation of reciprocity between two individuals is difficult, some researchers believe that balanced reciprocity may occur within a group as a whole which is called generalized exchange (Ekeh, 1974). Davernport and Prusak (1998) in their paper stated that norm of reciprocity and trust are significant factors driving knowledge sharing. Wasko and Faraj in their study on three Usenet newsgroups found that people share their knowledge in the online community because of moral obligation that they have to the other members in the group (Wasko & Faraj, 2000). Wang and Fesenmaier in their study on members of a virtual tourism community also found that future reciprocation is a major motivation of individuals' contribution. Lin in his study on 172 employees from 50 organizations found the significance association of reciprocal benefits with employees' knowledge sharing attitudes and intentions (Lin, 2007). Fehr and Gachter in their study differentiated norm of reciprocity from altruism in the way that norm of reciprocity shows a pattern of behavior in which people respond to the actions of others with the similar actions (Fehr & Gächter, 2000). These arguments lead to:

Hypothesis 1: the norm of reciprocity is positively associated with the quality of knowledge sharing.

Trust

Mayer and David (1995) defined Trust as follows:
"The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party."

The importance of trust has been discussed in various disciplines, such as IS group performance (Nelson & Cooprider, 1996), marketing (Geyskens et al., 1998), organizational value creation (Uslaner, 2000), online transactions (Chang et al., 2005), knowledge sharing in virtual communities, communication, leadership, management, marketing (Geyskens et al., 1998), and Information technology artifacts (Vance et al., 2008). In management and sociology literature, trust has been identified in two dimensions: 1) "a business view based on confidence or risk in the predictability of one's expectations" and 2) "a view based on confidence in another's good will" (Smeltzer, 1997). Nahapiet and Ghoshal suggested that trust has significant effect on the willingness of individuals to engage in cooperative interaction (Nahapiet & Ghoshal, 1998). Chiu and colleagues (2006) in their study find that trust has influence on individuals' knowledge sharing behavior. The belief that the other party will not take advantage of the situation is known as Trust (Gefen, Karahanna, & Straub, 2015). When there is no rule to make the members to be sure that others in the community will behave as they are expected, trust will be a substitute to the rules. Nonaka (1994) stressed the importance of trust for creating an atmosphere for knowledge sharing. Many studies have examined the effect of trust on the knowledge sharing of the individuals in virtual communities and identified it as a key element in fostering the level of knowledge sharing in virtual communities (Butler & Cantrell, 1994; Ridings et al., 2002). In a study on 73 members in several knowledge discussion forums hosted by PChome online (www.pchome.com.tw), Chen and Hung (2010) found the significant effect of interpersonal trust on knowledge sharing attitude. These arguments lead to:

Hypothesis 2: trust is positively associated with the quality of knowledge sharing.

Personal Outcome Expectation and Community Related Outcome Expectation

Based on social capital theory, "outcome expectations refer to the expected consequences of one's own behavior" (Bandura, 1997). Positive outcome expectation has been found to incentivize individuals to participate (Bandura, 1997). Members of virtual communities may share knowledge for their self-interest (such as having fun, self-satisfaction, improving performance, and etc.), or they may share knowledge for virtual community rather than self-interest (such as success of virtual community, enriching knowledge in virtual community, and etc.) (Wasko & Faraj, 2000). Therefore, members' outcome expectations can be divided into two dimensions: personal outcome expectations and community related outcome expectations (Hsu et al., 2007).

Enhancement of career opportunities is a personal outcome expectation which has been discussed as motivational factor for individuals' knowledge-sharing and participation in various communities. One of the most cited reasons for individuals' participation in open source communities is the enhancement of career opportunities (Lakhani & Von Hippel, 2003). Also social capital research in the context of organizations shows that collective action of individuals is driven by instrumental motivations, such as career advancement (N. Lin, Ensel, & Vaughn, 1981). Lin (2007) found the significance association of enjoyment in helping others with employees' knowledge sharing attitudes and intentions. Contribution to the community's knowledge and enhancing the community have been found in previous research as a motivation for individuals' knowledge sharing (Bock & Kim, 2001). These arguments lead to:

Hypothesis 3: personal outcome expectations are positively associated with the quality of knowledge sharing.

Community related outcome expectations can be extended to two perspectives. One of the perspectives reflects the values of sharing knowledge in the virtual community, such as how the individuals think they can benefits from the virtual community. The other perspective is the perceived compatibility. Gerrard and Cunningham (2003) defined compatibility as "the degree to which an innovation is perceived as being consistent with existing values, previous experiences, and potential needs, where existing values involve lives style/habit, work attitude/relevance, and concepts in knowledge sharing"(Gerrard et al., 2003), which indicates whether the virtual community shares the benefits that the individual needs. Researchers in various studies on virtual communities found the significant effect of perceived compatibility on individuals' motivation for knowledge sharing (Bock & Kim, 2001; Chiu et al., 2006; Hsu et al., 2007; Wasko & Faraj, 2000), where the fulfillment of perceived compatibility shows the another aspect of community related outcome expectation. These arguments lead to:

Hypothesis 4: community related outcome expectations are positively associated with their quality of knowledge sharing.

Shared Language and Shared Vision

Shared language addresses “the acronyms, subtleties, and underlying assumptions that are staples of day-to-day interactions” (Lesser & Storck, 2001). Shared codes and language cause individuals to understand each other’s goals (Tsai & Ghoshal, 1998). Shared vision “embodies the collective goals and aspirations of the members of an organization” (Tsai & Ghoshal, 1998). Tsai and Ghoshal showed that organizational members who share a vision are more likely to share their knowledge with each other (Tsai & Ghoshal, 1998). These arguments lead to:

Hypothesis 5: Shared vision is positively associated with the quality of knowledge sharing.

Hypothesis 6: Shared language is positively associated with the quantity of knowledge sharing.

Identification

Identification refers to “conception of one’s self in terms of the groups’ defining features”. Bagozzi and Dholakia (2002) found that identification has significant effect on individuals’ participation in virtual communities. Hogg and Abrams (1988) found that identification is a motivational factor for individuals’ engagement in virtual communities. These arguments lead to:

Hypothesis 7: identification is positively associated with the quality of knowledge sharing.

Moderating Effects

Moderating effects on the relationships between variables have attracted many researchers’ interest in online communities and crowdsourcing contests (Nov et al. 2009, Sun et al. 2012). Researchers argue that the contribution to model development will be larger if moderating variables are included in the research model. Previous research on knowledge sharing among bloggers in online social networks indicates that the impact of identification and reciprocity on knowledge sharing varies by gender (Chai et al., 2011). In this study, we would like to identify how gender moderates the effect of independent variables on knowledge sharing. Based on the theoretical background and the hypotheses discussed above, this study establishes a research model which suggests seven primary links and two moderating links for gender involved in understanding the motivations of sharing knowledge in a virtually community. In order to test the hypotheses, Table 1 is represented.

Table 1. Research Hypothesis

		Moderator (Gender)
Dependent Variable	Knowledge Sharing	
Independent Variables	Norm of Reciprocity	Yes
	Trust	No
	Personal Outcome Expectation	No
	Community Outcome Expectation	No
	Shared Language	No
	Shared Vision	No
	Identification	Yes

RESEARCH METHODOLOGY

Setting and Participants

Data were collected through a field survey at a large public university behavioral research lab. The participants were undergraduate students in business school who voluntarily registered to participate in the survey to earn extra credits for their participations. Indeed, virtual community is very popular among college students since they are highly active in online communities, especially in social networks. Therefore, this study focuses on examining the motivations among the young college students. We obtained a total of 100 responses from the students.

Measurement Development

All measurement items where adopted from the related literature. By reviewing previous studies, we figured out that the combination of the items from different studies is complementary and would result in better scale development. Therefore, multiple items from different studies have been used in the survey, and the most reliable measurements have been selected for the analysis process. The independent variable norm of reciprocity is measured by 2 items, trust is measured by 5 items, personal outcome is measured by 13 items, community outcome is measured by 4 items, shared language is measured by 3 items, shared vision is measured by 3 items and identification is measured by 4 items. The dependent variable is measured by the quality of knowledge sharing which includes three measurement items.

Factor Analysis

Exploratory factor analysis (EFA) is used to discover factor structure of the constructs. Latent root criterion is used to decide on the number of factors to extract. Using eigenvalue for establishing cutoff point is reliable for our dataset, since the number of variables is 35 which is within the acceptable range of 20 to 50 (Hair et al., 2013). The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy(MSA) was used to determine the appropriateness of using factor analysis. Overall, MSA value equals 0.744 that meets 0.50 criterion. Individual variables' MSA values are all above 0.50 in the factor matrix which indicates that using factor analysis is appropriate (Hair et al., 1995). Bartlett's Test of Sphericity value (which is approximately Chi-Square) was 2210.213, which is significant at $p = 0.0001$. This test showed that a significant correlation exists among at least some of the variables which indicates the appropriateness of factor analysis solution. Principal Component Analysis of factor analysis was used with VARIMAX rotation to identify the variables associated with the virtual community. To make sure that this is the best factor solution, different methods such as latent root criterion, scree-plot, and parallel analysis are used to extract number of factors. The results showed that latent root criterion gives the number of factors which leads to the constructs with the highest reliability in comparison to the other methods. The factor analysis identified 35 items with factor loadings greater than 0.45. But, we chose 0.7 as the threshold for factor loading since in the literature for the sample size around this study's sample size, the same threshold has been used. Table 2 shows an excessive consistence among items under each factor with all the factor loadings, where 23 items are used in this study.

Table 2. Finalized Indicators

	Component							
	1	2	3	4	5	6	7	8
Norm of Reciprocity_1	0.825							
Norm of Reciprocity_2	0.836							
Trust_1		0.875						
Trust_2		0.845						
Trust_3		0.800						
Trust_4		0.827						
Trust_5		0.851						
Personal Outcome_1			0.803					
Personal Outcome_2			0.792					
Personal Outcome_3			0.838					
Personal Outcome_4			0.763					
Personal Outcome_5			0.724					
Community Outcome_1				0.812				
Community Outcome_2				0.839				
Community Outcome_3				0.802				
Shared language_1					0.812			
Shared language_2					0.905			
Shared Vision_1						0.901		
Shared Vision_2						0.736		
Shared Vision_3						0.762		
Identification_1							0.791	
Identification_2							0.734	
Identification_3							0.812	
Quality of Knowledge sharing_1								0.853
Quality of Knowledge sharing_2								0.790
Quality of Knowledge sharing_3								0.770
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.								

The Exogenous Construct 1 (Norm of reciprocity) includes two variables: “members' helping in return of my help” and “being fair by helping others when they help me”. Construct 2 (Trust) includes five variables: members' “not taking advantage of other members”, “keeping promises”, “truthfulness”, “non-disruptiveness”, and “behaving consistently”. Construct 3 (Personal Expectations) is identified by communication and Joy which includes: “reputation”, “respect”, “strengthening tie between members”, “making friends” and “improve status between members”. Construct 4 (community-related outcome expectation) includes three variables: “enriching community's knowledge”, “successful functioning of community”, and “community's continuous operation”. Construct 5 (shared language) includes two variables: members' “using understandable communication pattern” and “using understandable narrative forms for posting messages”. Construct 6 (shared vision) includes three variables: members' shared vision of “helping others is pleasant”, “helping others solve their problems”, and “learning from each other”. Construct 7 (identification) includes three variables: “sense of belonging to the community”, “positive feeling toward community”, and “feeling proud for being member of the community”. The Endogenous construct (quality of Knowledge sharing) includes three variables: “Knowledge shared by members is accurate”, “Knowledge shared by members is complete” and “knowledge shared by members is reliable”.

Reliability and Validity

Confirmatory factor analysis is employed to test the reliability and validity of the measurement model (Hair et al., 2013). The ratio of number of factors to the number of observations meets the acceptable ratio of 1:10, and the cumulative variance is 76.76% which is satisfied with the standard level 60% (Hair et al., 2013). Reliability focuses on the consistency of the subjects' responses to a set of items. In fact, it means that the items of the instrument measure the same phenomenon. We assessed reliability using construct reliability. As recommended by Nunnally, we use 0.60 as a minimum value to ensure that our measurements are reliable and are free from measurements' error (Nunnally et al., 1967). All of the constructs in this study have construct reliability greater than 0.7 which means that they are reliable. Construct validity is the extent to which a set of measured variables represent the theoretical latent constructs they are designed to measure. Convergent and discriminant validity are terms used to distinguish between two aspects of construct validity. Results showed that the model fits the data well. The correlation matrix between factors also proves discriminant validity because the correlations among the constructs are significantly less than one with all of the correlations are less than 0.001, which indicates that none of the constructs have significant relationship with the others, therefore this test provides evidence for discriminant validity. As Table 3 indicated, all eight constructs have strong convergent validity and discriminant validity. Consequently, well-defined constructs are used in the next stage of structure equation model, and Partial Least Squared Model is used for exploring the results in this study.

Table 3. Confirmatory Factor Analysis for Eight Constructs

Component	Composite Reliability	Cronbach's Alpha	Average Variance Extract
Norm of Reciprocity	0.911	0.805	0.837
Trust	0.948	0.931	0.784
Personal Outcome	0.915	0.883	0.682
Community Outcome	0.906	0.845	0.764
Shared language	0.889	0.749	0.799
Shared Vision	0.912	0.855	0.775
Identification	0.931	0.888	0.817
Quality of Knowledge Sharing	0.902	0.836	0.755

Data Analysis and Results

Table 4 summarizes the results of the hypotheses tests for the dependent variable – Quality of Knowledge Sharing. The results show that Trust and Shared Vision have positive significant effect on quantity of the submissions. Trust means that members of virtual communities who are not taking advantage of other members and behaving consistently and non-disruptiveness for being a member of the community, and shared vision expressed the members whom self believe that members also will learn from each other, are more interested to share contributed knowledge with the other members of the community. Also, “Norm of Reciprocity”, “Personal Outcome”, “Community Outcome”, “Shared Language” and “Identification” do not have significant effect on quality of the knowledge sharing.

Table 4. Results of hypothesis tests with quality of knowledge sharing			
	Coefficient	P Value	VIF
Norm of Reciprocity	0.132	0.119	1.397
Trust	0.249	0.011**	1.827
Personal Outcome	-0.051	0.327	1.842
Community Outcome	0.029	0.401	1.432
Shared language	-0.1	0.187	1.148
Shared Vision	0.315	0.002***	1.531
Identification	0.006	0.478	2.077

Moderator Tests

In this section, we test moderator effect of gender and evaluate whether the effect of norm of reciprocity and identification differ for Men and Women. The results are shown in Table 5. Gender significantly moderates the effect of norm of reciprocity on the quality of knowledge sharing. The results indicate men share high quality knowledge compared to women when there are higher norm of reciprocity and identification, which implies that men and women have different attitudes toward knowledge sharing in virtual communities. Moreover, the results shows that men have more sense of belonging to the virtual community and believe that if they help the other members, the other members also will help them.

Table 5. Gender moderator effect of quality of knowledge sharing			
	Coefficient	P Value	VIF
Norm of Reciprocity	0.092	0.207	1.404
Trust	0.224	0.020**	1.941
Personal Outcome	-0.055	0.314	1.964
Community Outcome	-0.002	0.494	1.586
Shared language	0.095	0.199	1.232
Shared Vision	0.256	0.009***	1.571
Identification	0.011	0.461	2.182
Gender*Reciprocity	-0.155	0.081*	1.235
Gender*Identification	-0.157	0.078*	1.254

DISCUSSION AND IMPLICATIONS

Key findings

There are a couple of intriguing findings in this study. First, individuals' knowledge sharing behavior is positively associated with their attitude toward knowledge sharing. This study supports that "trust" and "shared vision" have significant positive effect on individuals' quality of knowledge sharing. Members who have positive feeling toward the virtual community and feel that other members do not take advantage from each other, are more likely to share their high quality knowledge with other virtual community members. Moreover, members share their knowledge in the virtual community because they believe that members in the virtual community would help them if they need it. The results also shows the moderating effect of gender on the relationship between norm of reciprocity and quality of knowledge sharing and the relationship between identification and quality of knowledge sharing. It shows that women and men have different attitudes toward sharing knowledge in the virtual communities.

Limitations

Before discussing the implications, the limitations of this study should be notified. First, the study's sample is limited to 100 students whose age ranges from 19 to 27, which constitutes a main concern of the current paper. The findings of this study may not be generalizable to all members of virtual communities. For instance, senior citizens or teenagers can behave differently from adults. In order to enhance our external validity, in the future research we should increase the sample size and variety. However, the student sample that we used can well represent the population of virtual community participants. They were "born and raised" in the Internet age and actively and ardently participate in various online activities.

Implications

The empirical results of our study show that "Trust" and "Shared Vision" have significant effect on individuals' quality knowledge sharing behaviors. It indicates that members who have a strong sense of belonging to the virtual community and who feel that they can be helped as long as they help others, are likely to share their knowledge in the community. Also, the results show that men and women have different attitudes toward knowledge sharing in virtual communities. Women are more willing to share their knowledge in virtual communities because they enjoy from participation and feel happy while helping each other. On the other hand, men have the feeling of belonging to the community and because of that they like to participate and share their knowledge. For practical applications, the findings of this study will help the online community sponsors to understand individuals' motivations and enhance online community platforms to motivate more people to participate in these communities and share high-quality knowledge.

Conclusion

Prior studies have found various factors that can motivate participants to share knowledge with others in virtual communities. Different constructs have been developed to measure individuals' motivations. In this study, we adopted items from various studies in prior literature and used them in designing our questionnaires that are distributed to the students. The results show that Trust and Shared Vision have significant positive effect on the quality of knowledge sharing in virtual communities. Also, gender moderates the impact of "identification" and "norm of reciprocity" on the "quality of knowledge sharing" in the virtual communities. It indicates that, individuals with sense of belonging tend to share more knowledge in the virtual community. Also, Women perform differently from men such that they like to share their knowledge in virtual community because of enjoyment they receive from helping others. Instead, men who have stronger sense of belonging to the virtual community are more willing to help others.

REFERENCE

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411-423.

Anderson, W. T., Challagalla, G. N., & McFarland, R. G. (1999). Anatomy of exchange. *Journal of Marketing Theory and Practice*, 7(4), 8–19.

Bagozzi, R. P., & Dholakia, U. M. (2002). Intentional social action in virtual communities. *Journal of Interactive Marketing*, *16*(2), 2–21.

Bandura, A. (1995). *Self-efficacy in changing societies*. Cambridge university press. Bandura, A. (1997). *Self-efficacy: The exercise of control*.

Beach, L. R., & Mitchell, T. R. (1990). Image theory-a behavioral-theory of decision-making in organizations. *Research in Organizational Behavior*, *12*, 1–41.

Bock, G.-W., & Kim, Y.-G. (2001). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. *PACIS 2001 Proceedings*, 78.

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology*, *130*(3), 353–363.

Burke, M., Marlow, C., & Lento, T. (2009). Feed Me : Motivating Newcomer Contribution in Social Network Sites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2009.*

Butler, B. S. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research*, *12*(4), 346–362.

Butler, Jr., J. K., & Cantrell, R. S. (1994). Communication factors and trust: an exploratory study. *Psychological Reports*, *74*(1), 33–34.

Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. *Information & Management*, *42*(4), 543–559.

Chen, C.-J., & Hung, S.-W. (2010). To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities. *Information & Management*, *47*(4), 226–236.

Chiu, C.-M., Hsu, M.-H., & Wang, E. T. G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, *42*(3), 1872–1888.

Compeau, D. R., & Higgins, C. A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research*, *6*(2), 118–143.

Compeau, D. R., & Higgins, C. A. (1995b). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189–211.

Davenport, T., and Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press.

Easley, R. F., Devaraj, S., & Crant, J. M. (2003). Relating collaborative technology use to teamwork quality and performance: An empirical analysis. *Journal of Management Information Systems,* 19(4), 247-265.

Ekeh, P. P. (1974). *Social exchange theory: The two traditions*. Heinemann London.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The*

Journal of Economic Perspectives, 14(3), 159–181.

Gefen, D., Karahanna, E., & Straub, D. W. (2015). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, *27*(1), 51–90.

Gerrard, P., & Barton Cunningham, J. (2003). The diffusion of Internet banking among Singapore consumers. *International Journal of Bank Marketing*, *21*(1), 16–28.

Geyskens, I., Steenkamp, J.-B. E. M., & Kumar, N. (1998). Generalizations about trust in marketing channel relationships using meta-analysis. *International Journal of Research in Marketing*, *15*(3), 223–248.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate Data Analysis*. Pearson Education.

Hall, H. (2001). Social Exchanges for knowledge exchange paper presented at managing knowledge: Conversations and critiques. University of Leicester Management Centre, 10-11.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes.* Taylor & Frances.

Hsu, M.-H., & Chiu, C.-M. (2004). Internet self-efficacy and electronic service acceptance. *Decision Support Systems*, *38*(3), 369–381.

Hsu, M.-H., Ju, T. L., Yen, C.-H., & Chang, C.-M. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. *International Journal of Human-Computer Studies*, *65*(2), 153–169.

Johnson, R. D., & Marakas, G. M. (2000). Research report: the role of behavioral modeling in computer skills acquisition: toward refinement of the model. *Information Systems Research*, *11*(4), 402–417.

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker motivation in crowdsourcing–a study on mechanical turk. *Proceedings of the Seventeenth Americas Conference on Information Systems*, 1–11.

Kim, M., Sharman, R., Cook-Cottone, C. P., Rao, H. R., & Upadhyaya, S. J. (2012). Assessing roles of people, technology and structure in emergency management systems: a public sector perspective. *Behaviour & Information Technology*, *31*(12), 1147–1160.

Lakhani, K. R., & Von Hippel, E. (2003). How open source software works:“free” user-to-user assistance. *Research Policy*, *32*(6), 923–943.

Lakhani, K., & Wolf, R. (2003). Why hackers do what they do: Understanding motivation and effort in free/open source software projects, 1–27.

Lam, J., & Lee, M. (2005). Bridging the digital divide-The roles of Internet self-efficacy towards learning computer and the Internet among elderly in Hong Kong, China. *HICSS'05. Proceedings of the 38th Annual Hawaii International Conference* on *System Sciences*.

Lesser, E. L., & Storck, J. (2001). Communities of practice and organizational performance. *IBM Systems Journal*, *40*(4), 831–841.

Lin, H.-F. (2007). Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of Information Science*, *33*(2), 135–149.

Lin, N., Ensel, W. M., & Vaughn, J. C. (1981). Social resources and strength of ties: Structural factors in occupational status attainment. *American Sociological Review*, 46(4), 393–405.

Lindenberg, S. (2001). Intrinsic Motivation in a New Light. *Kyklos*, *54*(2), 317–342.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-year odyssey. *American psychologist,* 57(9), 705.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

McAfee, A. P. 2006. Enterprise 2.0: The Dawn of Emergent Collaboration," *Enterprise*, 2, 15-26. Mun, Y. Y., & Davis, F. D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research*, *14*(2), 146–169.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review*, *23*(2), 242–266.

Nelson, K. M., & Cooprider, J. G. (1996). The contribution of shared knowledge to IS group performance. *MIS Quarterly*, 20(4), 409–432.

Nonaka, I. (1991). The knowledge-creating company. *Harvard Business Review*, *69*(6), 96–104. Nov, O., Naaman, M., & Ye, C. (2009). Analysis of participation in an online photo-sharing community: A multidimensional perspective. *Journal of the American Society for Information Science and Technology*, 61(3), 555-566.

Nunnally, J. C., Bernstein, I. H., & Berge, J. M. F. ten. (1967). *Psychometric theory* (Vol. 226). McGraw-Hill New York.

Ridings, C. M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *The Journal of Strategic Information Systems*, *11*(3-4), 271–295.

Ryan, R., & Deci, E. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67.

Smeltzer, L. R. (1997). The Meaning and Origin of Trust in Buyer-Supplier Relationships. *International Journal of Purchasing and Materials Management*, *33*(4), 40–48.

Sun, Y., Fang, Y., & Lim, K. H. (2012). Understanding sustained participation in transactional virtual communities. *Decision Support Systems*, *53*(1), 12–22.

Tajfel, H., & Turner, J. C. (2004). The Social Identity Theory of Intergroup Behavior. *Organizational Identity: A reader,* 56-65.

Teo, H.-H., Chan, H.-C., Wei, K.-K., & Zhang, Z. (2003). Evaluating information accessibility and community adaptivity features for sustaining virtual learning communities. *International Journal of Human-Computer Studies*, *59*(5), 671–697.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: An emperical study of intra-firm networks. The *Academy of Management Journal*, 41(*4)*, 464-476.

Uslaner, E. M. (2000). Social capital and the net. *Communications of the ACM*, *43*(12), 60–64.

Vance, A., Elie-Dit-Cosaque, C., & Straub, D. W. (2008). Examining trust in information technology artifacts: the effects of system quality and culture. *Journal of Management*

Information Systems, *24*(4), 73–100.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Wasko, M., & Faraj, S. (2000). “It is what one does”: why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, *9*(2-3), 155–173.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review*, *14*(3), 361–384.

DECISION SCIENCES INSTITUTE

Eco-Labeling Program Design and Adoption to Improve Environmental and Profit Performance

Yu Amy Xia[a*], Shilei Yang[b], Xu Hartling[c]
[a]College of William and Mary, Williamsburg, VA, USA. amy.xia@mason.wm.edu
[b]Southwest University of Finance and Economics, Chendu, China, syang@swufe.edu.cn
[c]Salem State University, Sale, MA, USA. xu.hartling@salemstate.edu

ABSTRACT

Eco-labels not only help consumers take environmental concerns into purchasing decisions, but also may help companies increase profit performance through enhanced environmental quality. This research studies the dynamics in the establishment and adoption of an eco-labeling program through a Stackelburg game between the third-party committee who designs the eco-labeling program and the company who adopts the program. Managerial insights are generated. We find that with complete information, a well-designed single-level eco-labeling program is sufficient, a sophisticated multi-level eco-labeling program is not necessary. Numerical tests are conducted to further investigate how the market responses to an eco-labeling program.

KEYWORDS: Eco-Labeling Program, Environmental Quality, Sustainability, Corporate Social Responsibility

INTRODUCTION

This paper studies the design and adoption of the eco-labeling program in an industry. An eco-labeling program commonly includes one or multiple levels of eco-standard in an industry. Products or services with environmental quality level exceeds certain standard level will be granted the respective eco-label, which demonstrates the product's performance in environmental quality to consumers, and in return influences consumers' purchasing decisions.

Eco-labeling programs have a wide range of applications in construction, agriculture, fashion, tourism, and consumer products industries (Burnett, 2007; Loureiro, McCluskey, & Mittelhammer, 2001; Nimon & Beghin, 1999; Font, 2002; Tang et al, 2004). Eco-labels such as "Green Stickers" and "Energy Star" have been very popular in food industry and consumer products. Take the Energy Star program for example. It was first established by the U.S. Environmental Protection Agency (EPA) in 1992. Since then Energy Star has successfully identified and promoted energy-efficient products and has been recognized as an umbrella of voluntary programs (Banerjee & Solomon, 2003). From 1993 to 2012, it is estimated that families and businesses have prevented 1,903 million metric tons of greenhouse gas (GHG) emissions and saved more than $239 billion on utility bills while using Energy Star products (Energy Star, 2014).

Unlike enforced regulations and policies, such as carbon taxes and emission trading programs, eco-labeling programs are voluntary programs. It acts as a form of sustainability measurement directed at consumers and intends to make it easy for consumers to take environmental concerns into account when making a purchasing decision. Through its influence on consumers, it performs as a "soft" policy instrument, which motivates companies to voluntarily adopt higher eco-standard in their products and services (Tang et al, 2004). The wide range of applications and the success stories of eco-labels prove the flexible, effective and less costly nature of this voluntary program. A properly designed eco-labeling program will likely attract more environmentally cautious companies to participate in the program; it will also demonstrate the authoritativeness and trustfulness to consumers to choose eco-labeling products. Ultimately, the consumer behavior will lead to higher market penetration of products with eco-labels and higher environmental performance of the industries.

An eco-labeling program is usually designed and examined by non-governmental organizations (NGOs), government offices, independent quality control organizations, or a combined force of the above. In this paper, we refer to the program designer as the third-party committee or the third-party. In order to make the eco-labeling program adoptable in the industry, the third-party often consults some major companies in the industry when making decisions in the design process. The third-party has the authority to design the eco-labeling program; the major companies have the resource, knowledge, and technology in leading the adoption of the eco-labeling program and implementing the eco-performance improvement in the industry. For the third-party, its goal is to promote eco-products and to improve the overall eco-performance of the industry. For a major company, its goal is to enhance company financial performance and to take corporate social responsibility in an affordable way.

The purpose of this paper is to analyze the third-party's decision making in designing the eco-labeling program and the major industry company's decision making in adopting the eco-labeling program. Under different situations of the cost structure, pricing strategy, consumer eco-awareness, and consumer distribution, the goal of the third-party and the goal of the major companies sometimes integrate together and sometimes set apart from each other. In this research, we aim to understand how companies respond to an eco-labeling program in determining the pricing strategy and the eco-performance level of its products, and how should the third party design an eco-labeling program incorporating the expectations of the company's decisions. We also attempt to quantify how consumers evaluate the products with different eco-labels and make their purchasing decision based on the price and eco-labels of the products. Game Theory is used to analyze and understand the complication and dynamics in the design and adoption of an eco-labeling program. Our analysis as well as the managerial insights generated from the research will help both the third-party and the major company in setting, compromising, and approaching their goals in the process. The ultimate goal of this research is to build sustainable supply chain through the eco-labeling programs. In particular, we aim to address the following research questions:

- How is the company eco-performance and its pricing strategy influenced by the consumer response to an eco-labeling program?

- How do the third-party and the major company interact through the design and adoption of the eco-labeling program?

- How should the third-party and the major company act pre-actively and dynamically in pursuing their goals given their different knowledge level in predicting the cost of technological advancement and consumer distribution?

The remainder of this paper is organized as follows. We review the related literature in Section 2. Then we introduce our modeling framework and main findings in Section 3 and Section 4 respectively. Numerical studies are conducted in Section 5. Finally, we conclude the paper with implications and limits of our research in Section 6.

LITERATURE REVIEW

In recent years, eco-labeling programs have attracted increasing attention among scholars. Many research projects in eco-labeling focus on various applications of eco-labeling programs and provide solid base to testify the application of eco-labeling programs in wide range of industries (Hartling, 2018; Yang, 2013; Burnett, 2007; Tang et al, 2004; Font, 2002; Loureiro et al, 2001). Their work justifies the importance and the implementation value of this research. In this section, we provide a brief review of the related literature.

The Implementation Value of Eco-Labeling Programs

Consumers in general prefer environmental quality in products and services. Positive environmental quality information of a brand tends to yield higher product evaluation and increased brand choice (Cho et al, 2017). However, observing environmental quality of a product and correctly evaluating its value is not an easy task for regular consumers. Unlike product quality attributes such as appearance, flavor, and functionality, which are usually revealed either pre-purchase or soon after purchase, environmental consequences of production processes and utility consumption of products are generally difficult to observe at purchase. Therefore, to promote eco-friendly products in the market, and to let consumers induce market changes by “voting” for green products with their shopping dollars (Dauvergne & Lister, 2010), an eco-friendly product needs to clearly demonstrate its environmental quality level. Eco-labeling programs provide consumers a way to easily collect and analyze such information and make their purchasing decisions accordingly.

Increased market demand is an essential motivation for companies to produce eco-friendly products. Higher demand of greener products will create a strong financial incentive for companies to improve their products’ eco-level in order to capture new market initiatives, increase their market shares, and perhaps generate higher profits on their eco-labeled products (Dauvergne & Lister, 2010). Eco-labeling programs provide opportunities for companies to demonstrate their products’ eco-quality to attract market attention, and in return eco-labels stimulate the eco-quality comparison and improvement effort in the industries (Xia et al, 2015).

A company’s products’ environmental performance is also positively correlated with its corporate social responsibility performance, which is a form of business behavior that leads companies to voluntarily contribute to a better society and a cleaner environment (Mir & Shah, 2018). Economic performance, environmental performance and social responsibility performance are three interrelated dimensions of corporate sustainability performance. Therefore, to protect the environment and fulfill their social responsibilities, and even to enhance their economic performance, companies are motivated to produce and redesign more environmentally friendly products. The voluntary adoption of the eco-label programs and the labels placed on their products not only highlight the product’s environmental quality, but also demonstrate the company’s effort in taking its social responsibility. Production of eco-labeled products can eventually be one of the best ways of pursuing competitive strategy (Orsato, 2006; Roy & Vezina, 2001).

From the third-party’s perspective, eco-labels inform consumers about the environmental impacts of the product’s manufacturing, usage, and disposal processes, thus may change consumer purchasing behaviors (Teisl et al, 2002). A significant change in consumer behavior may increase the demand of green products with eco-labels and motivate a company to adopt the eco-labels, which differentiate the company from its competitors and market its products with an environmental characteristic. An increasing supply of these green products may enhance consumer purchase simply through greater availability without even changes in individual awareness, and consequentially reduce ecological footprint of the products in the whole industry.

An Overview of Eco-Labeling Systems and Eco-Labeling Design

Eco-labels have been observed in the market for many years, although using eco-labels to promote carbon footprint reduction or sustainability is a more recent phenomenon (Banerjee & Solomon, 2003). While eco-labeling alone will not improve environment, experience has shown that a properly designed eco-labeling mechanism can be a significant stimulus for market transformation towards environmentally preferable products. Moreover, designing and introducing an appropriate eco-labeling program brings strategic benefits to companies, industries and society in general as we discussed in the earlier section.

Eco-labels are of different types. Salzman (1997) describes three types of eco-labels: single-issue voluntary labels, single-issue mandatory labels, and third-party voluntary labels. Single-issue voluntary labels, such as “Recyclable” mark, are self-claimed by manufacturers and rarely verifiable and accurate. Single-issue mandatory labels, such as “flammable” and “eco-toxic”, are a much smaller class and are required by law and by many governments. Third-party voluntary labels, which are observed as an environmental seal of approval, present the overall environmental quality of the products (Salzman, 1997). Fet and Skaar (2006) further categorize the third-party voluntary eco-labels into three sub-types based on the ISO 14000-series: Type I programs are multiple-criteria third-party programs awarding labels claiming environmental preferability. Type II programs contain a lot of self-declared environmental information on products. They are required to take into account of life cycle considerations. Type III programs need to conduct a life cycle assessment (LCA) of the product following the ISO 14040-standards. They consist of quantified environmental data on significant impacts of the product. The information should enable comparisons between products fulfilling the same function. Other requirements of type III programs include the LCA approval and the third-party verification of the declaration (Fet & Skaar, 2006). In this paper, we focus our discussion on the Type III eco-labeling programs of the third-party voluntary labels, so that we can compare product environmental quality on one criterion that set by a third-party. Note that earlier research has shown that eco-labeling system created by a third-party is more likely to be environmentally beneficial than the eco-labeling system created by the industry itself (Fischer & Lyon, 2014). Future study can be extended on more complicated eco-labels and eco-label programs.

The general procedure of designing a Type III eco-labeling program can be described as follows. The third-party committee determines the category of products to be covered and the specific objective standard that products must meet to obtain the label or labels. Given the standard, companies then determine if they would voluntarily participate in the program by submitting their products for evaluation. If a product meets the predetermined standard, the company pays a licensing fee and places the eco-label on the qualified products (Salzman, 1997). While the eco-labeling process seems straight-forward, there are two foremost design issues need to be addressed beforehand: whether to setup one level of label or multiple levels of labels in the program and what should be the proper standard level or levels.

Blengini and Shields (2010) point out that single-level eco-label is a binary indicator and it serves as an easy and clear standard for both the consumers and the manufacturers, and thus possesses marketing and managerial advantage. The concern of having one eco-label is that a product either meets the criteria (thus gets a label) or not. Therefore, receiving the eco-label only represents reaching the minimum predetermined eco-standard; it does not show the relative environmental performance among the products who receive the label. While in a multi-level eco-labeling program, products with various environmental quality performances obtain different levels of eco-labels, which inform the consumers more details about their relative green quality levels. On the other hand, multi-level eco-labeling system is usually more complicated in design and adoption, and it may cause confusion among consumers. Bleda and Valente (2009) develop a simulation model to show that competition drives the market toward cleaner technological patterns in the presence of multi-level eco-labels. Fischer and Lyon (2016) model competing eco-labels sponsored by an industry trade association and an environmental group. They find multi-level labeling system more attractive in a market with many producers and high production cost gap between different quality levels. Heys and Martin (2017) discover that the entries and competition of the third-party committee may affect the setting of eco-labels. Rysman et al, (2018) study the adoption of a multi-level environmental building certification system. They simulated the impact of switching from the three-level eco-label system (with four difference tiers) to the single-level eco-label system (with high-tier and low-tier), and found single-level system more efficiently stimulate environmental investment of low-tier builders to reach to the high-tier; however, the overall investment of all builders is higher in multi-level system.

Standard levels should be very cautiously set. The standard level or levels should effectively and efficiently promote carbon footprint reduction in the industry and proficiently transfer eco-labeling adoption into economic benefits for the companies that adopt the eco-labeling program. Too high the standards may eliminate participation from the majority of the companies due to the high cost of achieving such standards. However, if a standard is set too low and too easy to reach, the labeling program may lose credibility among consumers and reduce or even eliminate the value for companies to adopt the eco-label system (Harbaugh et al, 2011), and will not be able to truly promote carbon footprint reduction.

A successful eco-labeling program should bring economic benefit to the adopting companies, so that company's ability to become eco-friendly can be sustainably translated to their willingness to make the efforts. Till now, it remains a controversy whether environmental and financial goals are compatible or conflicting. Sammer & Wüstenhagen (2006) report empirical data on how the consumer behavior for household appliances is influence by eco-labels. Houe & Grabot (2009) suggest the label users a method in checking a product' compliance of an eco-label. Rysman et al (2018) show that vertical differentiation in adoption of building certification eco-label system generates economic benefits for builders and motivates them to invest in reaching higher eco-levels. Research works in sustainable supply frame work link company eco-performance with its economic performance to reach the goal of sustainability (Xia et al, 2015; Brandenburg, Govindan et al, 2014; Reuter, Foerstl et al, 2010). These research results provide support and evidences to transfer eco-labeling adoption into economic benefits, the idea we embrace in our model setup to evaluate an eco-labeling system.

In this research, we start with modeling analysis of the general multi-level eco-labeling system, seek to understand the game between the third-party committee who design the labeling program and the company who adopts the program. From the third-party's perspective, we examine the design of the labeling program on improve the eco-quality of the industry. From a company's perspective, we inspect its decisions on adopting various eco-quality levels and determining pricing strategy of its product/service according to the labeling program. Our analysis generates some interesting managerial insights in understanding the eco-labeling systems and improving the functionality of the system.

ECO-LABELING MODELING FRAMEWORK

We study a supply chain with a major company that provides a product and many consumers who make their purchase decisions based on both the price and the eco-quality level (such as water assumption for washing machine and gas-mileage for trucks in transportation) of the product. A third-party committee design the eco-labeling program for the industry of product. The company determines whether to adopt the eco-labeling program or not, and the eco-quality level to reach if it adopts the program. The decisions are made based on the analysis of its own cost structure and estimation of the market response. In the market, consumers observe the label on the product to evaluate the product's eco-quality level and make their independent purchasing decisions. For the third-party, its goal is to improve the overall eco-performance of the industry. For the major company, its goal is to enhance its own economic performance and to take affordable company social responsibility. For the consumers, they want good eco-quality product with affordable price. The third-party has the authority to design the eco-labeling program; the major company has the resource, knowledge and technology to adopt the eco-labeling program and implement the eco-performance improvement of the industry; the consumers make their purchasing decisions. The three groups all target at different goals, however their decisions influence each other consequently. In this paper, we focus on the game between the third-party committee and the major company. Consumers' aggregative decision is described as the overall market response to the eco-quality price and the price set by the company.

Without loss of generality, we assume the current mandatory industry eco-quality standard to be $L_0 > 0$ and normalize the highest eco-quality level reachable in the market as "1". We assume that the company can choose its product/service to reach any eco quality level q in between the mandatory and highest reachable standards ($L_0 < q \leq 1$). In order to reach eco-quality q, it endures a unit production cost $C = f(q)$. Without loss of generality, we assume $f(q)$ to be a convex increase function of q, e.g., $f'(q) > 0, f''(q) \geq 0$, since it is common knowledge that quality improvement at higher quality level tends to cost more than quality improvement at lower level (Craig, 2012). The third-party designs the eco-labeling program to includes n levels of eco-standards, each set to be L_i ($i = 1, \dots, n$ and $L_i < L_{i+1}$). We denote the designed labeling program as $\Psi' = (L_1, \dots, L_n)$. Incorporating the market mandatory eco-quality level L_0, we define the labeling program as $\Psi = (L_0, L_1, \dots, L_n)$. Clearly, the company should set its product service eco-quality level to be at-the-levels instead of in-between-levels as in-between-levels only adds additional cost, wouldn't be recognized by the eco-labeling program.

A cooperative advertising model is used to describe the market response to the eco-labels from consumers' perspective. Cooperative advertising model is well accepted in marketing and operations literature when market is influenced simultaneously by brand quality recognition and product price (Yue et al, 2006; Huang et al, 2002; Li et al, 1996). Here, the eco-label acts as the recognition of the product's eco-quality level. Therefore, the market share of the company when it chooses eco-level L_i and unit product price p can be described as

$$S(q = L_i, p|\Psi) = \left(1 - bL_i{}^{-\delta}\right)(p)^{-\theta}, i = 0,1, \dots n. \quad (1)$$

In the above equation (1), without loss of generality, we unify the potential market demand as "1". $b\ (0 < b < {L_0}^{\delta})$ is a market scaling parameter for the quality effect. Since we assume that products with eco-quality level lower than L_0 is not acceptable by the market. $1 - b$ represents the highest market potential of the company. We further define two positive constants δ and θ as the quasi-quality elasticity of demand and the quasi-price elasticity of demand respectively. To ensure the market demand sufficiently responsible to eco-quality and price changes, we assume both $\delta > 1$ and $\theta > 1$.

We now investigate the Stackelberg Game of the eco-labeling system with the third party setting up the labels and the company deciding whether to adopt the labeling program consequently. We first assume that the company has adopted level $i\ (i = 0, 1, .. n)$ of the ecolabeling system; $i = 0$ represents the scenario that the company maintains its eco-quality level to the minimum requirement of the market and does not adopt the eco-labeling system. The company's expected gross profit at label level i is then

$$\Pi_i(q = L_i, p_i\ |\Psi) = \left(1 - b{L_i}^{-\delta}\right)(p_i)^{(-\theta)}(p_i - f(L_i)),\ \ i = 0,1, \dots n. \tag{2}$$

Therefore, if the company adopts an eco-quality level, the following lemmas help up understand the setting and adoption of the quality levels. All proofs of this paper are included in the Appendix of this paper.

Lemma 1. At a desired enhanced eco-quality level L_i, the company would set its enhanced product price to

$$p_i = \frac{\theta}{(\theta-1)} f(L_i),\ \ i = 1, \dots, n. \tag{3}$$

and maximize its profit at the enhanced quality eco-level as

$$\Pi_i = \left(1 - b{L_i}^{-\delta}\right)\frac{1}{\theta-1}\left(\frac{\theta-1}{\theta}\right)^{\theta} f(L_i)^{-(\theta-1)},\ \ i = 1, \dots, n. \blacksquare \tag{4}$$

Lemma 2. The profit function of the company is a concave function of the eco-quality level. The maximum of Π_i is reached at $\hat{L}$, which satisfy the condition that $\left(\hat{L}^{\delta} - b\right)\frac{f'(\hat{L})\hat{L}}{f(\hat{L})} = \frac{\delta b}{\theta-1}$ $\blacksquare$

It is important to know that the profit function $\Pi\ (L)$ has only one optimal solution $\hat{L}$. Any eco-quality levels $L_i > \hat{L}$ will result in a lower profit than the profit at $\hat{L}$, are thus not attractive to the company and won't be approached by the company. Those eco-levels ($L_i > \hat{L}$) are infeasible quality levels for the company.

MAIN FINDINGS

The timing of our model frame can be explicated as a Stackelberg Game. Step 1, the third-party committee designs the eco-labeling program, hence set L_i, $i = 1, \dots, n$. Step 2, the major company chooses its product/service eco-quality level ($q^* \in \{L_i,\ i = 0, 1, \dots, n\}$) and price (p^*) of the product/service. Step 3, the market responds to the company's decision. In this section, to analyze the decision process and identify the equilibriums with the backward induction.

We start with the market's response, which is specified by the cooperative advertising model as stated in equation (1) of this paper. Assuming that the company choose eco-label level L_i, its profit is thus estimated by equation (2). Based on the profit function, the company should choose its pricing strategy as in equation (3). The company would not choose any higher eco-level which generates lower profit for it. Therefore, $\Pi_i > \Pi_{i-1}$ and $\Pi_i > \Pi_{i+1}$. The third-party committee has enhanced the eco-quality of the company by $L_i - L_0$, which is defined as the as program benefit $B(\Psi)$. The following Lemmas describe the choices of the eco-quality levels.

Lemma 3. The feasible quality levels of the labeling programs satisfy the condition that $L_i \leq \hat{L}$. ■

In other words, for an eco-labeling system with $L_k \leq \hat{L} < L_{k+1}$, only the eco-levels L_i, $(i \leq k)$ are feasible. Levels L_i $(i \geq k+1)$ are unnecessary quality levels and possess no practical value to the system. In addition, in the special case that $k = 0$, the labeling program does not engage the company to participate in the program. Therefore, the improvement of the eco-performance cannot be achieved, the program is thus a failed eco-labeling program. The analysis can be organized in the following Lemma.

Lemma 4. The maximum program benefit is $B^*(\Psi) = \hat{L} - L_0$. If $L_1 > \hat{L}$, the eco-labeling program fails to promote eco-quality improvement in the system. ■

Clearly, how should the third-party design the labeling program is determined by its knowledge of $\hat{L}$. If the third-party has the complete information of the company's cost structure $f(L)$ as well as the market parameters (b, δ and θ), we can easily find the following Proposition that describes the equilibrium of the game in terms of the pricing strategy and profit function for the company and the setting of labeling program for the third-party.

Proposition 1. With complete information shared in the labeling system, the third party should design a one level eco-quality labeling as $\Psi: (L_1^* = \hat{L})$. The company should choose to enhance product eco-quality level to $\hat{L}$ and set its optimal price as $p^* = \frac{\theta}{(\theta-1)} f(\hat{L})$, and generate the highest profit as

$$\Pi^* = \left(1 - b\hat{L}^{-\delta}\right)\frac{1}{\theta-1}\left(\frac{\theta-1}{\theta}\right)^{\theta} f\left(\hat{L}\right)^{-(\theta-1)}. \blacksquare \tag{5}$$

Corollary 1. A single level of labeling program is sufficient if complete information is shared in the labeling system. ■

Corollary 2. A labeling program are not functional and should not be introduced if $\hat{L} < L_0$. ■

To better understand the labeling program, we now investigate the case that production cost of the company is quadratic with respect to eco-quality level $f(L) = \alpha L^2$. A quadratic cost function is well accepted in literature. αL^2 is a convex increase function of L since quality improvement at higher quality level tends to cost more than quality improvement at lower level (Craig, 2012).

Corollary 3. With the quadratic production cost function, the third-party committee should set the optimal quality level as $L_1{}^* = \hat{L} = e^{\frac{1}{\delta}ln\left[b+\frac{\delta b}{2(\theta-1)}\right]}$*. As long as* $L_0 < L_1{}^*$*, the labeling program can improve quality in the industry, and a multi-level labeling mechanism is not necessary.* ■

Clearly, if $L_0 > L_1{}^*$, the company's profit drops if it adopts eco-program. A labeling program based on the voluntary adoption of the company is not feasible to promote higher eco-quality in the market. Corollary 3 shows that the optimal quality level $L_1{}^*$ decreases monotonically with the quasi-price elasticity θ. If consumers are very sensitive to price, the company should adopt a relatively low eco-quality level. On the other hand, $\frac{\partial L_1{}^*}{\partial \delta} = e^{\frac{1}{\delta}ln\left[b+\frac{\delta b}{2(\theta-1)}\right]} \times \left[-\frac{1}{\delta^2} ln\left[b + \frac{\delta b}{2(\theta-1)}\right] + \frac{1}{\delta} \cdot \frac{1}{2(\theta-1)+\delta}\right]$. Higher consumer environmental awareness may or may not leads to a higher eco-quality level equilibrium of the system. Due to the compacity of the market responses, we use numerical test to investigate the market as in Section 5. Notably, the quality function coefficient α possess no impact on the quality level decision.

Corollary 4. Assuming the company holds quadratic production cost function, the third-party committee doesn't need to know the cost function coefficient to set the labeling program. ■

The above result is very important since operational details are usually considered business secret of the company. It is very difficult and many times impossible for the third-party to collect the related information and make an accurate estimation. On the other hand, market survey to understand δ, θ, and b are relatively easier task for the third-party before it designs or adjusts the labeling program.

When the third-party can only guess a wide range of $\hat{L}$, e.g., $\hat{L} \in (\underline{L}, \overline{L})$, a multi-level of labeling program Ψ: $(L_1 \leq \underline{L}, \ldots, L_n \geq \overline{L})$ should be introduced in order to promote higher eco-quality product and capture the optimal labeling level. Depending on the information, the design of the multi-level labeling program requires further in-depth investigations.

From the company's perspective, the labeling program provide an opportunity for it to further enhance its profitability through eco-quality improvement. Its pricing strategy at any of the quality level L_i can be simply executed by the Equation (3) as long as it understands its own cost structure $f(L_i)$. It can also understand the market through the Third-party's design of the labeling program.

NUMERICAL STUDIES AND MANAGERIAL INSIGHTS

In this section, we use numerical examples to investigate the market response with respect to different parameters, and gain further insight into the company's eco-quality level desicison. We test the following base parameter set: $b = 1/10$, $\alpha = 2$, $\delta = 1.5$, and $\theta = 1.8$. All the following results are obtained by modifying the base set by one or two parameters at a time.

Earlier analysis tells that that the optimal quality level $L_1{}^*$ set by the third-party is a monotonic function with the price elasticity θ. However, the profitability of the major company may not be monotonic with θ or δ. The group of charts in Figure 1 demonstrate that the maximal profit decreases with δ, but interestingly the maximal profit decreases first with θ and then increases with it. So the company may be able to exploit the consumer's price sensitivity by adopting the labeling system designed by the third-party. Furthermore, the 3D figure reals that consumers' environmental awareness (shown by δ) have a larger impact on the company's profit when the consumers are more price sensitive.

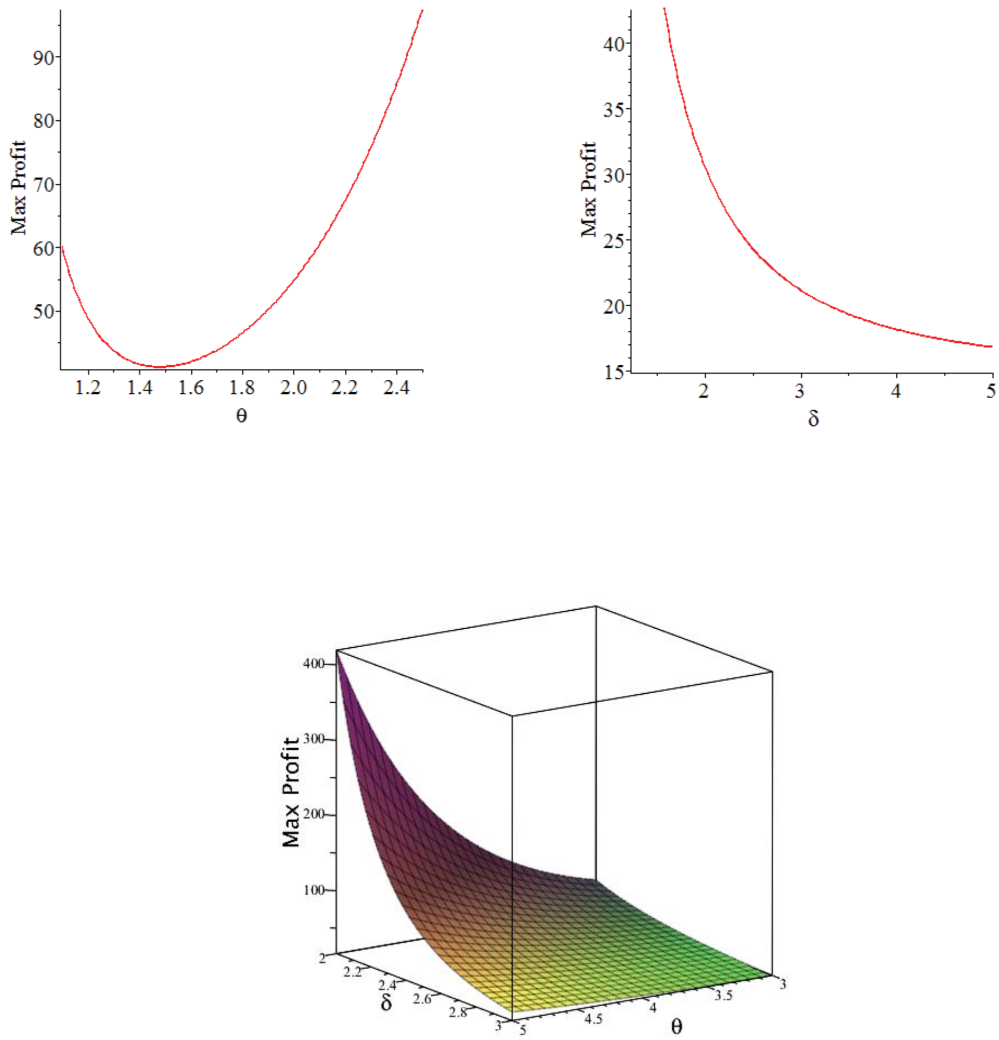


Figure 1 Relationship Between Maximal Profit and Two Elasticity Parameters

Because consumers are reacting more strongly to environmental concerns, their environmental awareness may increase over time. Figure 2 illustrates how such a trend on consumers' behaviors changes optimal decision of the third-party. As δ increases, the third-party should apparently design a higher eco-quality level. Therefore, if considering the evolving of consumers' environmental awareness, designing a multi-level labeling system might be a wise future-oriented solution for the third-party committee.

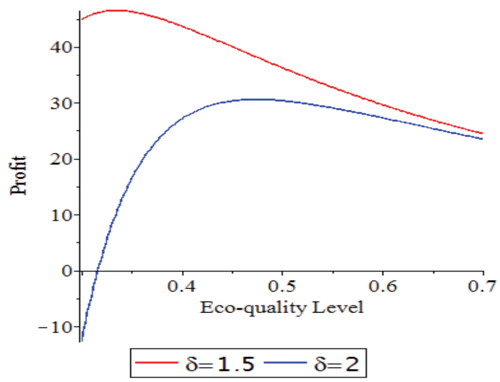


Figure 2 Impact of Consumers' Environmental Awareness

CONCLUSIONS AND FUTURE STUDIES

In this paper, we examine the strategies on how to design an eco-label program. We proposed the pricing strategy if a company wants to achieve certain eco-quality. We also provide the third-party committee with the insight on how to design the eco-labeling program to motivate the companies to enhance eco-quality. In addition, we also investigate how the market parameters influence the design and adoption of eco-labeling program. In particular, if consumers are very sensitive to price, a relatively low eco-quality level will be adopted by the company; on the other hand, if consumers' environmental awareness increases, a higher eco-quality level could lead to more profit for the company. Although multi-level eco-labeling programs are established to promote green quality, we find that it is not necessary to include in the program design and a one-level eco-labeling is enough if complete information is shared in system. This finding justifies the business practice that the majority of eco-labels only contain one level.

However, this paper has its limitations. First, we only consider one major company in the system to focus on the adoption of the eco-labeling program. In reality, it is quite common that there exist several companies competing together. Therefore, the competition among the companies and the impact of the competition deserves further investigation. In addition, because service quality is also a significant factor when consumers make choices, analyzing eco-labeling program design in together with service quality will certainly enrich our understanding on consumers' choices.

REFERENCES:

Amacher, G., Koskela, E., & Ollikainen, M. (2004). Environmental quality competition and eco-labeling. Journal of Environmental Economics and Management, 47, 284–306.

Banerjee, A., & Solomon, B. D. (2003). Eco-labeling for energy efficiency and sustainability: a meta-evaluation of US programs. Energy Policy, 31, 109–123.

Bjorner, T., Hansen, L., & Russell, C. (2004). Environmental labeling and consumers' choice—an empirical analysis of the effect of the Nordic Swan. Journal of Environmental Economics and Management, 47, 411–434.

Bleda, M., & Valente, M. (2009). Graded eco-labels: A demand-oriented approach to reduce pollution. Technological Forecasting & Social Change, 76, 512–524.

Blengini, G., & Shields, D. (2010). Green labels and sustainability reporting - Overview of the building products supply chain in Italy. Management of Environmental Quality: An International Journal, 21(4), 477-493.

Brandenburg, M., Govindan, K., Sarkis, J., & Seuring, S. (2014). Quantitative models for sustainable supply chain management: Developments and directions. European Journal of Operational Research, 233(2), 299-312.

Bratt, C., Hallstedt, S., Robèrt, K.-H., Broman, G., & Oldmark, J. (2011). Assessment of eco-labelling criteria development from a strategic sustainability perspective. Journal of Cleaner Production, 19, 1631-1638.

Brécard, D., Hlaimi, B., Lucas, S., Perraudeau, Y., & Salladarré, F. (2009). Determinants of Demand for Green Products: An application to eco-label demand for fish in Europe. Ecological Economics, 69(1), 115–125.

Burnett, J. (2007). City buildings—Eco-labels and shades of green! Landscape and Urban Planning, 8, 29-38.

Cho , Y., Soster , R., & Burton, S. (2017). Enhancing Environmentally Conscious Consumption through Standardized Sustainability Information . Journal of Consumer Affairs.

Craig, A. (2012). Measuring Supply Chain Carbon Efficiency: A Carbon Label Framework. Massachusetts Institute of Technology. Cambridge, MA: Massachusetts Institute of Technology.

Dauvergne , P., & Lister, J. (2010). The Prospects and Limitsof Eco-Consumerism:Shopping Our Way toLess Deforestation? Organization & Environment, 23(2), 132–154.

Energy Star. (2014). Retrieved from https://www.energystar.gov/about/

Fet, A., & Skaar, C. (2006). Eco-labeling, Product Category Rules and Certification Procedures Based on ISO 14025 Requirements. Int J LCA, 11(1).

Fischer , C., & Lyon, T. P. (2016, 1). A Theory of Multi-Tier Ecolabel Competition. Ross School of Business Paper No. 1319.

Fischer, C., & Lyon, T. P. (2014). Competing Environmental Labels. Journal of Economics and Management Strategy, 692-716.

Font, X. (2002). Environmental certification in tourism and hospitality: progress, process and prospects. Tourism Management, 23, 197–205.

Gallastegui, I. (2002). The use of eco-labels: a review of the literature. European Environment, 12(6), 316-331.

Global Eco-labeling Network. (2015).

Harbaugh, R., Maxwell, J. W., & Roussillon, B. (2011). Label Confusion: The Groucho Effect of Uncertain Standards . Management Science, 57(9), 1512–1527.

Hartling, X. (2018). Design a Successful Program for Voluntary Carbon Footprint Reduction. International Review of Advances in Business, Management and Law, 1(1).

Heyes, A., & Martin, S. (2017). Social Labeling by Competing NGOs: A Model with Multiple Issues and Entry. Management Science, 63(6), 1800-1813.

Houe, R., & Grabot, B. (2009). Assessing the compliance of a product with an eco-label: From standards to constraints. International Journal of Production Economics, 121(1), 21-38.

Huang, Z., Li, S., & Mahajan, V. (2002). An analysis of manufacturer–retailer supply chain coordination in cooperative advertising. Decision Sciences, 33, 469–494.

IPCC. (2007). IPCC fourth assessment report: Climate change 2007. Intergovernmental Panel on Climate Change.

Karl, H., & Orwat, C. (1999). ENVIRONMENTAL LABELLING IN EUROPE: EUROPEAN AND NATIONAL TASKS. European Environment, 9, 212-220.

Li, S., Huang, Z., & Ashley, A. (1996). Inventory, channel coordination and bargaining in a manufacturer–retailer system. Annals of Operations Research, 68, 47–60.

Loureiro, M. L., McCluskey, J. J., & Mittelhammer, R. C. (2001). Assessing Consumer Preferences for Organic, Eco-labeled, and Regular Apples. Journal ofAgricultural and Resource Economics, 26(2), 404-416.

Mera, A., & Palacios, T. (2004). Factors that determine the adoption of an eco-label as a marketing tool. esic market.

Mir, U. A., & Shah, F. A. (2018). Impact of Corporate Social Responsibility on Corporate Financial Performance: A Study Of The Consumer Goods Industry of India. Amity Global Business Review, 13(1), 50-59.

Nilsson, H., Tuncer, B., & Thidell, A. (2004). The use of eco-labeling like initiatives on food products to promote quality assurance—is there enough credibility? Journal of Cleaner Production, 12, 517–526.

Nimon, W., & Beghin, J. (1999). Are eco-labels valuable? Evidence from the apparel industry . American Journal of Agricultural Economics, 81, 801-811.

OECD . (2001). Experts Workshop on Information and Consumer Decision-Making For Sustainable Consumption.

OECD. (2002). Towards Sustainable Household Consumption? Trends and Policies in OECD Countries.

OECD. (2005). Eco-Labelling Schemes: Compilation of Recent Studies.

Orsato, R. (2006). Competitive Environmental Strategies: WHEN DOES IT PAY TO BE GREEN? 48(2), 127-143.

Reuter, C., Foerstl, K., Hartmann, E., & Blome, C. (2010). Sustainable global supplier management: The role of dynamic capabilities in achieving competitive advantage. Journal of Supply Chain Management, 46(2), 45-63.

Roy, M.-J., & Vezina, R. (2001). Environmental Performance as a Basis for Competitive Strategy: Opportunities and Threats. Corporate Environmental Strategy, 8(4), 339-347.

Rysman, M., Simcoe, T., & Wang, Y. (2018). Differentiation Strategies in the Adoption of Environmental Standard: LEED from 2000-2014. Working paper.

Salzman, J. (1997). Informing the Green Consumer - The Debate Over the Use and Abuse of Environmental Labels. Journal of Industrial Ecology, 1(2), 11-21.

Sammer, K., & Wüstenhagen, R. (2006). The influence of eco-labelling on consumer behaviour–Results of a discrete choice analysis for washing machines. The influence of eco-labelling on consumer behaviour–Results of a discretBusiness Strategy and the Environment, 15(3), pp.185-199.

Tang, E., Fryxell, G. E., & Chow, C. S. (2004). Visual and Verbal Communication in the Design of Eco-Label for Green Consumer Products. Journal of International Consumer Marketing, 16(4), 85-105.

Teisl, M. F., Roe, B., & Hicks, R. L. (2002). Can Eco-Labels Tune a Market? Evidence from Dolphin-Safe Labeling. Journal of Environmental Economics and Management, 43.

Truffer, B., Markard, J, J., & Wustenhagen, W. (2001). Eco-labeling of electricity—strategies and tradeoffs in the definition of environmental standards. Energy Policy, 29(11), 885–897.

United Nations. (1987). Kyoto Protocol to the United Nations Framework Convention on Climate Change.

Xia, Y., Zu, X., & Shi, C. (2015). A profit-driven approach to building a “people-responsible” supply chain. European Journal of Operational Research, 241(2), 348-360.

Yang, X. (2013). Green Policies and Green Businesses. Proceedings of the “Suzhou-Silicon Valley-Beijing” 2013 International Innovation Conference on Technology Innovation and Diasporas in a Global Era. Suzhou.

Yue, J., Austin, J., Wang, M., & Huang, Z. (2006). Coordination of cooperative advertising in a two-level supply chain when manufacturer offers discount. European Journal of Operational Research, 168(1), 65-85 .

APPENDIX : SOME PROOFS

Proof of Lemma 1:

$$\Pi_i(q = L_i, p_i \,|\Psi) = \left(1 - bL_i{}^{-\delta}\right)(p_i)^{(-\theta)}(p_i - f(L_i)),$$

At $q = L_i$, calculate

$$\frac{\partial \Pi}{\partial p} = \left(1 - bL_i{}^{-\delta}\right) \cdot \left[-\theta p^{-\theta-1}(p - f(L_i)) + \, p^{-\theta}\right]$$
$$= \left(1 - bL_i{}^{-\delta}\right) \cdot \left[(1-\theta)p^{-\theta} + \, \theta p^{-\theta-1} f(L_i)\right].$$

Let $\frac{\partial \Pi}{\partial p} = 0$, therefore, $p_i = \frac{\theta}{(\theta-1)} f(L_i)$

Then, substitute it into the profit function, the optimal profit at the voluntary product price is

$$\Pi_i = \left(1 - bL_i{}^{-\delta}\right)\left(\frac{\theta}{\theta-1} f(L_i)\right)^{-\theta}\left[\frac{\theta}{(\theta-1)} f(L_i) - f(L_i)\right]$$
$$= \left(1 - bL_i{}^{-\delta}\right)\left(\frac{\theta}{\theta-1} f(L_i)\right)^{-\theta} \frac{1}{(\theta-1)} f(L_i)$$
$$= \left(1 - bL_i{}^{-\delta}\right)\frac{1}{\theta-1}\left(\frac{\theta-1}{\theta}\right)^{\theta} f(L_i)^{-(\theta-1)}$$

Q.E.D.

Proof of Lemma 2

$$\frac{\partial \Pi_i}{\partial L_i} = \frac{1}{\theta-1}\left(\frac{\theta-1}{\theta}\right)^{\theta}\left[-(\theta-1)\left(1 - bL_i{}^{-\delta}\right)f'(L_i) + \delta b L_i{}^{-(\delta+1)} f(L_i)\right] f(L_i)^{-\theta}$$

Since $f'(L_i) > 0$, $f''(L_i) \geq 0$, we can easily find the condition for $\frac{\partial \Pi_i}{\partial L_i} = 0$ and prove that $\frac{\partial^2 \Pi_i}{\partial L_i{}^2} > 0$.

Q.E.D.

DECISION SCIENCES INSTITUTE

Consequences of Sustainability Orientation in Service Firms

Sandeep B. Jagani
Illinois State University
Email: SBJagan@ilstu.edu

ABSTRACT

Recently, internal and external sustainability drivers have increasingly mandated firms to improve their engagement with the two previously less exploited dimensions of sustainability, social and environmental dimensions. Sustainability orientation, an evolving strategic intent, has enabled firms to develop blueprint of policies and procedures of corporate sustainability for various business functions in their organizations. Manufacturing sector has received fair attention from the researchers because of their reputation for emissions of greenhouse gasses, polluting natural water sources and callous social practices across the globe. However, there has been very little research on implications of sustainability orientation on business practices and sustainability outcomes in service sector. Unlike, tangible goods, services need to be produced at the place of its consumption. Process alterations and tradeoffs, that a service system design makes to accommodate sustainability practices, will be noticed by the customers in service system unlike a manufacturing system. Customers can develop mixed perception about the company or the brand in their minds. This research investigates the role of sustainability orientation of service firms and its impact on the customers' readiness to adhere to the sustainability initiatives. The research tries to identify the sustainability outcomes in the form of social, environment and economic goals that are realized based as a result of sustainability orientation. As future research agenda, this paper suggests to conduct a survey of senior practitioners in service sector of United State of America using a comprehensive questionnaire on sustainability strategies, practices and outcomes designed from the thorough literature review to find empirical evidence of the propositions.

KEYWORDS: Sustainability orientation, environmental performance, social performance, sustainability implementation, sustainability outcomes, consumer perceptions

INTRODUCTION AND LITERATURE REVIEW

Increasingly, companies in manufacturing and service sector are positioning themselves as advocates of sustainability by designing strategies that enables business functions to use stainability practices (Hong, Kwon, & Roh, 2009; Kleindorfer, Singhal, & Wassenhove, 2005; Linton, Klassen, & Jayaraman, 2007; Wu & Pagell, 2011). Strategic intent for sustainability (sustainability orientation) is the result of the all the stakeholders' pressure (government compliances, customer preferences, labor unions' requirements, order qualifying criteria, etc.) on the firms to deliver products and services in more environment friendly operations, considering the frugal use of natural resources and becoming more socially responsible while guaranteeing profits. Sustainability orientation is translated to sustainability practices that transforms the traditional way of operations to realize the perceived sustainability outcomes.

Over past three decades, manufacturing sector has incorporated sustainability initiatives in their operations leading to noticeable sustainable outcomes in the terms of economic, social and environmental benefits (Chen et al., 2017; Hong, Jagani, Kim, & Youn, 2019; Tachizawa & Wong, 2015). However, there is a prominent difference in incorporating sustainability practices in manufacturing firms and that in service firms. This is because frontend activities (activities that happen in front of customers) and backend activities (activities that happen without the presence of customer) are very distinct in manufacturing and service operations. Modification of a process design in manufacturing firm to include a sustainability initiative will be less eminent to a customer or user of its product unless that initiative has modified the product design. The same is rarely true for a service firm. Services are different from tangible goods in many ways. Production of services requires involvement of customers. Services cannot be stored for later use. Services are produced at the location where they are delivered. The production and consumption of services occur simultaneously. Thus, the incorporation of sustainability initiatives in service form will alter both, the service system design as well as the service offering. Changes in service operations for sustainability can impact the customer perception about the company and their likelihood of making a repeat purchase.

The review of literature suggests that several researchers have attempted to study sustainability of service firms. In service sector, there are more conceptual papers and case studies on sustainability. Very few empirical studies in the field emerged from the literature review. Three of the most comprehensive studies on sustainability in service sector are briefed here. Hussain et al., (2016) reveals a framework for supply chain sustainability in service industry. However, (1) they look at the sustainability from a firms supply chain perspective not operations perspective. (2) They do not consider the service delivery process and service offerings inseparable. Finally, (3) They concentrate more on sustainability orientation (strategic intent of sustainability) and fail to address the practices related to sustainability that mediated the relationship between strategic intent and outcomes. Goodman (2000) uses case study of Scandic Hotels in Nordic Europe to find that environmental information systems, employee training programs and collaboration with suppliers are the mechanisms that enabled this hotel chain to be leader in sustainable operations. This study concentrates specifically on hotel industry and discusses service operations limited to sample size of one organization. Though this study identifies few mechanisms through which company implements sustainability initiative, it fails to discuss implementation of sustainability on customer perceptions. Lastly, Rothenberg (2007) argues that more manufacturing businesses are moving toward including services in their portfolio using the service dominant logic (Vargo & Lusch, 2008). She presents a series of case studies of manufacturers who have realized sustainability benefits on moving to the service sector. Though she argues about the practices these firms use to be more sustainable, her sample size is limited and her arguments are difficult to generalize across different industries in service sector. In summary, despite an exhaustive attempt by previous researchers, there are certain research gaps that will enable service firms to understand consumer behavior as a result of sustainability orientation and sustainability implementation.

RESEARCH MODEL AND PROPOSITIONS

This research intends to answer three questions related to sustainability orientation, sustainability implementation, sustainability outcomes and consumer perceptions.

1. What is the role of sustainability orientation in service firms in translation of generic service operations into sustainable service operations?

2. How the sustainable operations translate to measurable outcomes in service firms?

3. How customers' perception about service firm changes as the firm moves towards more sustainable practices?

Sustainability orientation and sustainability implementation is service firms

According to Hamel & Prahalad (2005), companies gain competitive advantage by clearly defining a roadmap or a blue print of new functionalities and acquisition of competencies. Sustainability orientation is one such type of strategic intent to redefine existing service process and products to attain sustainability goals. When companies develop clear policies for their economic intents, environmental intents and societal intents, it is actually building orientation on sustainability. Sustainability of a firm is the combination of how well a firm performs in three dimension; economic, environmental and social (Dyllick & Hockerts, 2002; Markard, Raven, & Truffer, 2012). However, given the nature of the service system design, considering that there is involvement of customers in production of services and heterogeneity in delivering services, implementation of sustainability action plans has its own challenges compared to manufacturing systems. Nevertheless, action plans of sustainability implementation are the result of the sustainability polices and standard operating procedures defines in sustainability orientation. Therefore, I propose:

P1. Sustainability orientation will positively influence sustainability influence sustainability implementation.

Epstein & Buhovac (2010) argues that effective implement sustainability action plans need the formal and informal mechanisms. But, most companies do not possess both mechanisms. However, sustainability orientation is a phenomena to build such mechanisms. Measureable performance outcomes enable companies to quantify their efforts onto success criteria. Though the perishable nature of service products offer a new challenge in measuring sustainability outcomes, specifics sustainability practices should reveal several sustainability outcomes. Therefore, action plans of earning high revenues, conservation of reserves, being less detrimental or environment, and inclusiveness of society will enable companies to achieve sustainability goals. Thus, I propose:

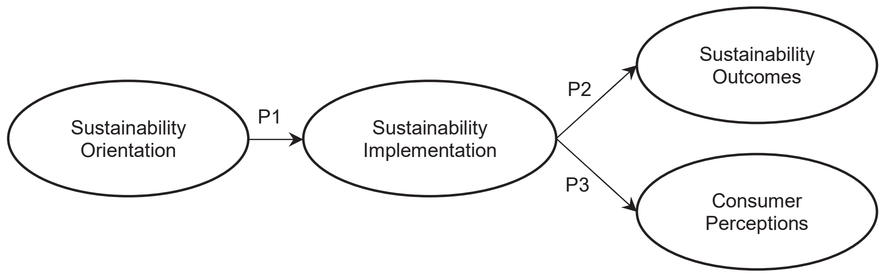
P2. Sustainability implementation will positively influence sustainability outcomes

Unlike manufacturing firms, service firms have a unique challenges. Services need to be produces at the place of consumption. Additionally, in service process design and product design are strongly tried together and are always done simultaneously. Implementation of sustainability orientation will require service firms to build and incorporate policies and standard operating procedures in their existing process and product design. This may require the companies to alter their existing processes and consequently change the specifications of the service offerings. This may result in change of consumer perception about the product and about the company. Consumer perception is a positive or a negative feeling associated with a brand or product (Carpenter & Nakamoto, 1989; McDonald & Oates, 2006). Altering a service process or service offering in order to incorporate sustainability practices may cultivate mixed consumer perceptions. However, in a controlled experiment of manufactured products, Choi & Ng, (2011) demonstrate that consumers favor sustainability in both dimensions by giving positive evaluations of the company and purchase intent. Therefore, I propose:

P3. Sustainability implementation will create positive consumer perception.

The research model is as shown in the below Figure 1 along with the three propositions.

Figure 1: Conceptual framework for the research



FUTURE SCOPE OF RESEARCH

To find empirical evidence of the relationships among the constructs in Figure 1, there is a need to conduct a survey of US firms of service sector. First, one should aim to develop valid and reliable measurement instrument for the constructs and sub-constructs used in the proposed conceptual framework in Figure 1. Next, using with Dillman (2000) and Huang (2012) methods, use four step approach to investigate the relationship: item generation, pre-test and structured interviews, pilot study using Q-sort, and finally launch large-scale survey and conduct data analysis.

IMPLICATIONS AND CONCLUSION

Sustainability outcomes and consumer perceptions are the key goals companies expect from the exercises of designing a sustainability strategy and implementing the action plan of sustainability practices. Linking consumer perception to the sustainability practices in a service production process of which consumers are co-producers will enable companies to decide the intensity of sustainability practices. Consequently, the results of the empirical investigation will allow companies to discover the tolerance of consumers' acceptance of alteration of services system design to incorporate sustainability criteria. This study attempts to fill the gap in the literature by focusing on service systems and linking strategic intent of sustainability to the measureable outcomes and consumer perceptions.

REFERENCES:

Carpenter, G. S., & Nakamoto, K. (1989). Consumer Preference Formation and Pioneering Advantage. *Journal of Marketing Research*, *26*, 285–298.

Chen, L., Zhao, X., Tang, O., Price, L., Zhang, S., & Zhu, W. (2017). Supply chain collaboration for sustainability: A literature review and future research agenda. *International Journal of Production Economics*, *194*, 73–87. https://doi.org/10.1016/j.ijpe.2017.04.005

Choi, S., & Ng, A. (2011). Environmental and Economic Dimensions of Sustainability and Price Effects on Consumer Responses. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-011-0908-8

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (Vol. 2). Wiley New York.

Dyllick, T., & Hockerts, K. (2002). Beyond the business case for corporate sustainability. *Business Strategy and the Environment*, *11*(2), 130–141. https://doi.org/10.1002/bse.323

Epstein, M. J., & Buhovac, A. R. (2010). Solving the sustainability implementation challenge. *Organizational Dynamics*. https://doi.org/10.1016/j.orgdyn.2010.07.003

Goodman, A. (2000). Implementing sustainability in service operations at Scandic Hotels. *Interfaces*. https://doi.org/10.1287/inte.30.3.202.11653

Hamel, G., & Prahalad, C. K. (2005). Strategic intent. *Harvard Business Review*. https://doi.org/Article

Hong, P., Jagani, S., Kim, J., & Youn, S. H. (2019). Managing sustainability orientation: An empirical investigation of manufacturing firms. *International Journal of Production Economics*, *211*, 71–81. https://doi.org/10.1016/J.IJPE.2019.01.035

Hong, P., Kwon, H.-B., & Roh, J. J. (2009). Implementation of strategic green orientation in supply chain: An empirical study of manufacturing firms. *European Journal of Innovation Management*, *12*(4), 512–532. https://doi.org/10.1108/14601060910996945

Huang, C. (2012). *Developing Circular Economy Capability: Antecedents, Mechanisms, and Outcomes in Chinese Manufacturing Industry* (Theses and). University of Toledo.

Hussain, M., Khan, M., & Al-Aomar, R. (2016). A framework for supply chain sustainability in service industry with Confirmatory Factor Analysis. *Renewable and Sustainable Energy Reviews*. https://doi.org/10.1016/j.rser.2015.07.097

Kleindorfer, P. R., Singhal, K., & Wassenhove, L. N. (2005). Sustainable Operations Management. *Production and Operations Management*, *14*(4), 482–492. https://doi.org/10.1111/j.1937-5956.2005.tb00235.x

Linton, J., Klassen, R., & Jayaraman, V. (2007). Sustainable supply chains: An introduction. *Journal of Operations Management*, *25*(6), 1075–1082. https://doi.org/10.1016/j.jom.2007.01.012

Markard, J., Raven, R., & Truffer, B. (2012). Sustainability transitions: An emerging field of research and its prospects. *Research Policy*, *41*(6), 955–967. https://doi.org/10.1016/j.respol.2012.02.013

McDonald, S., & Oates, C. J. (2006). Sustainability: Consumer perceptions and marketing strategies. *Business Strategy and the Environment*. https://doi.org/10.1002/bse.524

Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, *2*, 192–220.

Rothenberg, S. (2007). Sustainability Through Servicizing. *MIT Sloan*.

Tachizawa, E. M., & Wong, C. Y. (2015). The Performance of Green Supply Chain Management Governance Mechanisms: A Supply Network and Complexity Perspective. *Journal of Supply Chain Management*, *51*(3), 18–32. https://doi.org/10.1111/jscm.12072

Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-007-0069-6

Wu, Z., & Pagell, M. (2011). Balancing priorities: Decision-making in sustainable supply chain management. *Journal of Operations Management*, *29*(6), 577–590. https://doi.org/10.1016/j.jom.2010.10.001

DECISION SCIENCES INSTITUTE

A Student Engaged Project for Undergraduate and Graduate Introductory Financial/Managerial Accounting Courses

Judith A. Sage
Sage & Sage
Email: lloydsage@aol.com

Lloyd G. Sage
Sage & Sage
Email: lloydsage@aol.com

ABSTRACT

A student engaged Project was examined in introductory financial/managerial accounting for a midwest undergraduate course [MW]) and a southwest graduate course [SW]. This Project included many of the Pathways Commission (2015) Learning Objectives and AICPA Framework Core Competencies (2017/1999). The Project involved teamwork, written reports, oral presentations, and peer evaluations. The students strongly agreed/agreed (mean scores of 4.75 [MW] and 4.89 [SW]) that this Project increased understanding of selected financial/managerial accounting topics. It appears that this Project can be successfully used for teaching these topics in undergraduate and graduate courses.

KEYWORDS: Engaged project, Introductory financial/managerial accounting, Alternative teaching method

INTRODUCTION

Various committees for more than three decades have called for change in the accounting curriculum and teaching approaches. Recently, the Association of International Certified Professional Accountants discussed the skills needed by managerial accountants in the *CGMA Competency Framework: 2019 Update* (2019). Earlier, the *AICPA Pre-Certification Core Competency Framework* (2017) advocated for a skills-based curriculum for accounting students (public/industry/government/not-for-profit). Our research investigated a project that engaged students in activities to assist them in experiencing several needed skills.

Previously, the Pathways Commission (2012) indicated that accounting educators have responsibilities that comprise both curricular dimension (what we teach) and pedagogical scope (how we teach). Also, this Commission (2012) has endorsed, under the pedagogical scope pillar, approaches or learning experiences that are designed to help students in thinking, performing, and making decisions comparable to accounting professionals. Our research study examined a learning method that should assist students to think and make accounting decisions.

Apostolou *et. al.* (2013) advocated that accounting research should be conducted to identify approaches and/or methods to teach or learn core professional competencies (e.g., communication skills). In 2014 the Pathways Commission recommended that quality education should encompass the use of a variety of teaching techniques that stimulate students to enhance their ability to utilize professional judgment in decision-making. Also, Bandura and Lyons (2012) proposed that instructors should incorporate different learning approaches (e.g., lecture, problem-solving, or engaged learning) in their courses. In addition, Wessels (2010) indicated that the instructor's primary responsibility is to motivate students to use learning activities that will accomplish the desired learning course outcomes. This research paper explored whether a project that engaged the students was a good method of teaching selected financial and managerial topics to undergraduate and graduate students enrolled in introductory financial/managerial accounting courses.

In 2008 Jackson and Durkee suggested that the instructor's role must shift from being the presenter of facts to facilitator of active (engaged) learning to help students to achieve the essential professional skills. Previously, Albrecht and Sack (2000) underscored the need to transform the accounting education delivery methods. According to Helliar (2013), a vital component of accounting education should entail teaching methods that engage (enable) students (e.g., role playing, real-world case studies, and group activities). In this research our project involved group work and the use of real-world data (i.e., financial and nonfinancial data from two corporations in the same industry).

The Pathways Commission (2015) proposed the use of Schulman's (2005) five signature pedagogies criteria. Three of these criteria specified that students: (1) should be held accountable for their work to clients, student peers, and faculty; (2) must be active and interactive; and (3) should experience adaptive anxiety. In this research our project incorporated group work and required the students to actively engage in the learning process. Also, this project held the students accountable to their peers. In addition, the students in the preparation of their projects may experience some anxiety (e.g., uncertainty and complexity).

The Chartered Global Management Accountant (CGMA) in the CGMA Competency Framework [CGMA Framework] (2019) discussed five skilled knowledge areas. In addition, the Pathways Commission (2012) in Action Item 4.1.1 advocated the development of an accounting body of knowledge. The Pathways Commission's Knowledge and Pedagogy Task Force in November 2015 proposed a "Common Body of Knowledge Learning Objectives" for accounting. Earlier, the *AICPA Core Competency Framework for Entry into the Accounting Profession* [Framework] (1999) endorsed a set of essential competencies (encompassing100 elements) for accounting students who plan to enter the profession. These competencies are similar to the skill sets discussed by educators and practitioners in Albrecht and Sack (2000), Canadian Institute of Chartered Accountants [CICA] (2010), the Institute of Management Accountants (1994), Institute of Internal Auditors (1999), and the Accounting Education Change Commission [AECC] (1990). Our project was developed to help the students in undergraduate and graduate introductory financial/managerial accounting courses to experience several of the CGMA skilled knowledge areas, many of the Pathways Commission's learning objectives, and about one-half of the AICPA Framework (1999) elements.

THEORY

Learning-to-Learn

The Association of International Certified Professional Accountants (2019) recommended that accountants should recognize the necessity for lifelong professional learning and experience. Earlier, the CPA Vision Project (2017) indicated that there is a need for lifelong learning in the profession (i.e., continuous acquisition of new skills and knowledge). Also, the Pathways Commission (2014) stipulated that high-quality accounting education should prepare students for a lifetime of learning.

In addition, Albrecht and Sack (2000) stated that accounting educators should assist students in development of their ability-to-learn skills. Previously, the Accounting Education Change Commission (AECC) in 1990 indicated that the accounting curriculum "should lay the base on which life-long learning can be built." Further, Gammie and Kirkham (2008) suggested that the ability to "learn to learn" is a key competency that is essential for accountants to adapt to the rapidly changing business environment. Our financial/ managerial project should give the students an opportunity to develop or expand their ability to “learn to learn.”

Under the constructive learning theory, the role of students is to actively participate (engage) in activities that construct their knowledge base. According to this theory the learning environment should be aligned with learning outcomes, which should develop into independent learning. Seifried (2012) and Tan and Ferreira (2012) found that constructive alignment occurs when learning outcomes, teaching and learning activities, and assessment are carefully coordinated. Students need to “learn to learn” so that they will be able to discover needed tacit knowledge for their professional careers. The importance of life-long learning is reinforced as it is one of the top five core values recognized in the AICPA Vision Project (2017). Our learning project allowed the students to participate (engage) in an activity that should help them construct their knowledge base and to learn to learn.

Johnstone and Biggs (1998) advocated supplementing existing textbook material with real world activities (e.g., analysis of realistic data). Many cases have been created using companies’ annual reports and/or SEC 10-K reports for financial accounting courses [e.g., Taylor *et al.* (1998), Kern (2000), Stuart and Karan (2003), Williams and Koch (2004), and Williams (2005)]. However, only a relatively few cases have been developed for cost accounting or graduate managerial accounting courses [e.g., Carter (1999), Dikolli and Sedatole (2004), and Bamber and Bamber (2006)]. Both the Bamber and Bamber (2006) and the Adams *et al.* (1999) cases require an entire semester commitment to complete a series of mini-cases and involve only one company within an industry. Our project was developed to include both financial and managerial accounting topics for undergraduate and graduate students that required the students to analyze financial and nonfinancial data for business decision making.

This study contributes to the accounting literature by presenting an assignment (Financial/ Managerial Accounting Project) that exposes students enrolled in introductory financial/ managerial accounting courses to both financial and managerial information found in publicly traded company annual reports and/or SEC 10-K reports. For example, it was necessary for the students to identify what needs to be measured in the calculation of financial/managerial accounting ratios. Also, the project entailed obtaining managerial accounting information from the Management Discussion and Analysis (MD&A) section of the companies’ annual reports. Our project involved the analysis of two companies within one industry for each team.

Further, the project permitted the students to develop or refine the Pathways Commission’s Common Body of Knowledge Learning Objectives (2015), the CGMA Competency Framework (2019) knowledge areas, and the core competency strategies and techniques enumerated by the AICPA in1999 and again in 2017 while analyzing, writing, presenting, and evaluating their project. This study also contributed to the accounting literature by utilizing value added assessment to evaluate the effects of this learning exercise. In addition, student opinions of this project were obtained and analyzed.

Learning Objectives and Skills

The CGMA Competency Framework (2019) discussed interdependent knowledge areas (i.e., technical skills, business skills, people skills, leadership skills, and digital skills). For example, financial ratio analysis is utilized to assess the performance of the organization, preparation of analysis for decision-making purposes, and participation in teams and groups.

In addition, the Common Body of Knowledge (i.e., Accounting Competencies, Professional Foundational Competencies, and Broad Management Competencies), which was developed by the Pathways Commission (2015), has incorporated measurable learning objectives. Specifically, the Pathways Commission’s Common Body of Knowledge Learning Objectives suggested that accounting students should be able to interpret financial and nonfinancial data

located in external corporate annual reports (e.g., ratio analysis or trend analysis) for decision-making purposes (e.g., investment or credit decisions).
Previously, Albrecht and Sack (2000) specified that accounting courses should emphasize financial statement analysis instead of preparation of the statements. In addition, these authors recommended that an undergraduate accounting curriculum could utilize accounting information (e.g., annual reports) from different industries. Our project required each team to analyze and present information from a different industry for decision-making purposes.

Communication

The ability of the accounting students to communicate (both written and oral) is vital to the profession. Various professional committees have consistently advocated over many years that accounting education should place more emphasis on writing skills, oral skills, and critical thinking skills. More recently, the CGMA Framework (2019), the AICPA Framework (2017) and the Pathways Commission (2015) have indicated that beginning accounting professionals should be able to communicate clearly to the intended audience. Earlier, the AICPA in the Framework (1999) project listed communication as an element of "Personal Competencies." Our learning project permitted the students to write, make oral presentations, and employ critical thinking in making business decisions.
As previously mentioned, the Pathways Commission (2012) recommended that accounting students must possess professional skills (e.g., the ability to communicate effectively). In a follow-up report, the Pathways Commission (2014), stated that quality accounting education should result in an improvement in the students oral and written communication skills. In a survey, Bui and Porter (2010) found that employers considered communication skills to be indispensable for accounting graduates.
According to Conrad and Newberry (2012), despite the educators' best attempts, there still appears to be a gap between practitioners desired communication skills of graduates and what is presently possessed by new graduates. Taking into account this disparity, the Pathways Commission (2015) included under "Professional Foundational Competencies" learning objectives related to communication. Recently, the CGMA Framework (2019) included communication under its "People Skills—Proficiency Level." As suggested by both the CGMA Framework (2019) and the Pathways Commission (2015), our project gave the students an opportunity to make oral presentations and to write on financial statement analysis topics. Specifically, the students were assigned two different types of writing exercises (a team report and an individual report) and an oral presentation as part of their project requirements.
Further, the International Federation of Accountants (IFAC) in its International Education Standard 3 (IES3), *Professional Skills and General Education* (2008) identified the essential skills that are needed by individuals desiring to enter the accounting profession. The IFAC in IES3 emphasized that accountants entering the profession should have interpersonal and communications skills. In addition, the Institute of Chartered Accountants in Australia (ICAA) and Certified Practicing Accountants of Australia (CPA Australia) in their *International Accreditation Guidelines for Accounting Degree Programs* (ICAA/CPA Australia, 2009) indicated that cognitive skills (e.g., writing skills and thinking critically) and behavioral skills (e.g., interpersonal skills) are essential for accounting graduates to succeed. Our project required the students to write, make oral presentations, work in teams, and utilize critical thinking in making business decisions.

Analytical Thinking

The Pathways Commission (2014) denoted that high-quality education encompasses deep learning beyond memorization. The managing partners of the then "Big 8" CPA firms

(“Perspectives . . .”) (1989) indicated that accounting education should place emphasis on analytical and conceptual thinking instead of rote memorization. In addition, the Accounting Education Change Commission [AECC] (1990) recommended that students should actively participate in the learning process and not be just passive recipients of information. Both the Wilkin and Collier (2009) and the Flood and Wilson (2008) studies suggested that accounting programs should deviate from the traditional practice of communicating large amounts of technical knowledge and, instead, facilitate students’ understanding of the principles and concepts involved with accounting/business practices. Our project was designed to encourage the students to do some analytical and conceptual thinking and writing.

The CPA Vision Project (2017) indicated that one of the top five core competencies involves linking data, knowledge and insight to offer strategic decision-making advice. Also, the CGMA Framework (2019) suggested that professionals are expected to capture data, analyze information, advise and influence decision makers, and assist in ensuring required outcomes are achieved.

Earlier, the Pathways Commission (2012) indicated that to be competent an accountant must possess professional skills such as the capability to utilize knowledge to make reasoned judgments. Also, the Broad Business Perspective Competencies category of the *AICPA Core Competency Framework for Entry into the Accounting Profession* [Framework] (1999) suggested that strategic/critical thinking is a necessary component for professional success. In addition, Albrecht and Sack (2000) recommended that accounting education delivery methods be modified to allow students to develop critical skills. Further, students should learn to critically examine and interpret information and ideas according to the Canadian Institute of Chartered Accountants (CICA) in its *CA Skills and Competencies* (2010) report. Classroom techniques (e.g., student engaged projects) should be created to achieve these objectives.

The Pathways Commission’s Common Body of Knowledge Learning Objectives (2015) stated that accounting students should be able to locate appropriate accounting information for guidance in determining accounting measurement methods. Albrecht and Sack (2000) emphasized the importance of “teaching students how to find answers and how to learn.” Also, the AECC (1990) believed that students should have the ability to locate, obtain, and organize information. In addition, the need for entry-level accounting professionals to be able to organize and evaluate information was suggested in the AICPA Framework [Functional Competencies category under Decision Modeling] (1999). Our project required the students to locate, organize, and evaluate financial, managerial, and nonfinancial information.

One of the top five core competencies of the CPA Vision Project (2017) entailed interpretation of both financial and nonfinancial information. Also, in 2017 the CGMA in *Global Management Accounting Principles* stated that one of the main practice areas should involve an integrated and comprehensive view of the organization’s financial and nonfinancial performance along with evaluation of risks and strategy.

The CGMA Framework (2019) indicated that foundational skills (entry-level proficiency) encompass financial ratio analysis to evaluate the organization’s performance. In 2015 the Pathways Commission suggested that students should be able to interpret both financial and nonfinancial data to determine an organization’s operating efficiency and effectiveness. Previously, Albrecht and Sack (2000) recommended that an undergraduate accounting curriculum should include analysis and use of accounting information in making decisions. Also, accounting graduates should be able to analyze the information gathered and form an opinion on the impact of the information for a particular situation according to the CICA in its *CA Skills and Competencies* (2010). Our project required the students to analyze company financial/ nonfinancial accounting information (e.g., MD&A reports) as a basis for decision-making purposes (e.g., the choice of company in which to invest).

Leadership and Human Relations

Another one of the top five core competencies of the CPA Vision Project (2017) encompasses influencing, inspiring, and motivating others to accomplish results. Recently, the CGMA Framework (2019) indicated that accountants should encourage team performance, train and advise others, and motivate and inspire others. Previously, Albrecht and Sack (2000) suggested group activities should be used to teach both leadership and how to work together.

Under Professional Competencies of the AICPA Framework (2017), students should be able to work effectively with diverse individuals, promote inclusion, and enhance performance to accomplish objectives. Under Leadership Skills in the CGMA Framework (2019), entry-level professionals should welcome diverse ideas, interact effectively with everybody, contribute to groups or teams, and value the contribution and efforts of group or team members. Also, the IFAC in IES3 (2008) indicated the need for students to learn to work with others, to negotiate acceptable solutions, to listen effectively, and to solve any conflicts that may occur. Further, both the Kavanagh and Drennan (2008) and the Kennedy and Sorensen (2006) studies stated that employers expect accounting graduates to be able to successfully work on a team, in addition to having technical and analytical accounting skills. Our project involved students working together on teams.

Students should learn to think and perform as professionals during their educational experience according to the Pathways Commission (2015). Prior to this report, Sin, Reid and Jones (2012) advocated that accounting education should assist students in developing transferable skills that will help them to make the transition from their present learning environment to the work environment. These transferable skills include report writing and interpersonal skills. Stone and Lightbody (2012) found that interpersonal skills need to involve both listening and oral communication. Our project required the students to not only prepare a written report and make oral presentations, but also, to encourage them to learn to listen.

According to Schulman (2005) signature pedagogy should provide learning experiences that hold students accountable for their work (e.g., to clients, peers and faculty). In addition, the AICPA Framework [Personal Competencies category under Professional Demeanor] (1999) advocated that students should objectively respect the professional assessment or evaluation of others. Also, Boud *et al.* (2010) suggested that students should learn to make judgments about their own work and the work of others (i.e., peer review) in order to become successful continual learners and future accounting professionals. Our financial/managerial project gave the students an opportunity to evaluate their classmates' presentations (i.e., to perform peer review). As a result, the students received feedback/criticism from their peers.

Value-Added Assessment

One technique to evaluate students' learning is to use pre- and post-study measurements. According to Angelo and Cross (1993), the purpose of utilizing pre- and post-assessment techniques is to determine whether students have benefited from class discussions and assignments. The pre-test allows the instructor to establish a benchmark of what the students know on the subject matter being investigated before the study technique (project) is utilized. Gordon (1998) used the pre- and post-assessment technique to evaluate students' knowledge of social responsibility accounting.

In our study, Exam II, which was administered after discussing the homework problems on financial/managerial accounting topics (before the project was assigned), was designated as the pre-test. After Exam II there was no other class discussion or homework involving financial/ managerial accounting ratios or other questions included in the project. The Final Exam, which was given after the students completed the project, was considered to be the post-test.

In summary, the literature suggests that student engaged learning exercises can enrich the learning process and that writing projects can be utilized as a learning tool. The following hypothesis was used to test the benefit of this engaged learning technique (project):

> H_1: The distribution of exam scores for undergraduate (graduate) students on the financial/managerial accounting questions for Exam II (before the engaged learning exercise) and for the Final Exam (after the engaged learning exercise) are the same.

Student Perceptions

Ennis (1987) stated that attitudes strongly determine the ability to apply intellectual skills (e.g., written communication skills). Stout and Rebele (1996) indicated the need to examine student attitudes toward a teaching method. According to Ennis (1987), if students do not have positive attitudes toward a teaching method, desired learning outcomes (e.g., intellectual skills) may not occur.

Stone and Shelley (1997) used survey questions to measure student perceptions of the instructional processes. Ramsay *et al.* (2000) utilized a survey to determine the students' preferences for a cooperative learning method. Sawyer *et al.* (2000) administered a survey after the students received their grade and assessment sheet for a case to determine if they believed that the case was useful in meeting specified skills. Chu and Libby (2010) utilized a post-assignment questionnaire to evaluate an active learning assignment.

In this research study, students' opinions of the usefulness of the Financial/Managerial Accounting Project in accomplishing selected Pathways Commission's Common Body of Knowledge Learning Objectives (2015) and the AICPA's Framework Core Competencies (2017 and 1999 versions) were measured using a survey. This survey was given after the students received their project results.

Pathways Commission – Common Body of Knowledge

The Pathways Commission's Common Body of Knowledge Task Force began by examining the AICPA Core Competency Framework (1999) along with the work of other accounting task forces (e.g., Chartered Professional Accountants Canada, Chartered Accountants Australia, Institute of Management Accountants, and International Accounting Standards Board Education Initiative) to develop learning goals for accounting education (i.e., undergraduate, graduate, and beyond). To concentrate on undergraduate accounting education, a new task force of the Pathways Commission was formed (Knowledge and Pedagogy Task Force). This task force translated the goals of the Common Body of Knowledge Task Force into measurable undergraduate learning objectives.

The learning objectives were categorized into three competency areas (i.e., Accounting, Broad Management, and Professional Foundational). The Accounting Competency was divided into six subsections (i.e., external reporting and analysis; planning, analysis and control; taxation compliance and planning; information systems; assurance and internal control; and professional values, ethics, and attitudes). There are five subsections under the Professional Foundational Competencies (i.e., communication; quantitative methods; analytical thinking and problem solving; human relations; and technology). Finally, there are also five subsections under the Broad Management Competencies (i.e., leadership; organizational ethics and social responsibility; process management and improvement; governance, risk, management, and compliance; and additional core management competencies). The project's desired accounting learning objectives have been classified by the researchers into appropriate subsections of the Pathways Commission's (2015) Common Body of Knowledge Learning Objectives.

About half of the subsections of the Accounting Competencies have at least one project learning objective that the students were exposed to while preparing their engaged learning project. Specifically, all of the learning objectives of the professional values, ethics and attitudes subsection were integrated into our project. For the Professional Foundation Competencies, all of the subsections had at least one learning objective incorporated in the assigned project. At least half of the learning objectives were expected to be accomplished with our project for the communication; analytical thinking and problem solving; and human rights subsections. Of the Broad Management Competencies, all except one of the subsections were expected to have at least one learning objective reflected in the project. More than half of the learning objectives for the organization, ethics and social responsibility subsection were assimilated into our engaged learning project.

AICPA Framework

The AICPA Framework (2017 and 1999 versions) each provided a set of needed competencies for all students preparing to enter the accounting profession and/or the business world. The Framework (2017) competencies were divided into accounting, professional, and business. Our project encompassed elements in all of these competencies.

The Framework (1999) competencies were classified as functional, personal, and broad business perspective. Our project incorporated elements of the functional competencies category. Also, our project was expected to permit the students to experience elements of the personal and broad business perspective categories.

The team Financial/Managerial Accounting Project (project) was expected to result in elements of all three categories of the Framework (2017 and 1999 versions) competencies being integrated into introductory financial/managerial accounting courses (an undergraduate course and a graduate course) with a minimum of two to three hours of class time.

As previously mentioned, students' opinions of the usefulness of the project in accomplishing selected Pathways Commission's Learning Objectives (2015) and selected AICPA's Framework Core Competencies (2017 and 1999 versions) were measured using a survey. Our survey requested the students to indicate their opinions as to whether the competencies/learning objectives were accomplished in preparing, presenting, and evaluating the project.

Two Different Universities

One of the limitations of educational research that is conducted at only one university is whether the results will apply to other university settings. Accounting instructors are interested in teaching techniques or methods that might be successfully utilized in different university environments. One of the objectives of this research study was to determine if there would be different Final Exam score results when the outside of class project was used at two universities located in different regions of the country (midwest regional and southwest regional [95% English second language students]).

Another objective of this study was to determine whether there would be a different result when the introductory financial/managerial course was taught at the undergraduate level as compared to the graduate level. An additional objective was to ascertain whether the student-perceived benefits from using the engaged learning project vary between these two different university settings.

RESEARCH METHODS

One of the researchers taught both of the introductory financial/managerial accounting classes used in the research experiment. One of the universities was a midwest regional state university and the other was a state university with 96% minority students [95% English second language], which is located in the southwest. At the midwest university (MW-U) the students were undergraduate sophomores, while the students at the southwest university (SW-G) were graduate MBA students with no previous accounting knowledge.

The project involved financial/managerial accounting analysis for the two most recent years of annual reports and/or SEC 10-K reports for two companies in the same industry (e.g., Goodyear and Cooper Tire) for each team. Each team (consisting of 3-5 students) analyzed a different industry (e.g., office furniture, petroleum industry, or wireless components).

Project and Team Selection

The students at both universities were allowed to select the industry from pairs of companies identified by the instructor. The researchers have found that students are more motivated to do this project if at least one of the firms is a regional firm or at least a nationally recognized firm (e.g., Caterpillar and Deere).

In addition, if the company's annual report does not incorporate the SEC 10-K report, the project required the students to access their company's SEC 10-K report on the Internet. Other than the selection of the industry by the team, the entire project was designed to be prepared outside of class.

Written Reports and Oral Presentations

The students at both universities were required to prepare a written team report. This report had three parts. First, the team was required to calculate ten financial/managerial accounting ratios for the two most recent years. Then, the students were expected to use the ratio analysis results as the basis for answering a set of questions. For example, "Which company has the more favorable inventory turnover?" Another example was "Which company has the more favorable gross profit percentage trend?"

Finally, for the most recent year, each team was required to answer another set of questions related to each company's annual and SEC 10-K reports. For example, "What inventory valuation methods does each company use?" Another question required the students to identify two segments (e.g., product lines) reported by each firm and indicate if there was a product line or business segment that they would want to consider dropping in the future.

The third part of the team report was designed to expose the students to other sections of the company's annual report in addition to the financial statements and accompanying notes (e.g., Management's Discussion and Analysis (MD&A) report). The students were requested to use the MD&A section to determine if the company management made forward-looking or projection statements. Then, if company projections were presented in the MD&A section, the students were asked to identify two factors related to the projections that would also affect the preparation of a budget for each company. For example, "What assumptions discussed in the MD&A section were related to the development of sales forecasts?"

In addition, to assure that each student had writing experience during preparation of the project, each student was required to write a one-page report. This individual report required the students to give their opinions as to the firm they would select for investment purposes. They were expected to support their conclusions based on the team data developed in the first three parts of the project requirements.

Also, each team was required to orally present their analysis to their classmates. The students were expected to dress as if they were presenting to clients. The teams were informed that their presentations should be 15 to 20 minutes in length. The other classmates were expected to act as the client and were encouraged to ask questions of the presenting team.

Evaluations

As the Pathways Commission (2012) recommended, the project was designed to promote deep engagement of the students by holding them accountable to the instructor and fellow students through the use of an evaluation form during their oral presentations. Each student's presentation was evaluated as well as the team as a whole.

The students at both universities received (one week before the presentation) a copy of the evaluation form, which was to be used during their presentations. Upon reviewing the evaluation form, the students knew that they would be scored (5-1 Likert scale with 5 being the highest) based on (1) the content of their segment or topic presented, (2) the organization of their presentations, (3) the use of visual aids during their presentations, and (4) the effectiveness of the delivery of their presentations. These grading characteristics were defined on the evaluation form. For example, the characteristics for effectiveness of the delivery of the individual student's presentation were listed as: "Did the presenter's voice enhance the effectiveness of the presentation?" "Did the presenter maintain good eye contact?" "Did the presenter sound natural and professional?"

In addition, the evaluation form defined the characteristics on which the teams were evaluated. The team score was based on (1) the content of the team project presented; (2) the organization of the entire team presentation; (3) the coordination of visual aids used by the team; and (4) the cohesiveness of the team presentation.

The instructor evaluated the team and each individual on the team. As mentioned above, the non-presenting class members were also required to complete the evaluation form. Each classmate and the instructor had equal evaluation weight in determining the student scores for the presentations. This helped the students to consider the presentations more seriously.

The students were asked to keep the evaluation forms confidential and were told that only the instructor will compute the presenter's scores. The evaluation form permitted the evaluators to write comments and suggestions for each presenter and for the team presentation as a whole. One of the purposes of the evaluation form was, as suggested by the Pathways Commission (2015), to hold the presenting students accountable to their student peers and the instructor. Also, as implied by the Pathways Commission (2012), it was expected that the students will benefit from peer evaluation/observation and comments.

When the written reports were returned, each student at both universities also received his or her project score and the written instructor/peer comments for the individual and team presentations. The total score for the project was 70 points (team report, 40; individual report, 15; individual presentation, 10; and team presentation, 5).

Testing

Approximately a week following the discussion of the homework problems on the financial/managerial analysis topics, Exam II was administered. Exam II was given before the students started their engaged learning exercise (i.e., Financial/Managerial Accounting Project). In our study, Exam II was designated as the pre-test.

About a week after the students made their presentations and completed their team and individual reports, the Final Exam was administered. The Final Exam had questions related to the financial/managerial analysis topics that were different than those that were given on Exam II but were similar in the level of complexity. The Final Exam was considered as the post-test in our study. The results of Exam II and the Final Exam were used to measure the effect of this engaged learning exercise technique.

Student Perceptions – Pathways Learning Objectives – AICPA Framework

This study used a survey to measure student perceptions of this engaged learning technique as Stone and Shelly (1997) did in their research. The survey was utilized to measure the students' opinions of the usefulness of the Financial/Managerial Accounting Project in accomplishing selected Pathways Commission (2015) Learning Objectives and the AICPA Framework Core Competencies (2017 and 1999 versions). In addition, the survey was used to ascertain the opinions of the students on the usefulness of the project in understanding the financial/ managerial accounting analysis topics.

Like the Sawyer *et al.* (2000) study, our survey was given after the students received their project results (i.e., their grade and assessment sheet). The survey requested the students to rank (strongly agree = 5) whether selected Pathways Commission Learning Objectives and AICPA Framework Core Competencies were accomplished during their project preparation, team presentation, individual presentation, completion of the peer evaluation form, and reviewing peer evaluations.

RESULTS

The students at both universities were tested twice on the financial/managerial accounting analysis topics. First, an exam (Exam II) was administered after discussing the homework problems on these topics, but before the Financial/Managerial Accounting Project (project) was assigned. Exam II was considered as the pre-test. As previously discussed, there was no other class discussion or homework involving financial/managerial accounting ratios or other questions included in the project after Exam II. The second exam (Final Exam) was given after the students had completed the project. The Final Exam was designated as the post-test. Different questions related to the project topics were given on these exams.

At both universities, the majority of the student exam scores related to the financial/managerial accounting analysis topics increased or stayed the same after the project was completed. For the undergraduate midwest university (MW-U) course, the mean score increased from 43.2% on Exam II to 55.2% on the Final Exam. The mean score for the graduate southwest university (SW-G) course increased from 29.7% on Exam II to 53.0% on the Final Exam.

The students' Exam II and Final Exam scores were matched by names. The Wilcoxon signed rank test was utilized to test H_1 (EII ≥ FE). Since there was a significant difference ($p = .01$ at MW-U and $p = .05$ at SW-G), H_1 was rejected. The students' exam scores appeared to have significantly increased at both universities as a result of this engaged learning exercise. Thus, it appears that this type of learning exercise can be successfully used as a teaching method for financial/managerial accounting analysis topics at both the undergraduate and graduate levels.

Like Sawyer *et al.* (2000), the researchers administered a survey after the students received their grade and assessment sheet for the project. The survey was used to determine whether the students felt that the assigned project was useful in meeting specified Pathways Commission Learning Objectives and AICPA Framework skills while analyzing, writing, presenting, and evaluating this project.

Generally, the students stated that they "strongly agree" or "agree" that the selected core competencies were accomplished by the project. The average mean score for the students at MW-U was 4.109 and the median score was 4.125. At SW-G the average mean score for the students was 4.356 and the median score was 4.440. The scores ranged at MW-U from 4.750 to 3.500 and at SW-G from 4.780 to 3.560.

Pathways Commission Common Body of Knowledge Learning Objectives

The Pathways Commission Common Body of Knowledge Learning Objectives were categorized into Accounting Competencies, Broad Management Competencies, and Professional Foundational Competencies. Many of the Accounting Competencies subsections were included in our project. The “external report and analysis” subsection of the Pathways Commission Accounting Competencies was one of the main reasons for this project with most MW-U students indicating a ranking of 4.13 or higher and a mean score of 4.32 and most SW-G students indicating a ranking of 4.10 or higher and a mean score of 4.32. Generally, both the MW-U and SW-G students believed that the “planning, analysis, and control” subsection was accomplished with a mean score of 4.13 for MW-U students and 4.36 for SW-G students. Another purpose of the project was to introduce the students to ethics and social responsibility. Therefore, the researchers were interested in the results of the questions for the “professional values, ethics, and attitudes” subsection. Both the MW-U and SW-G students generally believed this project overall improved their knowledge of this subsection with most of the related questions having a ranking of 4.00 or higher for the MW-U students and 4.10 or higher for the SW-G students and a mean score of 4.27 (MW-U) and 4.50 (SW-G). The difference between these two groups could have resulted because there was a dysfunctional team at the MW-U university.

Also, the students’ opinions of the project’s exposure to all of the subsections of the Pathways Commission Professional Foundational competencies were examined. Our project was designed to increase student communication skills, which is not normally part of an Introductory Financial/Managerial Accounting course. Thus, the results of the “communication” subsection were important. The majority of the MW-U students’ scores were 4.00 or higher and all of the SW-G students’ scores were 4.11 or higher with a mean score of 4.13 for the MW-U students and 4.38 for the SW-G students. The difference between the MW-U and SW-G students could be because the MW-U students (sophomores) may not have had to make presentations in their freshman level courses (which made the students feel less confident) while the graduate students at SW-G probably had made presentations prior to this course.

Another learning objective of our engaged learning project was to motivate the students to do some analytical and conceptual thinking and writing. Specifically, it was hoped that the students in the project preparation would learn to critically examine and interpret information [as suggested by the Canadian Institute of Chartered Accountants (2010)]. Explicitly, our project expected the students to analyze company financial/nonfinancial accounting information [as recommended by The Pathways Commission (2015)] as a basis for decision-making purposes (e.g., the choice of company in which to invest). As a result, the student opinions of the “analytical thinking and problem solving” subsection were examined. Most of the MW-U students’ scores for these questions were 4.00 or higher and the majority of the SW-G students’ scores were 4.11 or higher with a mean score of 4.17 for the MW-U students and 4.37 for the SW-G students.

One of the advantages of the team project was the opportunity for the students to develop better “human relations” than is normally possible in an Introductory Financial/ Managerial Accounting course. Most of the MW-U students’ opinion scores were 4.00 or higher and the majority of the SW-G students’ scores were 4.33 or higher for the human relations subsection questions with a mean score of 4.19 for the MW-U students and 4.44 for the SW-G students. Probably the lack of previous teamwork experience by the MW-U students may explain the resulting score differences between the undergraduate (MW-U) and the graduate (SW-G) students.

Our project required some use by the students of technology (e.g., locating and downloading their companies’ annual reports and/or SEC 10-K reports). For the MW-U students, the majority

of their opinion scores were 4.00 or higher and all of the SW-G students' scores were 4.11 or higher with a mean score of 4.15 for the MW-U students and 4.36 for the SW-G students.
The Pathways Commission's final Common Body of Knowledge competency is the Broad Management category. Because the project involved teams, it was hoped that the students would have some leadership opportunities while working on the project. Most of the MW-U students' scores related to this competency were 4.00 or higher and the SW-G students' scores were 4.22 or higher with a mean score of 4.16 for the MW-U students and 4.43 for the SW-G students.
Another subsection of the Broad Management category relates to "governance, risk management, and compliance," which was not emphasized in the project. As a result, the student opinions were relatively low with a mean score of 3.79 for the MW-U students and 3.93 for the SW-G students.

AICPA Framework Core Competencies

As previously discussed, this Financial/Managerial Accounting Project should facilitate the students in accomplishing elements of all three of the AICPA Framework's broad skills categories (Functional Competencies, Personal Competencies, and Broad Business Perspective). Specifically, the researchers expected about one-half of the AICPA Framework's skills to be achieved by the students.
Of the elements that the researchers expected to be accomplished, the students ranked 70.49% and 93.44% as either "strongly agree" or "agree" at MW-U and SW-G, respectively. The ten highest elements at MW-U had a score of 4.50 or higher. Because of ties the 20 highest elements at SW-G had a score of 4.56 or higher.
The majority of the core competency elements achieved by this project were related to the personal competencies category. Personal competencies may be more difficult to achieve with traditional activities in an undergraduate or graduate introductory financial/managerial accounting class. Of the top ten elements at MW-U, 80% are from the personal competencies category. Of the top 20 elements (because of ties) at SW-G, 80% are from the personal competencies category.
The survey also requested the students to rank (strongly agree = 5) whether preparing the Financial/Managerial Accounting Project greatly assisted them in understanding the financial statement analysis and managerial accounting topics. The students "strongly agreed" or "agreed" (4.75 mean score at MW-U and 4.89 mean score at SW-G) that this project increased their understanding of these topics.

SUMMARY

Many of the students received higher scores on the Final Exam related to the financial statement analysis and managerial accounting topics after they completed the project than they did on Exam II, which involved only regular class discussion and homework assignments. There was a significant difference at p = .01 for MW-U (undergraduate course) and at p = .05 for SW-G (MBA graduate course) between the mean scores on Exam II (before the project) and the Final Exam (after the project). The students' exam scores appear to have significantly increased at both universities as a result of this project. Therefore, this engaged learning exercise (project) appears to be beneficial for both undergraduate and graduate students.
Many of the specified Pathways Commission (2015) Learning Objectives and the AICPA Framework Core Competencies (1999 and 2017 versions) were accomplished when the students prepared, presented, and evaluated their projects. On the survey, the SW-G students ranked 93% of the learning objectives/core competencies that the researchers expected to be accomplished as either "strongly agree" or "agree." The SW-G students' average mean score

was 4.37 and the median score was 4.44. The scores ranged from 4.78 to 3.56 for the SW-G students.

Of the expected elements, about 80% of the top scores at both universities were from the learning objectives/personal competencies category. This project gave the students an opportunity to experience types of skills recommended by the Pathways Commission (2015) Learning Objectives and the AICPA Framework Core Competencies (1999 and 2017 versions) that may not normally be accomplished in a typical undergraduate or graduate financial/ managerial accounting course (e.g., professional foundational/personal competencies). Further, the students indicated that they “strongly agreed” or “agreed” (4.75 mean score at MW-U and 4.89 mean score at SW-G) that this project increased their understanding of the financial statement analysis and managerial accounting topics.

Also, in the preparation of this project the students were active participants in the learning process as suggested by Schulman (2005) and the AECC (1990). The use of actual company data in this project permitted the students to analyze and interpret financial information as advocated by Albrecht and Sack (2000). Further, as recommended by both the Pathways Commission (2015) and the AECC (1990), the students were required to locate, obtain, and organize/analyze financial and nonfinancial information.

This project can be accomplished using a minimum of two to three hours of class time. Of course, the amount of class time spent on the project will vary depending on the size of the class. In conclusion, it appears that this engaged learning exercise (project) can be successfully used as a teaching method for financial statement analysis and managerial accounting topics in an undergraduate or graduate financial/managerial accounting course.

REFERENCES

References available upon request.

DECISION SCIENCES INSTITUTE

Lean Supply Chain: A Cluster Analysis of U.S. Manufacturing Industries

Full Paper Submission to the 2019 DSI Annual Meeting

Chaodong Han
Towson University
Email: chan@towson.edu

Barin N. Nag
Towson University
Email: bnag@towson.edu

ABSTRACT

Using cluster analysis, we present research examining the characteristics of lean operations by U.S. manufacturing industries based on U.S. Economic Census data. From an end-to-end supply chain perspective, we aim to develop insights into the learning practices by industry groups based on a set of industry characteristics related to lean operations, including manufacturing worker productivity, fuel consumption, capital expenditures, inventory, waste, etc.

KEYWORDS: Lean operations, Economic Census, Cluster analysis, Manufacturing industries

INTRODUCTION

Simply put, lean manufacturing in terms of a management philosophy is all about relentless efforts to eliminate waste from the manufacturing process. In lean manufacturing, waste is broadly defined to include any activity that does not add value to the end-to-end supply chain, including wasted materials, defected products, overstocked inventory, emissions, underutilized capacity, underutilized skills, lost efficiency and lost productivity due to unevenness of the workflow, etc. According to UK Lean Enterprise Research Center (LERC), nearly 60% of production activities in a typical manufacturing process are waste since they do not add value to the end customers.

Ford may be the first company to implement the pristine form of lean manufacturing by using standardized processes in the assembly line of the Model T automobiles where each stage was fitted together tightly to ensure a continuous and consistent production flow. However, it is Toyota that perfected the manufacturing process through its innovative "Toyota Production System" which identifies and eliminate expenditures of material, effort and time that do not generate value for customers in late 1990s. Toyota identified eight forms of waste from a customer's perspective, including transport, inventory, motion, waiting, overproduction, over processing, defects and underutilized workers. Those activities may be necessary steps in delivering final products to the end customers; however, customers do not appreciate their value and may be unwilling to pay for them. Therefore, manufacturing companies shall try their best to eliminate or minimize those wasteful activities.

In terms of best practices, lean manufacturing includes just-in-time production (JIT), which means "making only what is needed, only when it is needed, and only in the amount that is needed", total quality management (TQM), Six Sigma, automation and avoidance of human errors, 5S (Seiri, Seiton, Seiso, Seiketsu, and Shitsuke in original Japanese terms; Sort, Set in order, Shine, Stanardize, and Sustain in English translation), Kanban (pull systems), Kaizen (continuous improvement), total productive maintenance (TPM), supplier relationship development, value stream mapping (VSM).

Based on U.S. Economic Census data, in this research we set out to map the state of lean manufacturing for the U.S. manufacturing industries using cluster analysis. In the following sections, our paper progresses with literature review of lean philosophy and lean practices. We then propose a supply chain perspective of lean practices and methodologies. In the section of research methodology, we introduce data collection process and conduct cluster analysis to map the lean operation status of all U.S. manufacturing industries, complemented by case studies of selected industries in each cluster.

LITERATURE REVIEW

We review lean production literature from three perspectives: definition and evolution of lean production as a management philosophy; lean practices and methodologies; and performance indicators of lean operations.

Definition and Evolution of Lean Manufacturing/Production

Womack, Jones and Roos (1990) prescribed that lean production uses less of everything compared to mass production, including less human effort in factory, less manufacturing space, less investment in tools, fewer engineering hours in new product development, less inventory, fewer defects, more flexibility, greater variety of products, and greater efficiency. Shah and Ward (2007) defined lean production as an integrated socio-technical system whose main objective is to eliminate waste by concurrently reducing or minimizing supplier, customer and internal variability.

Garza-Reyes, Kumar, Chaikittisilp and Tan (2018) operationalized lean manufacturing as five methods and tools, including TPM, VSM, Automation, Kaizen/CI, and JIT and used four indicators to measure the environmental impacts of lean manufacturing: material use, energy consumption, non-product output and pollutant release. Material use goes beyond raw material to include water consumption, packing material, hazardous material and the use of recycled material. Energy consumption refers to the use of fuel and electricity. Generally, the use of electricity generates less waste. Non-product output and pollutant releases refer to waste and hazard for disposal, air emissions and landfill.

Aqlan and Al-Fandi (2018) provides a summary of waste considered by lean operations, including inadequate processing, unnecessary transportation, excess motion, defects, waiting, over production, and inventory. When waste activity is eliminated in a process, cost is reduced and the process becomes more efficient. Believing that lean production focuses on the elimination of waste and creation of value for customers, Womack and Jones (1996) identified three additional waste, including underutilization of people, environmental waste, and underutilization of facilities, and five critical elements of lean implementation, including value, value stream, flow, pull, and the pursuit of perfection. Value is emphasized to keep customer focus. Value stream analysis is used to identify non-value-added activities. The flow element is to ensure a smooth and leveled process without interruption and too much variation. Pull ensures that firms only produce what customers want and are willing to pay for. Perfection encourages firms to continue to improve.

First introduced by Motorola in the 1980s, Six Sigma is a process improvement philosophy aimed to reduce variation in the process and seeks to identify the causes of defects and develop effective corrective actions through "Define, Measure, Analyze, Improve, and Control" (DMAIC). Lean operation does not use statistical tools to achieve process capabilities; Six Sigma is not designed to optimize the process flow. Lean Six Sigma combines both lean operations and Six Sigma methodology.

Business sustainability and social and environmental responsibility in the 21st century call for companies to seek innovative solutions for cleaner technologies. Taddeo, Simboli, Di Vincenzo and Ioppolo (2019) claimed that environmental issues have gradually required manufacturers to pursue a more rational use of resources and a reduction in wastes production. The emerging Lean and Clean Production suggests that firms should create more value with less impact and that environmental eco-efficiency is of the same importance to companies as operational efficiency and market effectiveness. Based on a bibliometric and network analysis of lean and clean production literature, this study identifies keywords, key constructs, subject areas, author collaboration in Lean and Clean Production research.

Benefits of lean manufacturing include increased productivity, reduced cost, shortened lead times, increased volume flexibility and enhanced competitiveness (Shah and Ward, 2003).

Identification of Lean Manufacturing Practices/Methods

However, not all organizations have achieved desired outcomes and success when implementing lean production (Wickramasinghe & Wickramasinghe, 2017; Doolen and Hacker, 2005). Wickramasinghe & Wickramasinghe (2017) provided two interesting findings which may well account for performance variation of lean operations. First, lean operational practices complement each other and only a systematic implementation of all lean operation methodologies may achieve desired outcomes. Second, lean duration moderates the relationship between adoption of lean production practices and lean performance outcomes. Essentially, it takes time for firms to implement lean manufacturing and for lean manufacturing to generate benefits. Lean manufacturing is not an effective tool to solve short-term competitive problems.

Another underlying factor for performance variation is the unfortunate fact that there is no consensus on what qualifies as lean practice. James-Moore and Gibbons (1997) summarized five methods, including general waste elimination, flexibility, people utilization, process control, and optimization. Koufteros (1998) identified seven categories of lean operations, including shop floor employee involvement in problem solving, re-engineering set-up, cellular manufacturing, quality improvement, preventative maintenance, dependable suppliers, and pull production.

Flynn, Schroeder, and Flynn (1999) defined ten kinds of lean practices to include employee development, management technical competence, design for customer needs, worker participation, proprietary equipment, continuous improvement, process control, feedback of information, pull system, and JIT supplier relationship. Panizzolo (1998) proposed six broad categories: process and equipment, manufacturing planning and control, human resources, product design, supplier relationships, and customer relationships. Shad and Ward (2007, JOM, 25/4)'list includes supplier feedback, JIT delivery by suppliers, customer involvement, pull, continuous flow, setup time reduction, TPM, statistical process control, and employee involvement.

Hayes and Wheelwright (1984) summarized lean production methods into two major categories: structural, which includes capacity, facilities, technology, and vertical integration, and infrastructural, including workforce, quality, production planning, and organization. Sakakibara, Flynn, Schroeder, and Morris (1997) proposed slightly different two categories: infrastructure practices include quality management, workforce management, manufacturing strategy, organizational characteristics, product design; JIT practices, including setup time reduction, scheduling flexibility, maintenance, equipment layout, Kaban and JIT supplier relationships.

Wickramasinghe & Wickramasinghe (2017) complies a comprehensive list of lean manufacturing practices, including cellular manufacturing, multitasking workforce, lot size reduction, JIT, work delegation, total productive maintenance (TPM), set-up time reduction, total quality management (TQM), continuous flow production, agile manufacturing. Lean production also covers process capability measurement, safety improvement programs and human resource management.

Performance Effects of Lean Manufacturing

Manufacturing performance can be described in multiple dimensions, including manufacturing plant's labor efficiency (Arthur, 1994), manufacturing plant productivity (Ichniowski et al., 1997), machine efficiency (Youndt et al., 1996), on time delivery (Swink et al., 2007), inventory management (Hofer et al., 2012), production volume flexibility (Cua et al., 2001), and manufacturing cost efficiency (Bozarth et al., 2009). The ultimate benefit of lean manufacturing is to enhance a firm's manufacturing competitive capabilities (Jabbour et al., 2013). (Wickramasinghe and Wickramasinghe, 2017) applied seven measures for manufacturing plant performance, including waste elimination, defects minimization, cross-functional teams, continuous improvement, JIT and pull, information availability and employee involvement. Garza-Reyes, Kumar, Chaikittisilp and Tan (2018) examined the environmental performance of manufacturing firms that implemented lean methods and tools. Those performance measures include materials use, energy consumption, non-product output and pollutant releases.

Not surprisingly, many of those performance measures are interchangeably used with lean practices, such as waste elimination, cross-functional teams, continuous improvement, JIT, etc. To overcome those shortcomings in the literature, in this study we decide to use more objective measures for performance effects, such as inventory, waste releases, fuel consumed, etc.

CONCEPTUAL FRAMEWORK

In this study, we set out to map lean manufacturing practices from an end-to-end supply chain perspective. A manufacturing supply chain is comprised of upstream suppliers, internal operations and downstream customers. Lean manufacturing is a systematic approach for the entire supply chain. A supply chain is as lean as the fattest section. Therefore, it is meaningful to understand lean manufacturing practices from a supply chain perspective.

Based on our survey of lean operations, we find that some lean practices are more used to deal with supplier relationship management, such as JIT delivery; some are more likely to be used in internal operations, such as plant capacity utilization, machine setup time and down reduction, utilization of labor skills; and a few other practices deal with downstream customers, including design for customer needs, pull, and customer involvement. Therefore, lean practices are classified as follows from a supply chain perspective (Figure 1).

RESEARCH METHODOLOGY

Han & Nag   Lean Supply Chain: A Cluster Analysis of U.S. Manufacturing Industries

We collect archival data and perform cluster analysis to map the lean operation status of all U.S. manufacturing industries, complemented by case studies of selected industries in each cluster.

Data Collection

We collected the most recently available data on U.S. manufacturing industries from U.S. 2012 Economic Census (note that 2017 Economic Census data will be published until late 2019). The U.S. Economic Census includes all manufacturing establishments classified in sectors 31-33 with one or more paid employee at any time during the year. The unit of analysis in this study is six-digit North American Industry Classification System (NAICS) manufacturing industries. There are 364 six-digit NAICS manufacturing industries ranging from 311111 9 (Dog and cat food manufacturing) to 399999 (all other miscellaneous manufacturing, including candles, lamp shade, hot tubs, pet supplies, etc.).

Figure 1   A Supply Chain Perspective on Lean Practices

<table>
<tr><th>Upstream Suppliers</th><th>Internal Operations</th><th>Downstream Customers</th></tr>
<tr><td>Supplier JIT
On time delivery</td><td>Setup time reduction
Down time reduction
Utilization of plant facilities
Utilization of people skills
Production volume flexibility
Employee involvement
Work delegation</td><td>Pull
Design for customer needs
JIT delivery
Customer involvement</td></tr>
<tr><td colspan="3">Waste reduction including defects, rework and redundant work
Total Quality Management
Continuous improvement
Inventory management
Safety improvement
Capacity utilization
Information sharing</td></tr>
</table>

Below is the complete list of variables captured in U.S. 2012 Economic Census data for manufacturing industries:

- Number of companies
- Number of establishments
- Establishments with 0 to 19 employees
- Establishments with 20 to 99 employees
- Establishments with 100 employees or more
- Number of employees
- Annual payroll ($1,000)
- Total fringe benefits ($1,000)

- Employer's cost for health insurance ($1,000)
- Employer's cost for defined benefit pension plans ($1,000)
- Employer's cost for defined contribution plans ($1,000)
- Employer's cost for other fringe benefits ($1,000)
- Production workers average for year
- Production workers for pay period including March 12
- Production workers for pay period including June 12
- Production workers for pay period including September 12
- Production workers for pay period including December 12
- Production workers annual hours (1,000)
- Production workers annual wages ($1,000)
- Total cost of materials ($1,000)
- Cost of materials, parts, containers, packaging, etc. used ($1,000)
- Cost of resales ($1,000)
- Cost of purchased fuels consumed ($1,000)
- Cost of purchased electricity ($1,000)
- Cost of contract work ($1,000)
- Quantity of electricity purchased for heat and power (1,000 kWh)
- Quantity of generated electricity (1,000 kWh)
- Quantity of electricity sold or transferred (1,000 kWh)
- Total value of shipments and receipts for services ($1,000)
- Value of resales ($1,000)
- Value added ($1,000)
- Total inventories, beginning of year ($1,000)
- Finished goods inventories, beginning of year ($1,000)
- Work-in-process inventories, beginning of year ($1,000)
- Materials and supplies inventories, beginning of year ($1,000)
- Total inventories, end of year ($1,000)
- Finished goods inventories, end of year ($1,000)
- Work-in-process inventories, end of year ($1,000)
- Materials and supplies inventories, end of year ($1,000)
- Gross value of depreciable assets (acquisition costs), beginning of year ($1,000)
- Total capital expenditures ($1,000)
- Capital expenditures for buildings and other structures ($1,000)
- Capital expenditures for machinery and equipment ($1,000)
- Capital expenditures for automobiles, trucks, etc. for highway use ($1,000)
- Capital expenditures for computers and peripheral data processing equipment ($1,000)
- Capital expenditures for all other machinery and equipment ($1,000)
- Total retirements ($1,000)
- Gross value of depreciable assets (acquisition costs), end of year ($1,000)
- Total depreciation ($1,000)
- Total rental payments or lease payments ($1,000)
- Rental payments or lease payments for buildings and other structures ($1,000)
- Rental payments or lease payments for machinery and equipment ($1,000)
- Total other operating expenses ($1,000)
- Temporary staff and leased employee expenses ($1,000)
- Expensed computer hardware and other equipment ($1,000)
- Expensed purchases of software ($1,000)

- Data processing and other purchased computer services ($1,000)
- Communication services ($1,000)
- Repair and maintenance services of buildings and/or machinery ($1,000)
- Refuse removal (including hazardous waste) services ($1,000)
- Advertising and promotional services ($1,000)
- Purchased professional and technical services ($1,000)
- Taxes and license fees ($1,000)
- All other operating expenses ($1,000)

Cluster Analysis

Based on literature review and out conceptual framework, we believe that the following variables from the dataset can be used to capture the lean characteristics of U.S. manufacturing industries.

Productivity-related variables:
- output per worker per hour
- value added per worker per hour

Variable related to green operations:
- Cost of purchased fuels (non-clean energy)
- Cost of purchased electricity (clean energy)

Inventory variables:
- Raw material inventory
- Work-in-Process Inventory
- Finished goods inventory

Variables related to capital expenditures on PPE:
- Capital expenditures for buildings and other structures
- Capital expenditures for machinery and equipment
- Capital expenditures for automobiles, trucks, etc. for highway use
- Capital expenditures for computers and peripheral data processing equipment
- Capital expenditures for all other machinery and equipment

Waste variables:
- Refuse removal including hazardous waste

To map lean operations of all U.S. manufacturing industries, we conduct a cluster analysis based on lean-related variables listed above.

In cluster analysis, a cluster is considered a group of observations such that the distance of the observation from the cluster centroid is the shortest of the distances between that observation and the centroids of all the clusters. Typically, the assignment of an observation to a cluster is a result of an iterative process during which the cluster centroids may change as more observations are added to a given cluster, and some observations previously assigned to one cluster may be reassigned to another cluster.

There are multiple ways to start cluster analysis. In this research, we follow Nag et al. (2015) to employ non-hierarchical, *k-means* clustering, a process that demonstrates stability and relatively low sensitivity to data outliers. *K-means* clustering typically starts with a user-defined or randomly generated number of *k* initial clusters.

Each six-digit NAICS industry in the dataset is assigned to one cluster with the nearest Euclidean distance to the centroid of the cluster. As one cluster gains an observation, the centroid of the cluster is recalculated, and observations may be reassigned accordingly. The iterative algorithm stops when the within-group sums of squared errors, as measured by Euclidean distances from each observation, are minimized.

Our objective is to group U.S. manufacturing industries of similar lean characteristics into clusters so that manufacturing industries in the same group have the most commonality of lean operations and manufacturing industries across different groups differ most from each other in terms of lean operations.

Case Studies

We will conduct case analysis of selected industries in each of the clusters identified in the stage of cluster analysis. We believe that a multiple case study will provide further insights into the lean operations of U.S. manufacturing industries.

DISCUSSIONS AND CONCLUSIONS

This study aims to map the lean operations of all U.S. manufacturing industries based on cluster analysis of U.S. Economic Census data. This paper has completed literature review and conceptual development. We have also collected data, identified research methodologies appropriate to our research objectives and data. It is promising when data analysis is complete at a late stage. We then shall be able to conduct further case studies to complete cluster analysis and provide further insight into the lean operations of U.S. manufacturing industries.

REFERENCES

Aqlan, F., and Al-Fandi, L. (2018). “Prioritizing process improvement initiatives in manufacturing environments”, *International Journal of Production Economics*, 196, pp. 261-268.

Arthur, J.B. (1994). “Effects of human resource systems on manufacturing performance and turnover”, *Academy of Management Journal*, Vol. 7 No. 3 pp. 670-687.

Bozarth, C.C.., Warsing, D.P., Flynn, B.B., and Flynn, E.J. (2009). “The impact of supply chain complexity on manufacturing performance”, *Journal of Operations management*, Vol., 27 No.1 pp. 78-93.

Cua, K.O., McKone, K.E., and Schroeder, R.G. (2001). “Relationships between implementation of TQM, JIT, and TPM and manufacturing performance”, *Journal of Operations Management*, Vol. 19, No. 6, pp. 675-694.

Doolen, T.L. and Hacker, M.E. (2005). “A review of lean assessment in organization: an exploratory study of lean practices by electronic manufacturers”, *Journal of Manufacturing Systems*, Vol. 24, No. 1, pp. 55-67.

Flynn, B.B., Schroeder, R.G., and Flynn, E.J. (1999). “World class manufacturing: an investigation of Hayes and Wheelwright’s Foundation”, *Journal of Operations Management*, Vol. 17 No. 3, pp.249-269.

Gara-Reyes, J.A., Kumar, V., Chaikittisilp, S., and Tan, K.H. (2018). “The effect of lean methods and tools on the environmental performance of manufacturing organisations”, *International Journal of Production Economics*, 200, pp. 170-180.

Hayes, R.H., and Wheelwright, S.C. (1984). *Restoring our Competitive Edge: Competing through Manufacturing,* Wiley, New York, NY.

Hofer, C., Eroglu, C., and Hofer, A.R. (2012). “The effect of lean production on financial performance: the mediating role of inventory leanness”, *International Journal of Production Economics*, Vol. 138, No. 2, pp. 242-253.

Ichniowski, C., Shaw, K., and Prennushi, G. (1997). “The effects of human resource management practices on productivity: a study of steel finishing lines”, *American Economic Review*, Vol. 87 No. 3, pp 291-313.

Jabbour, C.J.C., de Sousa Jabbour, A.B.L., Govindan, K., Teixeira, A.A., and de Souza Freitas, W.R. (2013), “Environmental management and operational performance in automotive companies in Brazil: the role of human resource management and lean manufacturing”, *Journal of Cleaner Production*, Vol. 47, pp. 129-140.

James -Moore, S.M., and Gibbons, A. (1997). “Is lean manufacture universally relevant? An investigative methodology”, *International Journal of Operations and Production Management*, Vol. 17, No. 9, pp. 899-911.

Koufteros, A.A., Vonderembse, M.A., and Doll, W.J. (1998). “Developing measures of time-based manufacturing”, *Journal of Operations Management*, Vol. 16 No. 1, pp.24-41.

Panizzolo, R. (1998). “Applying the lessons learned from 27 lean manufacturers – the relevance of relationship management”, *International Journal of Production Economics*, Vol. 55 No. 3, pp. 223-240.

Sakakibara, S., Flynn, B.B., Schroeder, R.G., and Morris, W.T. (1997). “ The impact of just-in-time manufacturing and its infrastructure on manufacturing performance”, *Management Science*, Vol. 43 No.9, pp. 1246-1257.

Shah, R., and Ward, P.T. (2003). “Lean manufacturing: context, practice bundles, and performance”, *Journal of Operations Management*, Vol. 21 No. 2, pp. 129-149.

Shah, R., and Ward, P.T. (2007). “Defining and developing measures of lean production”, *Journal of Operations Management*, Vol. 25 No. 4, pp. 785-805.

Swink, M., Narasimhan, R., and Wang, C. (2007). “Managing beyond the factory walls: effects of four types of strategic integration on manufacturing performance”, *Journal of Operations Management*, Vol. 25, No.1, pp. 148-164.

Taddeo, R., Simboli, A., Di Vincenzo, F., and Ioppolo, G. (2019). “A bibliometric and network analysis of Lean and Clean (er) production”, *Science of the Total Environment*, 653, pp. 765-775.

Wickramasinghe, G.L.D., and Wickramasinghe, V. (2017). “Implementation of lean production practices and manufacturing performance”, *Journal of Manufacturing Technology Management*, Vol. 28, No. 4, pp. 531-550.

Womack, J.P., and Jones, D.T. (1996). “Beyond Toyota: how to root out waste and pursue perfection”, *Harvard Business Review*, Vol. 74, September – October, pp. 140-151.

Womack, J.P., Jones, D.T., and Roos, D. (1990), *The Machine that Changed the World, Harper Perennial*, New York, NY.

Youndt, M.A., Snell, S.A. Jr and Lepak, D.P. (1996). “Human resource management, manufacturing strategy, and firm performance”, *Academy of Management Journal*, Vol. 39 No. 4, pp. 836-866.

DECISION SCIENCES INSTITUTE

Alignment of Benchmarking Performance Metrics: A Perspective on Managerial Positions

(Full Paper Submission)

Mohammad Z. Meybodi
Indiana University Kokomo
mmeybodi@iuk.edu

ABSTRACT

This focus of this article is to examine consistency of benchmarking performance metrics at various levels of an organization. The article also examines the role of managerial positions on deployment of consistent strategic and tactical benchmarking performance metrics. Statistical results show evidence of misalignment between organizational strategy and proactive development of organizational core competencies. The results indicate that managers at the higher-level positions typically place more emphasis on strategic benchmarking performance measures while managers at the lower level of the organization place more emphasis on tactical benchmarking performance metrics. Utilizing consistent benchmarking performance metrics require a thorough understanding of organizational strategy and deployment of the strategy into tactical strategies.

KEY WORDS: Benchmarking Performance Metrics

INTRODUCTION

In today's global competitive market, world-class organizations consistently utilize benchmarking to improve their competitiveness in the market by improving elements such as cost, quality, delivery, agility, and customer service. Benchmarking may be defined as the process of learning from the best-in-class organizations, determining how the best-in-class achieve those performance levels, and utilizing the best practices to their own company to achieve superior performance (Ferreira de Castro and Frazzon (2017), Watson (1992), Camp (1989), Whiting (1991)). Benchmarking is a flexible tool that can be used for gradual continuous improvement, as well as for major changes of process reengineering. Benchmarking is an effective means for learning and change because it exposes employees to new approaches, systems, and procedures (Welch (1993), Kuebler (1993)). Deming (1982) and other quality advocates (Graham (1993) and Venetucci (1992)) have strongly recommended the use of benchmarking as an essential component of continuous improvement.

The use of benchmarking as an effective quality improvement tool was started by Xerox in the early 1980's to overcome severe international competition. Also, since 1987, benchmarking has been a major component of the Malcolm Baldrige National Quality Award criteria. Since 1987, out of a total of 1000 Baldrige points, benchmarking has consistently influenced more than 500 points, Bogan and English (1994). No other quality elements, such as process management, employee involvement, and quality planning, have had such a broad influence on the Baldrige criteria than benchmarking. More recently, the practice of benchmarking is being widely used by

six sigma process, agile manufacturing, supply chain management, and for organizations seeking ISO 9000 certification (Arya et.al (2017), Faisal et.al.(2017).
Since the early 1980's, application of benchmarking in various businesses, ranging from manufacturing to health care, marketing, supply chain, human resources, and accounting has been widely reported. Harrison (1999) presents a detailed analysis of the evolution of different aspects of benchmarking activities. A comparison of the Xerox and Kodak benchmarking process has been reported by Bogan and English (1994). Although the two benchmarking methods utilize different numbers of steps, their overall logic is quite similar. Zairi and Whymark (2000) report successful results of the application of benchmarking at British Royal mail. Applications of benchmarking to world-class purchasing and to US service sectors have been reported respectively by (Newman, Hanna, and Duffett (1995) and Roth et.al. (1997)). The use of benchmarking as an effective organizational learning tool has been presented by (Tasopoulou and Tsiotras (2017), Souliotis et.al. (2017), Senge (1990), Garvin (1993), Ford and Evans (2001), Smith (1997), Hambly (1997), Watson (2001), Chen and Paetsch (1995), O'Dell and Grayson (2000), Routroy and Pradhan (2013), and Evans and Dean (2003)). A comprehensive list of legal and ethical issues of benchmarking is presented by Brue (2002), Vaziri (1992), and Bogan and English (1994)).
Although the content of the above articles is diverse, their primary focus, however, has been on short term financial and operational aspects of benchmarking. These articles generally address the technical aspects of departmental benchmarking along with limited success stories. As argued by (Furey (1987), Goldwasser (1995), Kaplan (1992), and Talluri and Vazacopoulos (1998)) effective benchmarking is more than comparative analysis of quantitative metrics from one company to another. To take full advantage of benchmarking, the benchmarking activities need to be integrated into organizational strategy and the process employs a broad range of performance measures consistent with organizational strategy.
The aim of this article is to examine consistency of organizational strategy and to investigate the impact of managerial positions on the alignment of benchmarking performance metrics. Specifically, the focus of the article is to answer the following questions:

1. Are long term corporate goals and objectives consistent with their core competencies?
2. Are long term corporate goals and objectives consistent with detecting environmental factors?
3. Are corporate competitive priorities consistent with their long-term strategies?
4. Are corporate competitive capabilities consistent with their competitive priorities?
5. Are there relationships between managerial positions and selection of benchmarking performance metrics?

BENCHMARKING AND PERFORMANCES MEASREMENTS

Total quality management (TQM), Just-in-time (JIT) systems, proliferation of new technologies, and a number of other important events during the last two decades have helped organizations to recognize the importance of benchmarking and performance measurements in managing complex processes. Managers across various industries have recognized the importance of managing processes and the truth that what gets measured is what gets managed and improves. Special focus of Malcolm Baldrige National Quality Award on benchmarking and performance measurements is a clear indication of the critical role of these elements in managing and improving organizational processes (Ferreira de Castro and Frazzon (2017), Dewan et al. (2013), Potdar et al. (2017), Kafetzopoulos, Gotzamani, and Psomas (2013)).
In the past, organizations generally used performance measurements that contributed mainly to short-term financial and technical results. How the organization achieved those results and their impact on the entire organization was unimportant Eccles and Nohria (1992). Today, managers understand that focus on short-term financial and technical results without consideration to

overall organizational strategy could produce devasting results over the long term. As a result, organizations are learning to manage the system in a way that crosses traditional departmental boundaries. In this new horizontally integrated system, organizations need to accept a long-term perspective and utilize balanced, financial and nonfinancial performance measures to carefully improve the competitiveness of the entire organization. The approach requires that benchmarking organizations develop a complete understanding of their own business strategy and deployment of the strategy into functional strategies. This process will ensure that there is a consensus within the organization about long term and short-term performance measures that are consistent with organizational mission and goals (Papke-Shields et. al. (2000), Madigan (1992)).

METHODOLOGY

A questionnaire-based survey was used to examine the above questions. The part of the survey related to this article contains a series of questions on the use of strategic and tactical benchmarking metrics. Strategic questions are concerned about organizational mission and goals, as well as attitude toward customers, competition, technology, globalization, quality, and employee developments. Tactical items are related to specific technical performance measures such as cost, quality, and delivery.
The target population for this study consisted of manufacturing firms in the Midwestern United States. A sample of manufacturing firms with more than 50 employees was chosen from Midwestern manufacturers' directories. The sample covers organizations in a variety of industries ranging from fabricated metal, communication, electronics, automotive, toots, chemicals, rubber, and paper products. In addition to general organization and managerial profile items, the survey contained a series of questions regarding organizational goals and objectives, competitive priorities, manufacturing performance objectives, and manufacturing action plans.
The survey data indicates majority of respondents had various high level managerial positions from organizations with less than 500 employees. Presidents and vice presidents accounted for 29% and plant managers accounted for 30% of the sample. About 35% of the sample had other managerial positions such as operations/production managers, quality managers, and the remaining 6% were production line supervisors. In terms of manufacturing experience, about 28% of the respondents had between 10 to 20 years and 60% had more than 20 years of manufacturing experience.

RESULTS

Table 1 shows the ranking of the mean importance score for each element of corporate goals and objectives and. Table 2 also shows the mean importance rating for corporate core competencies and elements of environmental factors. The respondents were asked to rate each element based on the degree of importance (1=low importance, 5=high importance) to their company for the next five years.

Table 1. Mean Importance Rating for Corporate Goals and Objectives
(1=low importance, 5=high importance)

Element	Mean	SD
Build market share	4.76*	1.27
Maximize profits	4.61*	1.32
Focus on customer satisfaction	4.52*	1.15

SD = Standard Deviation, * = Statistically larger than the Means in Table 2 at α = 0.05

Table 2. Mean Importance Rating for Core Competencies and Environmental Factors

Element	Mean	SD
Build and exploit core competencies	3.74	1.39
Understand competitors' strategy	4.14	1.21
Understand global strategies	3.88	1.37
Understand technology	3.91	1.32

SD = Standard Deviation; The means are statistically smaller than the means in Table 1 at α = 0.05

Table 1 indicates that the respondents' top three corporate goals and objectives are building market share, maximizing profits, and focusing on customer satisfaction. Being in a better competitive position with respect to quality and customer satisfaction are reasonable explanation for market expansion and profit making. However, the mean rating for the four elements of core competencies and environmental factors in Table 2 are significantly lower than the mean rating of the three elements of corporate goals and objectives in Table 1 at 0.05 level of significance. This is perhaps an indication of traditional reactive strategy in which the primary focus of managers is on marketing and financial goals. Understanding external environmental factors such as competition, global issues, technology, and development of core competencies to effectively deal with these factors is considered to be secondary. This is a rather disturbing posture because in today's global market world-class organizations focus more on building core competencies than on achieving marketing and financial goals. They develop core competencies first, then utilize a proactive strategy and look for opportunities to exploit their core competencies to achieve a competitive advantage. Understanding the causes for such strategic misalignment between corporate goals and objectives and proactive development of core competencies is extremely important.

Table 3. Mean Importance and Capability Rating for Competitive Priorities
(1=low importance, 5=high importance) (1=weak strength, 5=strong strength)

	Importance		Capability			
Element	Mean	SD	Mean	SD	t-value	p-value*
Product reliability	4.69	1.16	3.56	1.24	4.21**	0.000
Conformance quality	4.58	1.28	3.84	1.14	1.34	0.090
Delivery reliability	4.47	1.24	3.78	1.32	2.57**	0.005
Product customization	4.35	1.36	3.42	1.09	3.46**	0.000
NPD speed	4.29	1.23	3.35	1.36	3.50**	0.000
Price	4.16	1.34	3.96	1.25	0.75	0.230
Fast delivery	4.03	1.32	3.82	1.19	0.78	0.210
Performance	3.98	1.22	4.22	1.28	0.89	0.190
Service after sales	3.84	1.44	4.18	1.23	1.27	0.100
Volume flexibility	3.62	1.31	4.32	1.37	2.61**	0.005

* Small p-value indicates the difference between two means is statistically significant
** Mean strength is less than mean importance at alpha = 0.05

Table 3 shows the mean importance and capability rating for each element of competitive priorities. The importance rating is similar to the one stated earlier. From Table 3, the respondents ranked product reliability, conformance quality, delivery reliability, product customization, and new product development speed as the top five important competitive priorities. The ranking of product reliability and conformance quality as the top two competitive priorities is consistent with corporate goals and objectives. It indicates that managers believe that quality factors are still important elements of competitive advantage. However, the ranking of delivery reliability, product customization, and new product development speed as the next three competitive priorities indicate that the respondents also believe on the importance of time based competition and product customization. Overall, one can argue that the importance ranking of top five competitive priorities is consistent with recent manufacturing literature and with organizational strategy of market expansion, profitability, and customer satisfaction. Table 3 also shows that low price as an element of competitive priorities ranked sixth. This rather interesting result indicates, unlike traditional thinking, the responding managers believe that low price is no longer the primary elements of competitive advantage. Relative low ranking of price along with the last four competitive priorities is perhaps an indication that these elements represent order qualifiers and the top five factors are order winners. In other words, competitive market considers the last five competitive priorities as given. To attract customers, organizations need to perform on the basis of top five competitive elements.

To understand organizational competitive capabilities, the respondents were asked to rate relative competitive capabilities of their organization with respect to the competitors who are doing best in that area. A five-point scale, where 1 corresponds to weak and 5 corresponds to strong, is used to indicate managers' perceptions of the company's current competitive capability relative to the best competitors. The mean competitive capability scores for each element of competitive priorities are shown in Table 3. Statistical tests indicate that, for the top five competitive priorities, the mean competitive capability is significantly lower than the mean importance. This is evident because with the exception of conformance quality, the p-value for other four competitive priorities is less than 0.005. This indicates, although managers ranked product reliability, conformance quality, delivery reliability, product customization, and new product development speed as the top five important competitive priorities, organizational competitive capability for those elements, however, is not that strong. On the other hand, the mean competitive capability for the last three competitive priorities is greater than the mean importance indicating that managers believe their competitive capabilities in the areas of performance quality, service after the sales, and volume flexibility are quite strong. This imbalance between importance and actual capability for the top five competitive priorities is perhaps a critical area that needs to be investigated.

Table 4. Strategic and Tactical Benchmarking Performance Elements

Strategic Elements	**Tactical Elements**
Develop mission and goals (DMG)	Reduce percent defects (RPD)
Develop core competencies (DCC)	Reduce percent errors (RPE)
Understand competitors' strategies (UCS)	Reduce labor costs (RLC)
Develop global strategies (DGS)	Reduce materials costs (RMC)
Develop technology strategies (DTS)	Reduce overhead costs (ROC)
Focus on customer satisfaction (FCS)	Reduce inventory costs (RIC)
Adopt TQM philosophy and practices (TQM)	Reduce set-up/changeover costs (RSUC)
Change organizational culture (COC)	Increase labor utilization (ILU)
Improve interfunctional communication (IIFC)	Increase equipment utilization (IEU)
Improve employee training (IET)	Improve process capability (IPC)
Improve employee empowerment (IEE)	Improve quality circle practices (IQCP)
Improve employee team work (IETW)	Utilize quality improvement tools (UQIT)
Install continuous improvement (ICI)	Utilize statistical process control charts
Adopt quality at the source (AQS)	Increase delivery speed (IDS)
Improve supply chain management (ISCM)	Increase product development speed
Improve supplier relationships (ISR)	Reduce manufacturing lead time (RMLT)
Eliminate wastes (ELW)	Increase delivery reliability (IDR)
Reengineer new product development (RNPD)	Increase new product customization

Table 4 shows the listing of strategic and tactical benchmarking metrics. Table 5 and 6 lists respectively the mean importance score for strategic and tactical benchmarking metrics. Each Table shows the mean and the standard deviation of importance rating for two managerial positions. In Tables 5 and 6, the responses from the presidents, vice presidents, and plant managers are grouped under managerial positions at the higher level of the organization and the responses from operations/production managers, quality managers, and supervisors are grouped under managerial positions at the lower level of the organization.

Table 5. Ratings for Strategic Benchmarking Performance Elements
(1=low importance, 5=high importance)

	Managerial Positions				
	Higher		Lower		
Element	Mean	SD	Mean	SD	t-value
DMG	4.63	1.92	3.56	2.11	2.34*
DCC	4.32	1.83	3.24	1.74	2.70*
UCS	4.65	1.76	3.76	1.85	2.21**
DGS	4.51	1.85	3.52	1.63	2.54*
DTS	4.49	1.68	3.46	1.75	2.69*
FCS	4.72	1.64	4.53	1.67	0.51
TQM	4.62	1.55	3.66	1.64	2.69*
COC	4.32	1.87	3.34	1.62	2.51*
IIFC	4.46	1.78	3.82	1.83	1.59
IET	3.92	1.66	4.20	1.47	0.80
IEE	3.75	1.45	4.14	1.69	1.33
IETW	4.51	1.62	3.78	1.85	1.88**
ICI	4.57	1.34	4.43	1.95	0.37
AQS	4.34	1.40	4.53	1.68	0.55
ISCM	4.47	1.87	3.67	1.48	2.12**
ISR	4.56	1.74	3.98	1.78	1.47
ELW	4.58	1.67	3.55	1.57	2.84*
RNPD	4.46	1.66	3.53	1.52	2.61*

SD= standard deviation, * = significant at 0.01, ** = significant at 0.05

In Tables 5 and 6, the respondents were asked to rate each element based on the degree to which they believe the factors are important (1=low importance, 5=high importance) to be used for benchmarking performance metrics. Table 5 indicates that overall managers at the higher level positions rated strategic elements higher than the managers at the lower level positions. This is evident because, with the exception of two, the mean ratings for these elements are above 4.00. Also, statistical tests indicate that out of 18 tests, 11 were statistically significant at least at a 0.05 level of significance; meaning for the strategic elements managers at the higher level positions rated these elements significantly higher than the managers at the lower level positions. For the strategic elements such as customer satisfaction, inter-functional communication, employee training, employee empowerment, continuous improvement, quality at the source, and supplier relationships, the mean ratings for managers at the higher level positions were not significantly different than the ratings for mangers at the lower level positions. One possible explanation for such result would be the popularity of these elements. Since these are well known TQM elements, managers at both levels believe on the importance of these strategic benchmarking elements. However, believing on the part of the managers at the lower levels does not necessarily translate to consistent actions. It is interesting to note that managers at the lower level positions rated the strategic elements employee training, employee empowerment, and quality at the source higher than the managers at the higher level positions.

This result was expected because managers at the lower levels are closer to employee related issues than managers at the higher levels.

Table 6. Ratings for Tactical Benchmarking Performance Elements
(1=low importance, 5=high importance)

	Managerial Positions				
	Higher		Lower		
Element	Mean	SD	Mean	SD	t-value
RPD	4.02	1.83	4.73	1.87	1.74**
RPE	4.07	1.65	4.75	1.86	1.75**
RLC	3.65	1.92	4.63	1.72	2.40*
RMC	3.54	1.82	4.58	1.69	2.65*
ROC	4.22	1.75	4.71	1.58	1.25
RIC	4.13	1.62	4.75	1.69	1.82**
RSUC	4.52	1.73	4.69	1.67	0.45
ILU	4.64	1.48	4.42	1.83	1.13
IEU	3.83	1.67	4.68	1.76	1.72**
IPC	4.13	1.59	4.76	1.55	1.79**
IQCP	3.63	1.78	4.71	1.58	2.87*
UQIT	4.05	1.59	4.73	1.93	1.72**
USPC	4.08	1.73	4.78	1.61	1.87**
IDS	4.22	1.68	4.37	1.72	0.39
IPDS	4.23	1.71	4.38	1.57	0.41
RMLT	4.04	1.84	4.78	1.73	1.85**
IDR	4.66	1.42	4.14	1.76	1.45
INPC	4.55	1.72	3.62	1.81	2.36*

SD= standard deviation, * = significant at 0.01, ** = significant at 0.01

Table 6 shows, unlike strategic elements, overall managers at the lower level positions rated tactical factors higher than the managers at the higher level positions. With the exception of one, the mean ratings for managers at the lower level positions are above 4.00. Also, statistical tests indicate that out of 18 tests, 12 were statistically significant at least at a 0.05 level of significance, meaning for tactical elements, managers at the lower level positions rated these factors significantly higher than the managers at the higher level positions. This result is consistent with manufacturing literature because managers at the lower levels positions have tendency to focus more on tactical cost cutting and quality improvement measures. This is perhaps due to miscommunication with the higher level managers or the result of inconsistent evaluation and reward system for managers at the lower levels. That is, regardless of organizational strategy, managers at the lower levels are often rewarded based on their cost cutting measures or efficiency in capacity utilization. The ratings for the tactical elements reducing overhead and set-up costs, increasing delivery and product development speed, and increasing delivery reliability for the two managerial levels were not significantly different. Perhaps due to popularity of these elements, managers at both levels believe on the improvement of these tactical benchmarking elements.

CONCLUSION

This article demonstrates how understanding organizational strategy is crucial to improve the effectiveness of benchmarking metrics. The article investigates the impact of managerial positions on the deployment of strategic and tactical benchmarking performance metrics. Five questions were asked to investigate the consistency of the decisions and to examine the relationship between managerial positions on the selection of strategic and tactical benchmarking performance metrics. Summary of the results indicate: 1) Possible misalignment between corporate goals and objectives and proactive detection of environmental factors. 2) Inconsistencies between corporate goals and objectives and development of corporate core competencies to effectively deal with the environmental factors. 3) Potential mismatch between organizational competitive priorities and their core competencies. 4) Overall, managers at higher level positions rated strategic elements significantly higher than the managers at the lower level positions. 5) Managers at the lower level positions rated tactical elements higher than the managers at the higher level positions. This result is consistent with manufacturing literature because lower level managers often have tendency to focus more on tactical cost cutting and quality improvement measures. 6) For strategic elements such as customer satisfaction, interfunctional communication, employee training, employee empowerment, continuous improvement, quality at the source, and supplier relationships the mean ratings for managers at the higher level positions were not significantly different than the mean ratings for managers at the lower level positions. Since these are well known TQM elements, perhaps managers at both levels believe on the importance of these strategic benchmarking factors. 7) The mean ratings for tactical elements of reducing overhead and set-up costs, increasing delivery speed and delivery reliability for the two managerial levels were not significantly different. Since these are well-known just-in-time elements, perhaps managers at both levels believe on the improvement of these tactical benchmarking elements.

In summary, the article showed evidence of misalignment, inconsistencies, and lack of consensus among managers at different levels of organization in choosing benchmarking performance metrics. To be effective, managers must develop a complete understanding of their own strategy and choose long term and short term benchmarking performance metrics that are consistent with organizational strategy.

REFERENCES

Arya, V., Sharma, P., Singh, A. and P.T.M. DeSilva (2017) “An exploratory study on supply chain analytics applied to spare parts supply chain”, *Benchmarking: An International Journal*, Vol. 24, No. 6, pp. 1571 – 1580.

Bogan, C. E. and English, M. J. (1994), Benchmarking for Best Practices, McGraw Hill, New York.

Brue, G. (2002), Six Sigma for Managers, McGraw Hill, New York.

Camp, R. C. (1989) “Benchmarking: The Search For Industry Best Practices That Lead to Superior Performance”, *Quality Progress*, pp. 285-286.

Chen, I. J. and Paetsch, K. (1995) “Benchmarking: Quality Managers’ Perspective”, Proceedings of the Decision Sciences Institute, pp. 1623-1625.

Deming, W. E. (1982), Quality Productivity and Competitive Position, MIT Center for Advanced Engineering Study, Cambridge, MA.

Dewan, M. Z. I. Zunder, T.H. and Jorna, R. (2013) “Performance evaluation of an online benchmarking tool for European freight transport chains”, *Benchmarking: An International Journal*, Vol. 20 No. 2, pp. 233 – 250.

Eccles, R. G. and Nohria, N. (1992), Beyond the Hype: Rediscovering the Essence of Management, Boston: Harvard Business School.

Evans, J. R. and Dean, J.W. (2003), Total Quality Management, Organization and Strategy, 3rd edition, Thomson/Southwestern.

Faisal, M.N., Al-Esmael, B. Sharif, K.J. (2017), “Supplier selection for a sustainable supply chain: Triple bottom line (3BL) and analytic network process approach”, *Benchmarking: An International Journal*, Vol. 24, No. 7, pp. 1956 – 1976.

Ford, M. W. and Evans, J. R. (2001) “Baldrige Assessment and Organizational Learning: The Need for Change Management”, *Quality Management Journal*, Vol. 8, No. 3, pp. 9-25.

Ferreira de Castro, V. and Frazzon, E.M. (2017), “Benchmarking of best practices: an overview of the academic literature”, *Benchmarking: An International Journal*, Vol. 24, No. 3, pp. 750 – 774.

Furey, T. R. (1987) “Benchmarking: The Key to Developing Competitive Advantage in Mature Markets”, *Planning Review*, Vol. 15, pp. 30-32.

Garvin, D. (1993) “Building a Learning Organization”, Harvard Business Review, pp. 41-48.

Goldwasser, C. (1995) “Benchmarking: People Make the Process“, *Management Review*, Vol. 84, pp. 39-45.

Graham, J. (1993) “Benchmarking: What You Need to Make it Work*”, Purchasing*, Jan. 14, pp. 63-75.

Hambly, L. (1997) “Sun Microsystems Embeds Quality into Its DNA”, *The Quality Observer*, pp. 16-20, 45.

Harrison, N. J. (1999) “A Research Agenda For Benchmarking”, Proceedings of the Decision Sciences Institute, p. 1394

Kafetzopoulos, D. Gotzamani, K. and Psomas, E. (2013) "Quality systems and competitive performance of food companies", *Benchmarking: An International Journal*, Vol. 20 No. 4, pp.463 - 483

Kaplan, R. S. and Norton, D. P. (1992) “The Balance Scorecard-Measures that Drive Performance”, *Harvard Business Review*, pp. 71-79.

Kuebler, A. (1993) “The Quest for Excellence 5”, Annual Conference of Baldrige Award Winners”, AT&T Universal Card Services Presentation, Feb. 15-17.

Madigan, J. M. (1992), Measures Matrix Chart, Eastman Kodak Company, a paper presented to Strategic Planning Institute.

Newman, R.W., Hanna, M., and Duffett, W. (1995) “ World Class Purchasing: A Benchmarking of the Best”, Proceedings of the Decision Sciences Institute, pp. 1617-1619.

O’Dell, C. and Grayson, C. J. (2000) “Identifying and Transferring Internal Best Practices”, APQC White Paper.

Papke-Shields, K. E., Malhotra, M. K., and Grover, V. (2002) “Strategic Manufacturing Systems and Their Linkage to Planning System Success”, *Decision Sciences*, Vol. 33, No. 1, pp. 1-30.

Potdar, P.K. Routrov, S. and Behera, A. (2017), “Analyzing the agile manufacturing barriers using fuzzy DEMATEL”, *Benchmarking: An International Journal*, Vol. 24, No. 7, pp. 1912 – 1936.

Roth, A. V., Chase, R. B., Voss, C., Rose, E., and Menor, L. (1997) “Benchmarking US Services: Results and Implications”, Proceedings of the Decision Sciences Institute, pp. 1510-1512.

Routroy, S., and Pradhan, S. K. (2013) "Evaluating the critical success factors of supplier development: a case study", *Benchmarking: An International Journal*, Vol. 20 No 3, pp.322 - 341

Senge, P. M. (1990), The Fifth Discipline: The Art and Practice of Learning Organization, Doubleday Currency, New York.

Smith, G. P. (1997) “A Change in Culture Brings Dramatic Quality Improvement”, *Quality Observer*, January 1997, pp. 14-15, 37.

Souliotis, A. Giazitzi, K. and Boskou, G. (2017),) "Benchmarking between vegetable suppliers in Greece", *Benchmarking: An International Journal*, Vol. 24, No. 6, pp. 1649 – 1662.

Talluri, S. and Vazacopoulos, A. (1998) “Efficiency Evaluation and Benchmarking in Small and Medium Enterprises”, Proceedings of the Decision Sciences Institute, pp. 1421-1423.

Tasopoulou, K. and Tsiotras, G. (2017), "Benchmarking towards excellence in higher education", *Benchmarking: An International Journal*, Vol. 24, No. 3, pp. 617 – 634.

Vaziri, H. K. (1992) “Using Competitive Benchmarking to Set Goals”, Quality Progress, November, pp. 81-85.

Venetucci, R. (1992) “Benchmarking: A Reality Check for Strategy and Performance Objectives”, *Production and Inventory Management Journal*, Vol. 33, No. 4, pp. 32-36.

Watson, G. H. (2001) “ Cycles of Learning: Observations of Jack Welch”, *Six Sigma Forum Magazine*, Vol. 1, No. 1, pp. 13-17.

Watson, G. H., (1992), The Benchmarking Workbook: Adopting Best Practices for Performance Improvements, Productivity Press, Portland, Oregon.

Welch, J. (1993) “Jack Welch’s Lessons for Success”, Fortune, January 25th, pp. 86-93.

Whiting, R. (1991) “Benchmarking: Lessons From Best-in-Class”, *Electronic Business*, Vol. 17, pp. 128-134.

Zairi, M. and Whymark, J. (2000) “The Transfer of Best Practices: How to Build Culture of Benchmarking and Continuous Learning- Part 1”, *Benchmarking: An International Journa*l, Vol. 7, No. 1, pp. 62-78.

DECISION SCIENCES INSTITUTE

Achieving Customer Specifications Through Lean Six Sigma: A Design of Experiment Case Study

Brian Vo
University of Bridgeport
brianvoct@gmail.com

Elif Kongar
University of Bridgeport
kongar@bridgeport.edu

Manuel Francisco Suárez Barraza
Universidad de las Americas Puebla | UDLAP
manuel.suarez@udlap.mx

ABSTRACT

Lean Six Sigma (LSS) rises as a viable alternative with its ability to aid in transforming a culture through continuous improvement and saving on cost. In order to provide data on its application and impact, this case study considers an injection molding process and demonstrates how to improve the quality of manufactured plastic components, eliminate waste, reduce the component variation and increase the output yield by applying Lean Six Sigma methodology. Design of Experiment (DOE) is used to identify the key process input variables (KPIVs) which have significant influence on the output of the plastic component.

KEYWORDS: Lean Six Sigma, customer satisfaction, design of experiment, molding, technology management.

INTRODUCTION

As global competition continues to rise, the pressure to improve efficiency becomes more and more intense. Under constant pressure to achieve more with less, only the firms that deliver superior value and produce world class quality in its all processes will remain competitive in the market. Corporations that can reduce costs, eliminate waste, minimize defects, decrease lead times, improve quality, enhance customer satisfaction, and transform a culture through continuous improvement will be the ones to operate successfully in the long term (Zhang et al., 2016b). To ensure their sustainability many of these corporations today are turning to process improvement methodologies such as Lean Six Sigma to improve and sustain their business operations. Lean Six Sigma is a continuous improvement methodology that combines two of the most powerful improvement disciplines. Lean was developed by Toyota Motor Corporation (Jiang et al., 2004) and Six Sigma was developed by Motorola (Dakhli et al., 2016). Lean focuses on elimination of waste, non-value added activities and increasing process efficiency through a focus on improvement in speed and cost. Six Sigma on the other hand, focuses on reduction of process variation and drastic elimination of defects with an emphasis on improvement in quality (Ben Ruben et al., 2017). The two methodologies are harmonized, and there is strong evidence that total improvement in a process is greater if Lean and Six Sigma are implemented together (Zhang et al., 2016b).

Lean Six Sigma is a well-established methodology for improvement the organizational effectiveness. The goals of Lean Six Sigma include building a world-class organization, developing leaders, creating a culture of continuous improvement (Erdil et al., 2018). Lean Six Sigma consists of five basic phases, Define, Measure, Analyze, Improve, and Control (Desai, D. A., 2006), and hence the process is also known as DMAIC(Antony et al., 2017).

Lean Six Sigma has been implemented very successfully in many large corporations such as Honeywell, Johnson & Johnson, General Electric, Boeing, Merck, Aetna, and Quest Diagnostics (Chakrabarty and Tan, 2007) and in some small and medium sized manufacturing enterprises (SMEs) (Kumar et al., 2006).There is, however, still a gap when it comes to the applications of Lean Six Sigma tools (Vo et al., 2019), specifically on how Lean Six Sigma is implemented in small organizations (Lande et al., 2016). The objective of this research is to investigate the steps and effects of Lean Six Sigma program on a consumer packaging company which is located in the United States.

LITERATURE REVIEW

Lean Manufacturing

Lean manufacturing is considered to be a long term philosophy that aims at continuous growth by creating value for the customer, societies, and economy by decreasing costs, improving quality and delivery through the elimination of waste (Houborg, 2010). Lean production system, aka Toyota Production System, is a world renowned production system developed and practiced by Toyota for a long time (Jiang et al., 2004). As Jones et al. (1999) points out the principles of 'lean thinking' aim at eliminating waste and focusing on those activities which add value for the customers (Jones et al., 1999). Waste, often known as muda, consists of seven types: Defects, Overproduction, Transportation, Waiting, Inventory, Motion, and Processing (Sternberg et al., 2013). Apparently, eliminating of waste relatively looks simple and straightforward, but it is often difficult to identify what is waste and what is not in most organization (Ghosh, 2013).

Six Sigma

Six Sigma is a management philosophy developed by Motorola to improve processes by reducing process variation and eliminating defects (Antony et al., 2017). The Greek letter sigma is used to denoted variation from a standard. In order for the organization to achieve Six Sigma Process, it can't produce more than 3.4 defects per million opportunities or 99.99966% defect free over the long-term. Six Sigma methodology makes use of many tools and techniques to enable continuous improvement. The most frequent statistical tools used are pareto chart, histogram, gage R&R, attribute agreement analysis, and process capability (Prashar, 2014).

Lean Six Sigma

Lean Six Sigma is a process improvement approach that consists of tools and methods of both Lean and Six Sigma (Zhang et al., 2016a) and aims to provide customers with the best possible quality, cost, and speed (Dakhli, et al., 2016). In the recent years, there has been a noticeable increase in LSS popularity and deployment in manufacturing and service industry (Lim et al., 2015). A synergy between Lean and Six Sigma would maximize the strengths and minimize the weakness of both methodologies when used alone (Timans et al., 2012)

Design of Experiments (DOE)

Design of Experiments (DOE) and statistical analysis of variance approach were developed by Fisher in 1930 (Box, 1980). DOE is commonly used to optimize industrial process (Dowlatshahi, 2004) and in this study it is utilized to optimize the injection process. Optimization is a mathematical approach to tackle the problems and widely used in many different areas. A full factorial design DOE with replicate and center points are suggested for a study if applicable. The next section explains how to carry the DOE design and analysis needed to identify the factors that have the most influence on the injection molding process (Timans et al., 2014).

METHODOLOGY

The purpose of this research is to explore Six Sigma tools and techniques that were applied to improve the process capability of torque retention for a cosmetic pump. The case study describes the impact of Lean Six Sigma on reducing component defects by optimizing the injection molding process. Rigorous use of the Lean Six Sigma methodology and application of each step of the DMAIC roadmap are essential in process improvement implementation. The design of the study and the type of data are clearly defined and described for the injection molding process. The statistical analysis, interpretations of the statistical results and graphical tools are presented along with the conclusions and recommended actions at the end of this article.

The execution of Define process for this project included approved project charter, a process map/SIPOC, and project Y's. During Measure phase, C&E matrix, FMEA and project baselines were established. Analysis included basic statistics and graphical summary, and hypothesis testing. The Improve phase included the identification of the solutions implemented above. Finally, pFMEA, validation and verification, and control plan were updated during the Control Phase.

CASE STUDY

The company in this case study was founded over half a century ago and is publicly held. It provides consumers all over the world with state-of-the art packaging materials in the beauty, health care, personal care, home care, food, medical, and prescription drug segments. The product in focus is a dispensing pump system which comprises of dust cap, actuator, insert, fixture, gasket, cartridge, and dip tube. All plastic components were molded and assembled into a finished pump by the company which recognizes as a leader in the global packaging and dispensing industry.

Recently, the company received a high count of customer complaints related to dust cap separated from the pump during transportation. In addition, it has been noticed that a number of lots failed to meet a minimum torque retention at the final inspection. The project was initiated to decrease the customer complaints and reduce cull/rework related defects. The Lean Six Sigma methodology was applied to identify the root causes of the defects and unitized design of experiment (DOE) to improve the process to achieve higher customer satisfaction and to reduce operational cost.

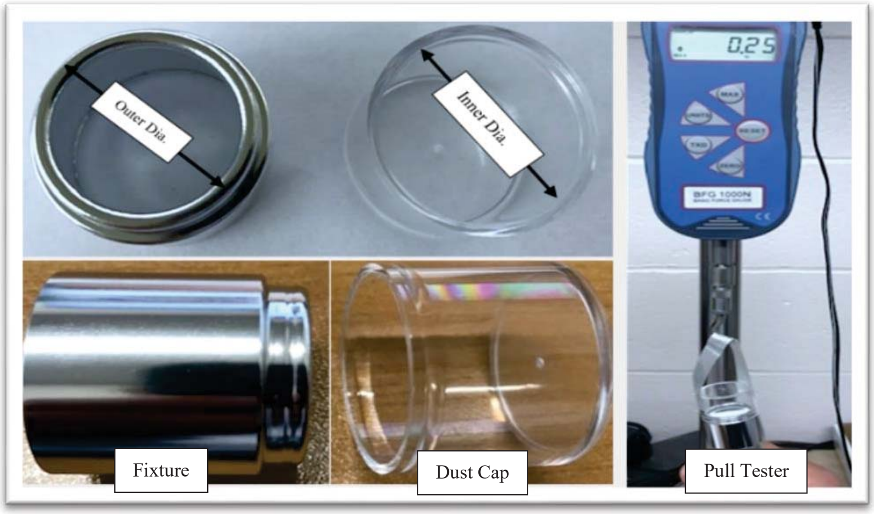


Fig. 1: Product

Table 1: Product specifications

Torque retention specification	1.5 – 4.0 lbf
Product scrap rate (OOS)	15%

The resin material which used to mold the dust cap is Polypropylene (PP) with 35 percent glass reinforced which is a high impact grade PP resin for injection molding. The mold has 32 cavities in which every cycle of the injection molding process produces 32 dust caps. The process of injection molding of the dust cap comprises a combination of heat, pressure and time parameters. The resin is compressed in a heated cylinder by means of a reciprocating screw and injected into a mold cavity. The mold is cooled to solidify the part before it is ejected out of the cavity.

Define Phase

The first phase of the DMAIC methodology is to define the problems through the voice of the customer/business, SIPOC, and project map. The critical to quality (CTQ) characteristics in the study is an interference fit of two components (dust cap and fixture) in which the inner diameter of the dust cap slightly exceeds the outer diameter of the fixture (Fig. 1). It is critical that the dust cap must be applied tight enough to maintain a good seal to the fixture, and there is a gentle pull onto the dust cap to remove it from the fixture. The customer specification of cap removal (torque retention) is from 1.5 lbs. to 4 lbs.
Given the mold design and material resin, the goals of the project are defined as follows in the Define phase:

- Identify the process parameters that have impact on torque retention
- Establish a range of the process parameters for those significant parameters.
- Determine optimal set-points for the process parameters.

Furthermore, a SIPOC map of an injection molding process, the Suppliers (the 'S'), the Inputs (the 'I'), the Process (the 'P'), the Outputs (the 'O'), and the Customers (the 'C'), was used to identify all pertinent elements of a process improvement project before work began. The project team consisted of the facility manager, the Director of Quality, the Manager of the Advanced Quality Planning, the floor supervisor, and the machine operator. The map allowed the team to identify the areas of opportunity for improvements and determined where resources should be allocated.

Suppliers (S)	Inputs (I)	Process (P)	Outputs (O)	Customers (C)
Resin supplier Colorant Supplier	Materials Machine Personnel Equipment Mold	W/O & Schedule ↓ Pickup Mold from Tool ↓ Hang Mold ↓ Mold & Machine ↓ Mold & Machine ↓ First Run ↓ Visual Inspection 1 ↓ Run ↓ Final Inspect ↓ Packaging	Molded components Quality records	Assembly Manufacturing

Fig. 2: SIPOC

Measure Phase

In order to determine the mutual relationships and contributing causes, a cause and effect (C&E) diagram was created for the injection molding process. The C&E identified which process inputs of the molding process had the most impact on the customer requirements. The Key Process Output Variables (KPOVs) from the process mapping were streamlined into project goals, and the numbers from 1 to 10 were scaled the importance to the project. Each of the input process steps was rated on a scale of 0, 1, 3, and 9. The highest scores were then determined to be included in the process failure modes and effects analysis (pFMEA). This funnel tool is a highly effective way of focusing on priorities of the project. The pFMEA, was then completed for the injection molding process to identify and evaluate the failure modes. The severity, occurrence, and detection multiplied together to form the risk priority number (RPN) which was used to rank and to focus the team on specific action items (Sekar and Santhosh, 2012). Any items with high RPNs, the recommended actions were identified to reduce the RPN, and the control plan was created to sustain corrective actions. Once the actions were implemented; the RPN was recalculated.

As part of the analysis, 100 samples were randomly collected and measured for the torque retention quality characteristic from the current multiple production lots. The torque retention data was then used in the process capability analysis to identify whether the process was producing results outside the specification limits to prevent further production of unacceptable products. Based on the results, Cp and Cpk indices are 1.64, and 0.41 respectively. These values indicated that the process is capable but is not centered and not performing within the specification limits, the process needs to be improved. Consequently, a high number of pumps failed to meet the torque retention specification as shown in Fig. 3.

Table 2: Summary Statistics

Variable	Count	Mean	StDev	Min	Median	Max	Data Normal?	Cpk
Torque Retention (N*m) (Current Production)	100	1.8	0.24	1.2	1.8	2.3	Yes	0.41

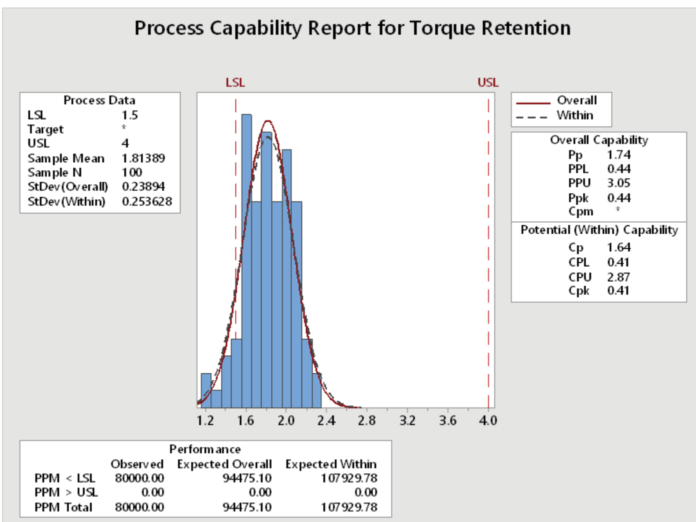


Fig. 3: Cpk Analysis of Torque Retention before Implementing DMAIC Methodology

Analyze Phase

In the analysis phase, the team identified the potential causes of variations in the process that affected the retention force using a cause and effect diagram. Figure 4 depicts all possible major causes that involve the injection molding process, i.e., temperature, pressure, speed, and time, and speed. Based on the resin material, mold design, flow mold analysis, pFMEA, and the data gathered from additional statistical analyses, the team determined four process parameters to be included in the design of experiments.

- Barrel Temperature (Tbarrel)
- Injection Pressure (Pinjection)
- Injection Speed (Sinjection)
- Hold Time (Thold)

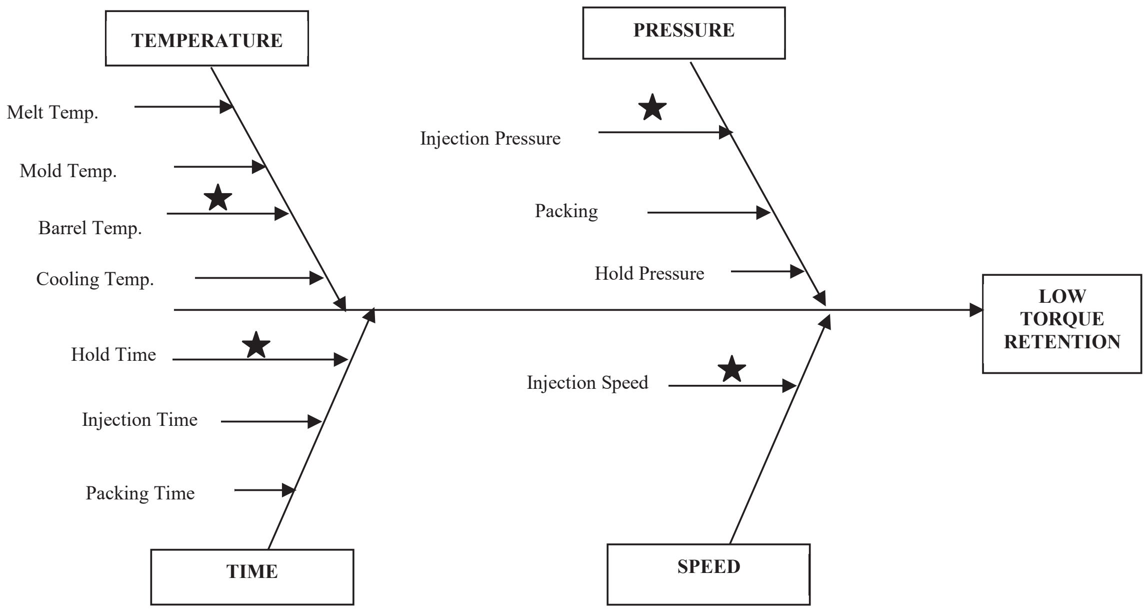


Fig. 4: Cause and Effect Diagram

Improve Phase

In order to study the effects of process parameters on the quality characteristics of interest, a design of experiment was carried out. A full factorial design consisted of four factors with two levels (low and high). Due to the production constraint, one replicate was applied in this study. A full factorial of 16 runs ($2^4 = 16$) was set. The quality characteristic of interest for this study was torque retention measured in pound-force unit. The torque retention will be the response output in the study.

Table 3: Level Values of Input Factors

Factors		Levels		Units
		Low	High	
A	Barrel Temperature (Tbarrel)	215	235	Degree
B	Injection Pressure (Pinjection)	30	45	PSI
C	Injection Speed (Sinjection)	40	50	ml/sec
D	Holding Time (Thold)	1.5	2.5	second

Table 4: Full Factorial Format with Torque Retention Data Applied

StdOrder	RunOrder	Barrel Temperature	Injection Pressure	Injection Speed	Holding Time	Torque Retention (lb-f)
16	1	235	45	50	2.5	2.3
12	2	235	45	40	2.5	2.3
2	3	235	30	40	1.5	2.4
1	4	215	30	40	1.5	0.7
6	5	235	30	50	1.5	2.8
13	6	215	30	50	2.5	0.9
8	7	235	45	50	1.5	2.9
15	8	215	45	50	2.5	2.5
10	9	235	30	40	2.5	2.6
7	10	215	45	50	1.5	2.6
3	11	215	45	40	1.5	2.5
11	12	215	45	40	2.5	2.5
5	13	215	30	50	1.5	0.6
4	14	235	45	40	1.5	2.7
9	15	215	30	40	2.5	0.8
14	16	235	30	50	2.5	2.7

Table 4 shows a total of 16 runs for each setting condition. Each trial was carried out by allowing the machine to run at least 10 cycles to stabilize the process condition before samples to be collected for testing. For instance, the first run order of the experiment, the process parameters were set for Barrel Temperature at 235, Injection Pressure at 45, Injection Speed at 50, and Hold Time at 2.5. All samples from the 16 runs were collected and stored for at a minimum of 24 hours at the ambient environment. The dust cap was assembled onto the pumps and then tested using the pull tester. The results were calculated and analyzed using Minitab v18.

To validate the data before analyzing, normality assumption was selected and checked for the error variables have a normal distribution. The normal probability plot was generated and showed that the data was normally distributed (Fig. 5). This plot illustrates that normality assumption is met for torque retention.

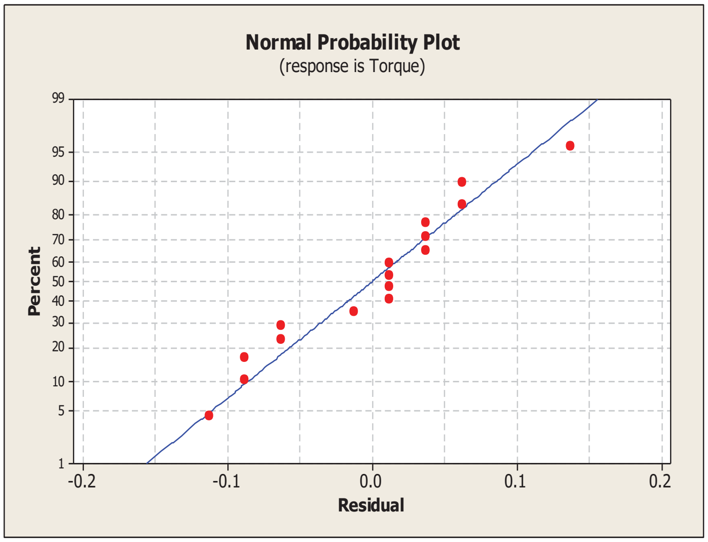


Fig. 5: Normal Probability Plot

The first step in the analysis was to identify any main and interaction factors that have influence the mean of torque retention. According to the Pareto chart of the standardized effects in Fig. 6, factors A (Tbarrel), B (Pinjection), AB (Tbarrel and Pinjection), BD (Pinjeciton and Tholding), and AD (Tbarrel and Tholding) have significant effect on the response. Since C, D, AC, CD, and BC terms are insignificant, these terms can be removed from the model. The design can be reanalyzed by choosing only significant factors A, B, AB, BD, and AD.

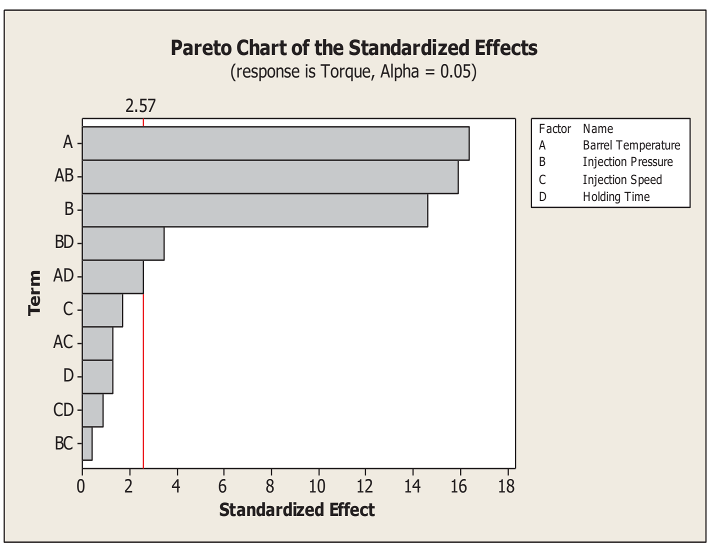


Fig. 6: Pareto Chart of the Standardized Effects

After eliminating the insignificant factors, the analysis which was rerun using significant factors that were identified in Fig. 6. The results confirmed that the factors A (Tbarrel), B (Pinjection), AB (Tbarrel and Pinjection), BD (Pinjeciton and Tholding), and AD (Tbarrel and Tholding) remained as significant factors. Given that factor AD had significant interaction; factor D is included in the analysis (Fig. 7)

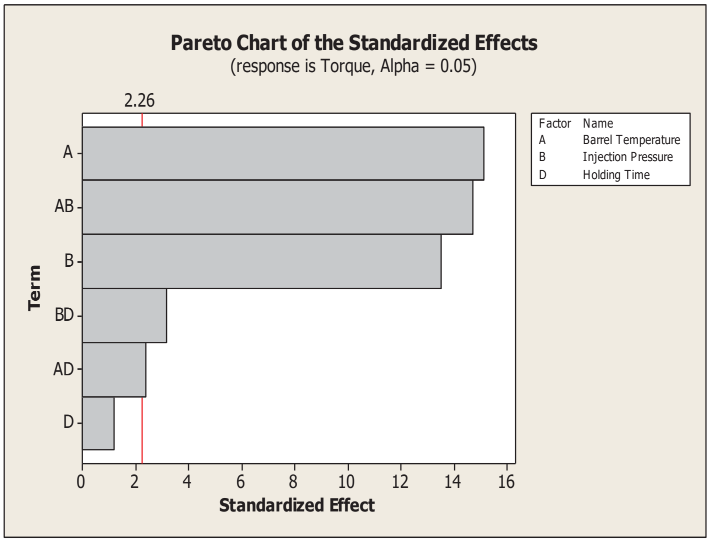


Fig. 7: Pareto Chart of the Standardized Effects

Next, since there are various factors that affect the response, an analysis of variance (ANOVA) was performed to determine the influential ones. If p-values were less than or equal to 0.05 (95% confidence level) they were considered as significant factors. According to the ANOVA output, the following factors were influenced to the torque retention: Two factors (Barrel Temperature and Injection Pressure) were significant. In addition the three interaction factors, “Barrel Temperature” and “Injection Pressure”, “Barrel Temperature” and “Holding Time”, and “Injection Pressure” and “Holding Time” were also significant.

Interpreting the ANOVA Results

Estimated Effects and Coefficients for Torque (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		2.1125	0.03146	67.15	0.000
Barrel Temperature	**0.9500**	**0.4750**	**0.03146**	**15.10**	**0.000**
Injection Pressure	**0.8500**	**0.4250**	**0.03146**	**13.51**	**0.000**
Holding Time	-0.0750	-0.0375	0.03146	-1.19	0.264
Barrel Temperature* Injection Pressure	**-0.9250**	**-0.4625**	**0.03146**	**-14.70**	**0.000**
Barrel Temperature*Holding Time	**-0.1500**	**-0.0750**	**0.03146**	**-2.38**	**0.041**
Injection Pressure*Holding Time	**-0.2000**	**-0.1000**	**0.03146**	**-3.18**	**0.011**

S = 0.125831 PRESS = 0.450370
R-Sq = 98.62% R-Sq(pred) = 95.64% R-Sq(adj) = 97.70%

Analysis of Variance for Torque (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F
Main Effects	3	6.5225	6.52250	2.17417	137.32
Barrel Temperature	1	3.6100	3.61000	3.61000	228.00

Injection Pressure	1	2.8900	2.89000	2.89000	182.53
Holding Time	1	0.0225	0.02250	0.02250	1.42
2-Way Interactions	3	3.6725	3.67250	1.22417	77.32
Barrel Temperature*Injection Pressure	1	3.4225	3.42250	3.42250	216.16
Barrel Temperature*Holding Time	1	0.0900	0.09000	0.09000	5.68
Injection Pressure*Holding Time	1	0.1600	0.16000	0.16000	10.11
Residual Error	9	0.1425	0.14250	0.01583	
Lack of Fit	1	0.0225	0.02250	0.02250	1.50
Pure Error	8	0.1200	0.12000	0.01500	
Total	15	10.3375			

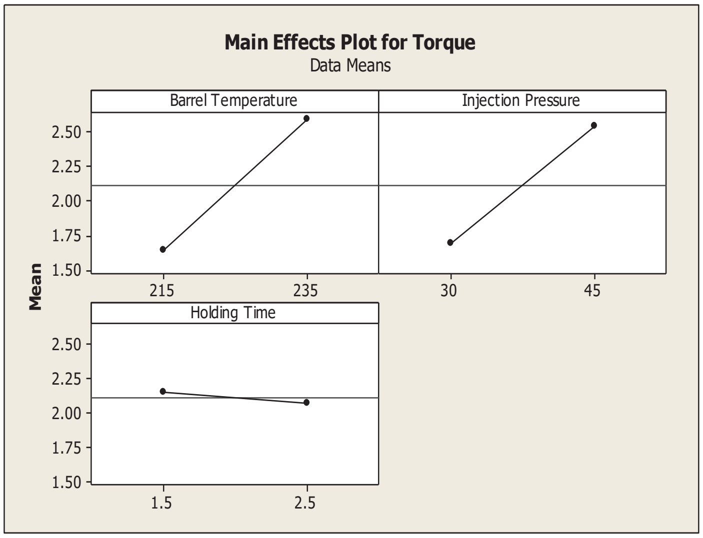


Fig. 8: Main Effects Plot

The Main Effects Plot in Fig. 8 indicates that A (Tbarrel) and B (Pinjection) affected the torque retention because the lines of mean were not horizontal. Both factors showed torque retention increased when the factors moved from the low level to the high level. In contrast, D (Thold) factor had no effect to the torque retention due to horizontal line (parallel to the x-axis).

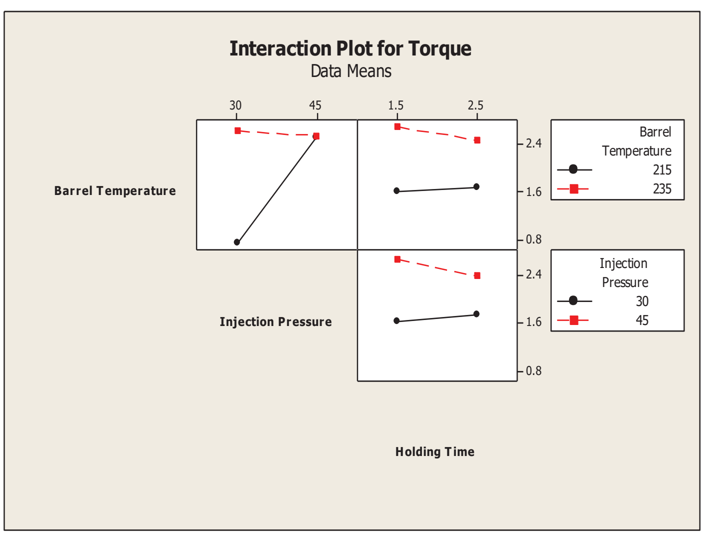


Fig. 9: Interaction plot

From the interaction plot in Fig. 9 it can be observed that the torque retention increases when A (Tbarrel) and B (Pinjection) are high.

The contour plot in Fig. 10 predicted that the torque retention response with the highest yield obtained when both A (Tbarrel) and B (Pinjection) factors were set at the high conditions.

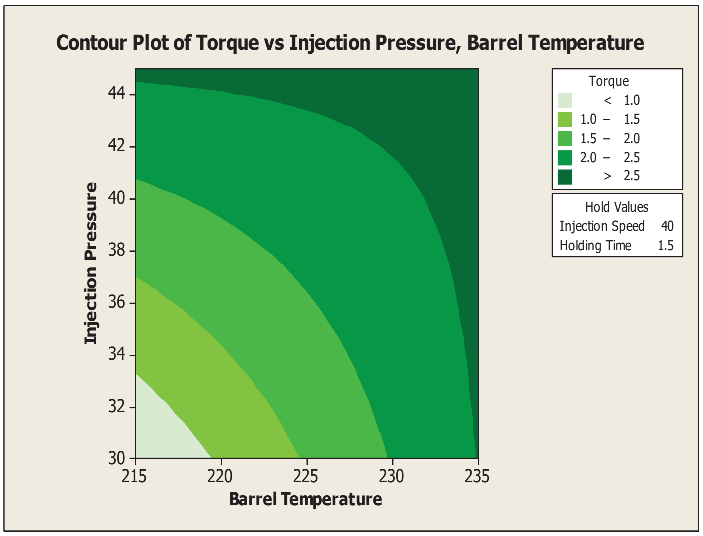


Fig. 10: Contour Plot

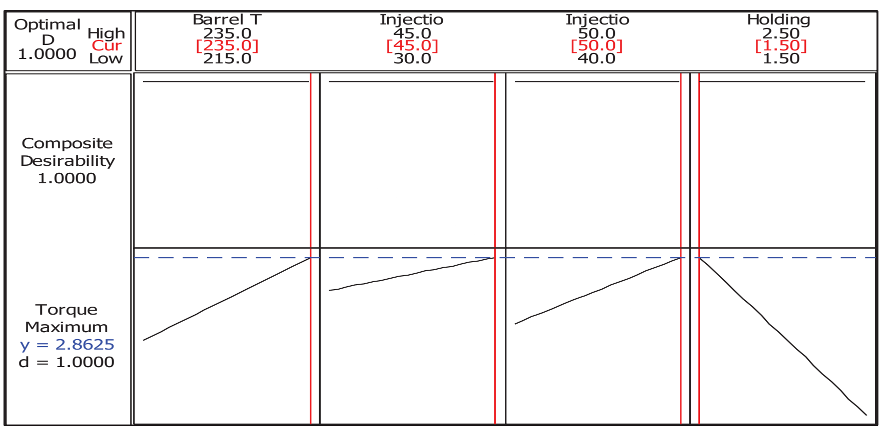


Fig. 11: Optimizer Plot

Applying the response optimization analysis, the process parameter settings at optimal conditions are as specified in Fig. 11. The composite desirability value was at 1 which is close to the best possible prediction.

In order to obtain the value of 2.86 lbf torque retention, the following process parameter settings should be set as follows; Factor A (Tbarrel) at 235º, Factor B (Pinjection) at 45 PSI, Factor C (Sinjection) at 50 seconds, and Factor D (Thold) at 1.5 seconds.

Control Phase

Update pFMEA

The pFMEA was updated to reflect the new process, including the changes, which eliminated several process steps. In addition, the incorporation of the recommended actions reduced the highest risks associated with the process.

Validation and Verification (V&V)

The injection molding process of the fixture is validated under change control request, including Installation Qualification (IQ), Optimization Qualification (OQ), and Performance Qualification (PQ). Proven Acceptable Range (PAR) challenges are completed and verified the process/product capability within specifications throughout the listed process ranges. These process parameters and their ranges are shown below:

Table 5: Proven Acceptable Ranges

Process Parameter	Low	Nominal	High
Barrel temperature (°C)	230	233	235
Injection Pressure (PSI)	42	44	45
Injection Speed (ml/sec)	45	47	50
Hold Time (sec)	1.5	1.6	1.8

Pilot run

One lot was run at the optimized settings of Barrel temperature would be set at 233º, Injection Pressure at 44 PSI, Injection Speed at 47 seconds, and Hold Time at 1.6 seconds. Quality Assurance collected and tested approximately 100 samples and recorded the results. The results are not only used to compare against the baseline data to show the process improvement but also are to validate if the desirable response (torque retention) is met.

Table 6: Individuals of Torque Retention (before vs. after) Improvement

Variable	Count	Mean	StDev	Min	Median	Max	Data Normal?	Cpk
Torque Retention (lbf) (Current Production)	100	1.8	0.24	1.2	1.8	2.3	Yes	0.41
Torque Retention (lbf) (After improvement)	100	2.9	0.20	2.3	2.8	3.4	Yes	1.87

2-Sample T-test

H_0: Torque Retention baseline ≥ Torque Retention after improvement.
Ha: Torque Retention baseline < Torque Retention after improvement.

Two-Sample T-Test and CI: Torque, Stage

```
Stage        N   Mean  StDev  SE Mean
After      100  2.853  0.200    0.021
```

```
Baseline   100  1.814  0.239    0.024

Difference = μ (After) - μ (Baseline)
Estimate for difference:  1.0388
95% CI for difference:  (0.9764, 1.1013)
T-Test of difference = 0 (vs ≠): T-Value = 32.81  P-Value = 0.000  DF = 189
```

Since the p-value= 0 and is less than .05, reject the null hypothesis. It can be concluded that the torque retention mean after improvement is higher than the torque retention baseline. Furthermore, the individual (I-Chart) clearly showed a step change after improvement (Fig. 11).

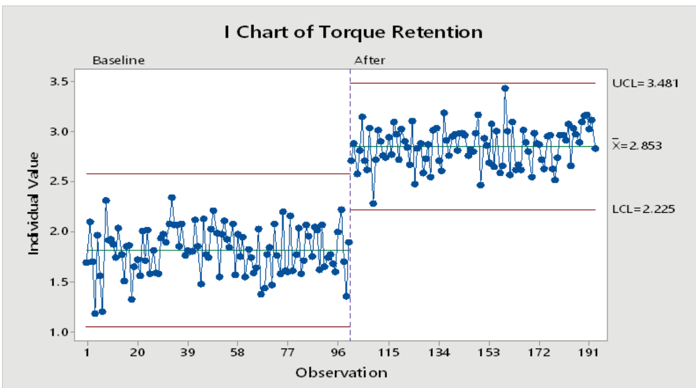


Fig 12: Individual Chart of Baseline vs. After Improvement

Lastly, Process capability analysis was re-analyzed for both Baseline and After Improvement. Both Cp = 2.037 and Cpk = 1.87(After Improvement) were achieved quality level and centering the specification range. This meant that the pumps passed torque retention specification, and the process was capable.

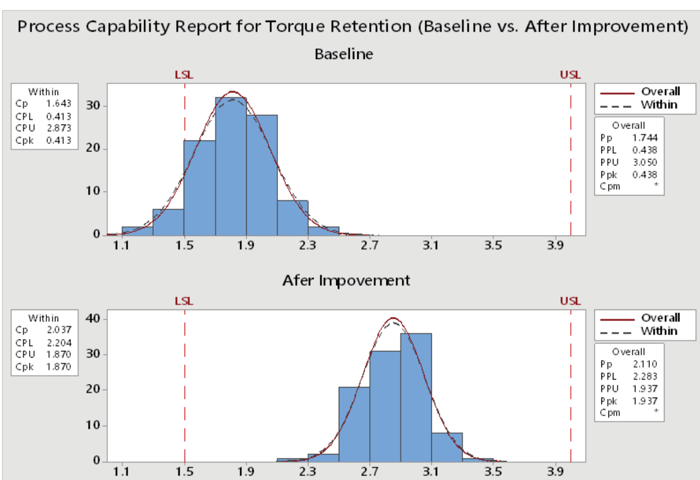


Fig. 13: Cpk Analysis of Torque Retention Baseline vs. After Improvement

Control plan

The optimal process parameters were confirmed after the study and documented in Standard Operational Procedure (SOP) including reaction plans to make corrective actions when the injection mold process is out of the control, especially torque retention will not meet specification limit. Control charts are trended for torque retention, and the results are reported in a monthly KPI review meeting. The control limits are re-calculated on a quarterly basis and revised as needed.

CONCLUSION

The five phases of the Lean Six Sigma were successfully implemented in the injection molding process. The design of experiment case study illustrated the effective use of Lean Six Sigma methodology to reduce process variation and to improve the torque retention output. The design of experiment analysis is applied to the injection molding process using full factorial design. The main and interaction factors are identified. Based on the DOE analysis, it is found that the Barrel Temperature and Injection Pressure had significant impact on torque retention; Injection Speed and Holding Time did not. The design of experiment technique also allows practitioners to run all possible combinations with different levels instead of running one-factor-at-a-time (OFAT) and missing-out the factor interactions. The optimization analysis indicated that parameter settings have led to significant improvements in the torque retention quality characteristics. With the optimized settings, the injection molding process became capable of producing a stable torque retention greater than 2.86 lbf. Lean Six Sigma proved to be a valuable tool to reduce defects and to improve yield of the manufacturing processes. This study provided a manufacturing case study to technology management practitioners detailing a step by step application of Lean Six Sigma techniques. Statistical analysis of the findings is also provided to validate the results. Moreover, the case study reiterates the significant impact of Lean Six Sigma as a tool that can help companies achieve their business objectives to deliver superior levels of performance to overcome operational and strategic challenges.

REFERENCES

Abraham Z., Wen L., Yangyan S., Song T. C., Zhi H. X. S. (2016). Lean and Six Sigma in logistics: a pilot survey study in Singapore. International Journal of Operations & Production Management, 36 (11), 1625-1643.

Antony, J., Snee, R. & Hoerl, R. (2017), "Lean Six Sigma: yesterday, today and tomorrow", *The International Journal of Quality & Reliability Management,* Vol. 34, No. 7**,** pp. 1073-1093.

Ben Ruben, R., Vinodh, S. & Asokan, P. (2017), "Implementation of Lean Six Sigma framework with environmental considerations in an Indian automotive component manufacturing firm: a case study", *Production Planning & Control,* Vol. 28, No. 15**,** pp. 1193-1211.

Box, J. F. (1980), "R. A. Fisher and the Design of Experiments, 1922-1926", *The American Statistician,* Vol. 34, No. 1**,** pp. 1-7.

Chakrabarty, A. & Tan, K. C. A survey on Six Sigma implementation in Singapore service industries. 2007 IEEE International Conference on Industrial Engineering and Engineering Management, 2-4 Dec. 2007 2007, pp. 1428-1432.

Dakhli, Z., Lafhaj, Z. & Bos, A. (2016), "Experiencing Lean Six Sigma in the French residential construction", *International Journal of Lean Six Sigma,* Vol. 7, No. 4**,** pp. 346-368.

Dowlatshahi, s. (2004), "An application of design of experiemtns for optimization of plastic injection molding processes", *Journal of Manufacturing Technology Management,* Vol. 15, No. 6**,** pp. 445-454.

Erdil, N. O., Aktas, C. B. & Arani, O. M. (2018), "Embedding sustainability in lean six sigma efforts", *Journal of Cleaner Production,* Vol. 198, No.**,** pp. 520-529.

Ghosh, M. (2013), "Lean manufacturing performance in Indian manufacturing plants", *Journal of Manufacturing Technology Management,* Vol. 24, No. 1**,** pp. 113-122.

Houborg, C. (2010), "Implementing a successful Lean programme: where do you begin?", *Pharmaceutical Technology Europe,* Vol. 22, No. 9**,** pp. 52-57.

Jiang, J.-C., Shiu, M.-L. & Cheng, H.-J. Integration of six sigma and lean production system for service industry. Proceedings of The 5th Asia-Pacific Industrial Engineering and Management Systems Conference, 2004, pp. 2-22.5.

Jones, C., Medlen, N., Merlo, C., Robertson, M. & Shepherdson, J. (1999), "The Lean Enterprise", *BT Technology Journal,* Vol. 17, No. 4**,** pp. 15-22.

Kumar, M., Antony, J., Singh, R. K., Tiwari, M. K. & Perry, D. (2006), "Implementing the Lean Sigma framework in an Indian SME: a case study", *Production Planning & Control,* Vol. 17, No. 4**,** pp. 407-423.

Lande, M., Shrivastava, R. L. & Seth, D. (2016), "Critical success factors for Lean Six Sigma in SMEs (small and medium enterprises)", *TQM Journal,* Vol. 28, No. 4**,** pp. 613-635.

Lim, S. A. h., Antony, J. & Albliwi, S. A. (2015), "A systematic review of Lean Six Sigma for the manufacturing industry", *Business Process Management Journal,* Vol. 21, No. 3**,** pp. 665-691.

Prashar, A. (2014), "Adoption of Six Sigma DMAIC to reduce cost of poor quality", *International Journal of Productivity and Performance Management,* Vol. 63, No. 1**,** pp. 103-126.

Sekar, V. & Santhosh, D. (2012), "Application of FMEA to an automotive leaf spring manufacturing organization", *TQM Journal,* Vol. 24, No. 3**,** pp. 260-274.

Sternberg, H., Stefansson, G., Westernberg, E., Rikard Boije af, G., Allenström, E. & Malin Linger, N. (2013), "Applying a lean approach to identify waste in motor carrier operations", *International Journal of Productivity and Performance Management,* Vol. 62, No. 1**,** pp. 47-65.

Timans, W., Antony, J., Ahaus, K. & Solingen, R. 2012, *Implementation of Lean Six Sigma in Small- and Medium-Sized Manufacturing Enterprises in the Netherlands*.

Zhang, A., Luo, W., Shi, Y., Chia, S. T. & Sim, Z. H. X. (2016a), "Lean and Six Sigma in logistics: a pilot survey study in Singapore", *International Journal of Operations & Production Management,* Vol. 36, No. 11**,** pp. 1625-1643.

Zhang, A., Luo, W., Shi, Y., Chia, S. T. & Zhi Hao Xavier, S. (2016b), "Lean and Six Sigma in logistics: a pilot survey study in Singapore", *International Journal of Operations & Production Management,* Vol. 36, No. 11**,** pp. 1625-1643.

DECISION SCIENCES INSTITUTE

Quality Management Analytics (QMA) Curriculum for Organizational Performance Improvement

Dr. Borinara Park
Illinois State University
Email: bpark@ilstu.edu

Shehrose Javaid
Illinois State University
Email: sjavaid@ilstu.edu

ABSTRACT

Quality Management Analytics (QMA) is a comprehensive quality improvement approach for (current & future) professionals to increase organizational performance by leveraging the power of the analytical and analytics methodologies. Specifically, Quality Management Analytics (QMA) program focuses on teaching students and professionals on how to: 1) Enhance the organizational performance by reducing business process inefficiencies through the proven systematic quality improvement methodologies; and 2) Promote excellence in business decisions and operations using modern analytics tools, increasing the chance of delivering quality services and products for their customers. This article describes the program for its unique approach, the curriculum structure, and the analytics methods utilized.

KEYWORDS: Quality Management, Performance Improvement, Organizational Excellence, Exploratory Analytics, Predictive Analytics, Prescriptive Analytics

QUALITY MANAGEMENT ANALYTICS (QMA)

Quality Management Analytics (QMA) is a comprehensive quality improvement approach for (current & future) professionals to increase organizational performance and meet demanding customer expectations by creating high quality processes for services and products. It achieves this goal by leveraging the proven power of the integrative analytics methodologies so as to reduce business process wastes and operational variations.

Specifically, Quality Management Analytics (QMA) enables practitioners to: 1) Enhance the organizational performance by reducing business process inefficiencies through the proven systematic quality improvement methodologies; and 2) Promote excellence in business decisions and operations using modern analytics tools, increasing the chance of delivering quality services and products for their customers. Program participants will gain knowledge and skill sets in quantitative and qualitative decision making practices to solve complex business operational problems and quality issues.

Intuitive data analytics tools and qualitative techniques are used to effectively deliver performance improvement solutions. Hands-on exercises and software packages are extensively utilized to help participants learn the concepts and applications in real world problems. In summary, the QMA program aims at facilitating organizational excellence by engaging professionals in data-driven analytics decision practices. The program places emphasis on the fact organizational performance is largely defined by how business entities create, manage, and utilize data related to their business and operation processes. Student will gain knowledge in extracting and analyzing data from business processes and confidence in presenting their findings.

QMA PROGRAM STRUCTURE & CONTENTS

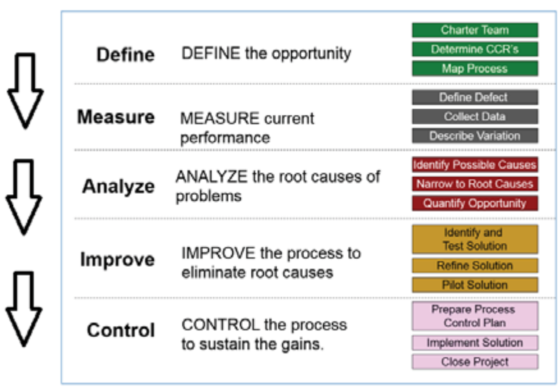
QMA program focuses on providing solutions for industry managers to efficiently handle four major organizational performance issues: 1) System Variations; 2) System Wastes; 3) System Risks; and 4) System Predictions. To this end, the program offers a structured curriculum as summarized below in Figure 1 with conceptual frameworks and analytical/ analytics methods. In the following sections of this article, each curriculum category will be described in detail.

Curriculum Category		Application Domain	Course Name
I. Organizational Excellence (System Performance **Management**)	(Business Functions)	- Organizational Decisions <Integrated Functional Decisions>	TEC xxx
	(Operational Functions)	- Quality Management & Project Management <Operational Excellence & Lean Six Sigma >	TEC xxx
II. Exploratory Analytics (System Performance **Description**)		- Insight Analytics <Visual Mining>	TEC xxx
		- Process Analytics <Process Mining>	TEC xxx
III. Predictive Analytics (System Performance **Prediction**)	A. Inductive Prediction	- Predictive Analytics <Machine Learning>	TEC xxx
	B. Deductive Prediction	- Decision Simulation <Monte-Carlo Simulation>	TEC xxx
		- Process Simulation <Discrete-Event Simulation>	TEC xxx
		- Behavior Simulation <Agent-Based Modeling>	TEC xxx
		- System Alternatives <System Dynamics>	TEC xxx
IV. Prescriptive Analytics (System Performance **Solution**)		- System Optimization <Stochastic Optimization Method>	TEC xxx

Figure 1. QMA Program Structure and Topical Contents

CURRICULUM CATEGORY I: ORGANIZATIONAL EXCELLENCE

In this part of the curriculum, the focus is on how important customer satisfaction is to the business success of any organizations. Students need to understand that the “quality” of any products/ services is one of the major factors in the customer satisfaction. Students also learn how to operationalize quality management as part of any regular business processes. Quality is not a concept but rather a measurable performance indicator by constantly comparing business performances against customer expectations. Students learn how companies can establish and implement such quality management practice using the Lean Six Sigma (LSS) methodology. The methodology aims at explaining how system variations and system wastes lead to customer dissatisfaction and how these fundamental issues can be resolved. QMA students follow the entire quality improvement process, the DMAIC (*Define-Measure-Analyze-Improve-Control*) methodology as a way of systematically and analytically investigating system performance issues and proposing solutions as depicted in Figure 2. Based on a real case or a scenario, students define a quality-related business problem, measure the current level of performance using KPIs, analyze the root causes of variations and sources of inefficiencies, and provide alternative solutions to improve and control system behaviors. At the end of the course, students have learned the fundamental concepts and tools that can be used to improve business system performance.



Define	Measure	Analyze	Improve	Control
Project Charter	Key Performance Indicators (KPI)	Process Mapping	Design of Experiments (DOE)	Statistical Process Control (SPC)
Voice of the Customer	Measurement System Analysis (MSA)	Cause & Effect Diagram	Theory of Constraints	Standardized Processes
House of Quality	Gauge R&R	Hypothesis Testing	Pugh Matrix	Control Plans
Kano Analysis	Process Capability	Confidence Intervals	5S	Total Preventative Maintenance
Critical to Quality (CTQ) Trees	PPM & DPMO	Failure Mode and Effect Analysis (FMEA)	Kanban	Project Report
SIPOC Map	Short/ Long Term Variation	Affinity Diagram	Work Flow Improvement	Procedures and Work Instructions
Gemba	Data Collection Methods	Correlation & Regression	Visual Management	Training Requirements

Figure 2. DMAIC Improvement Methodology and Conventional LSS Analytical Tools

If you are a manager responsible for improving the system performance of a factory, a distribution center, a supply chain, or any business systems, you have a daunting task because these systems are complex, dynamically changing, operating under uncertainties, and heavily influenced by your decisions. Because of these reasons, finding the source of poor system performance and providing the solutions to quality problems are very time-consuming and costly endeavors. In fact, typical analytical and statistical methods and tools used in the conventional Lean Six Sigma (LSS) practices have created these bottlenecks to the level that companies consider the LSS non-agile.

To overcome this very issue, the QMA program integrates various modern analytics and modeling methods as a way of facilitating the DMAIC process to make the quality management practice to be significantly nimble and comprehensive. As shown in Figure 1, these methods are organized into the exploratory, predictive and prescriptive analytics categories in the QMA curriculum to describe, estimate, and direct system behaviors. Specific methods covered include: insight visual analytics; process mining; predictive machine learning; monte-carol simulation; discrete-event simulation; and agent-based modeling. Emphasis on hands-on experience with computer applications is to ensure students to acquire working knowledge beyond the concepts and theories in making decisions in business and operations. Various industry-based decision situations are modeled and analyzed throughout the courses. Let us take a look first how the exploratory analytics is used in the quality management context.

CURRICULUM CATEGORY II: EXPLORATORY ANALYTICS

In this part of the curriculum, the focus is on finding potential reasons and sources of system variations and efficiencies. To illustrate, Figure 3 shows a process analytics methodology, called a process mining (the figure was generated using DISCO by fluxicon), to describe the system behaviors and its embedded wastes. The process map shown, automatically generated from a set of transaction dataset of business operations, allows users to identify typical system wastes such as reworks, dropped cases, and bottlenecks by navigating to various areas of the system in question. Tools such as this are much more powerful and time-saving than the conventional cause-effect tools such as fishbone diagram because users do not have to guess where the inefficiencies come from.

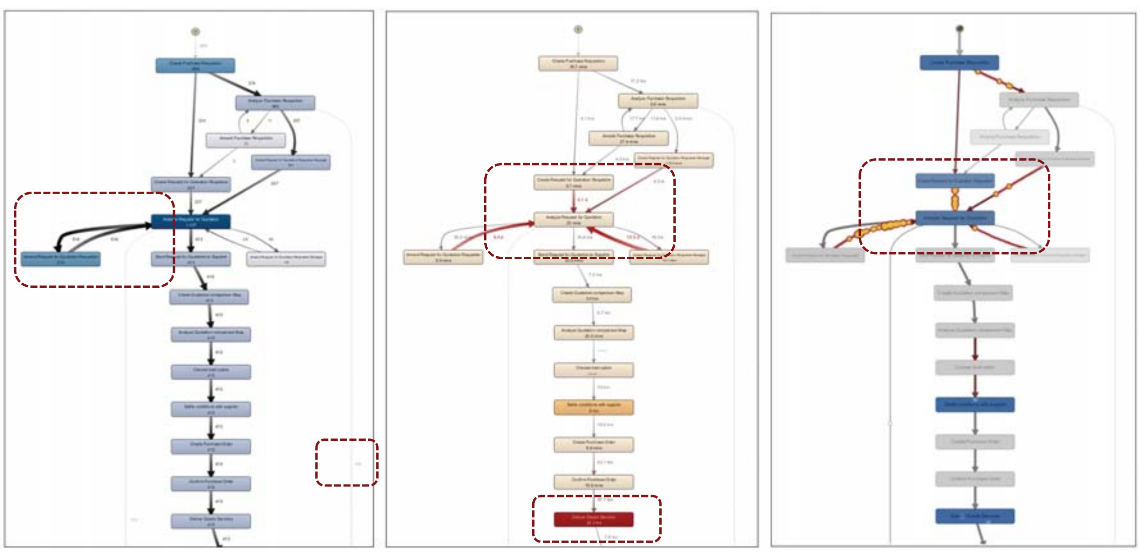


Figure 3. Process Analytics to Identify Wastes and Efficiencies

Figure 4 demonstrates another example of how a visual analytics technique can be used as a quality management tool. If you are a manager overseeing the sales in US and you know the sales in the Eastern states has been a struggle, the visual insight below (created by Tableau) indicates more specific sources of the struggle in terms of location and time in an objective and quantitative fashion instead of relying on the manager’s qualitative judgment. Again, the emphasis is on the speed and accuracy that can be adopted in the quality management practice.

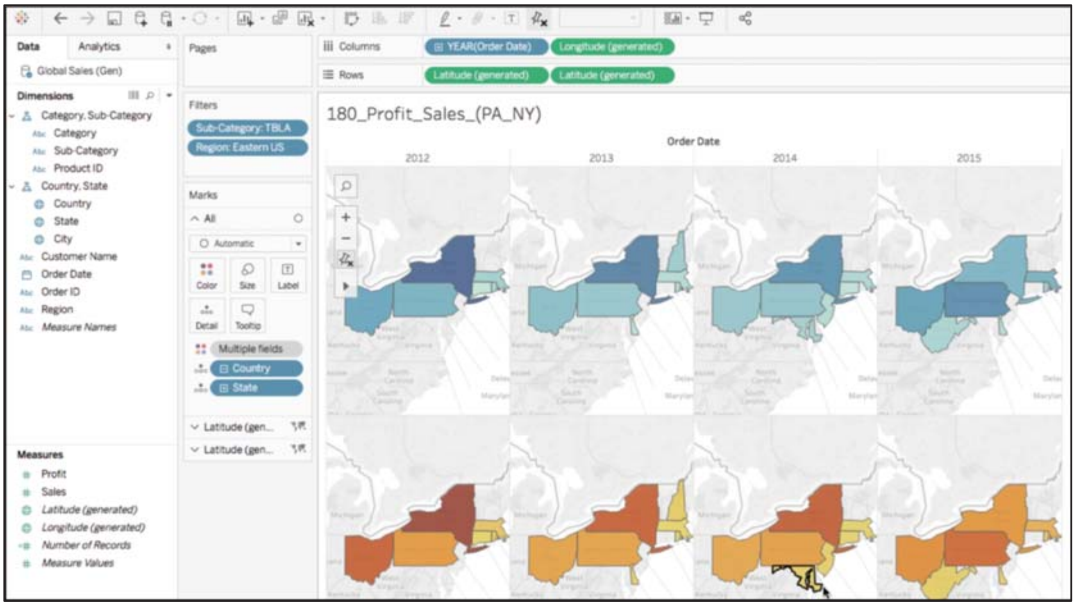
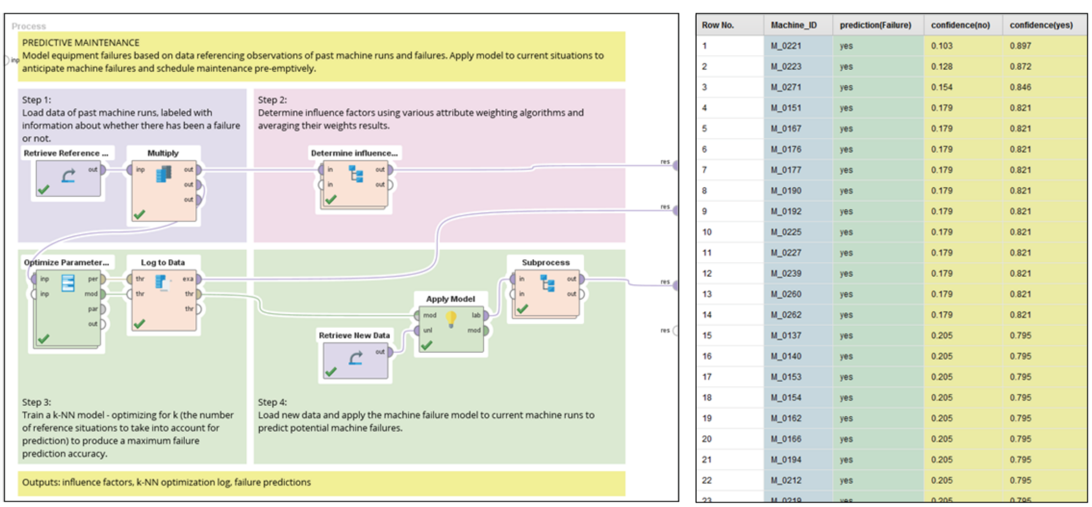


Figure 4. Visual Insights on Sources of Sales and Profit Declines

CURRICULUM CATEGORY III-A: PREDICTIVE ANALYTICS (INDUCTIVE PREDICTION)

If you care for system performance at a certain level, it is always a good idea to predict what could/ would happen before it happens. This may be the best quality management practice after all, if possible at all. One possibility is that now as a manager you have a lot more data that you can turn into a predicting machine. Figure 5 shows a predictive analytics model (created by the RapidMiner software) based on an operational data to take a preventative measure on potential machine failures. This type of analytics use will bring in a completely different managerial mindset in that now you want to be proactive in improving quality of product or services.



Row No.	Machine_ID	prediction(Failure)	confidence(no)	confidence(yes)
1	M_0221	yes	0.103	0.897
2	M_0223	yes	0.128	0.872
3	M_0271	yes	0.154	0.846
4	M_0151	yes	0.179	0.821
5	M_0167	yes	0.179	0.821
6	M_0176	yes	0.179	0.821
7	M_0177	yes	0.179	0.821
8	M_0190	yes	0.179	0.821
9	M_0192	yes	0.179	0.821
10	M_0225	yes	0.179	0.821
11	M_0227	yes	0.179	0.821
12	M_0239	yes	0.179	0.821
13	M_0260	yes	0.179	0.821
14	M_0262	yes	0.179	0.821
15	M_0137	yes	0.205	0.795
16	M_0140	yes	0.205	0.795
17	M_0153	yes	0.205	0.795
18	M_0154	yes	0.205	0.795
19	M_0162	yes	0.205	0.795
20	M_0166	yes	0.205	0.795
21	M_0194	yes	0.205	0.795
22	M_0212	yes	0.205	0.795

Figure 5. Predictive Analytics for Machine Failure

CURRICULUM CATEGORY III-B: PREDICTIVE ANALYTICS (DEDUCTIVE PREDICTION)

In this category, the approach is to put together a model closely emulating the behaviors of the business system. This is a deductive analytics approach because the system itself is replicated as a digital model or twin to describe and predict its behaviors. (*Note that in the previous categories, the focus is on analyzing the data produced from the system under consideration for prediction of its behavior. This is an inductive approach.*) Various simulation methodologies are used to support the deductive prediction modeling practice and some of the powerful ones include: agent-based modeling (ABM), monte-carol simulation (MC), and discrete-event simulation (DE) techniques. They provide more convincing and quantitative solutions to system improvement questions. Traditionally, this is where the conventional DMAIC LSS tools such as Pugh Matrix and FMEA (Failure Mode Effect Analysis) are used. Though analytical, they are qualitative and subjective in their analysis capabilities and they are incapable of handling complex business systems.

On the other hand, a simulation-based application such as the discrete-event simulation shown in Figure 6 (created by AnyLogic) can represent individual system components and interactions as precisely as the system understanding allows. The simulation is capable of predicting how the system would operate in dynamically changing and uncertain conditions.

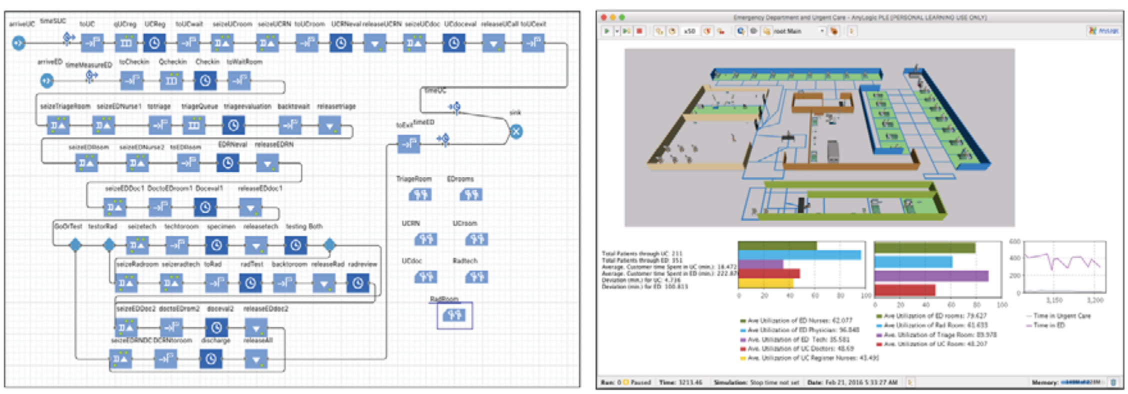
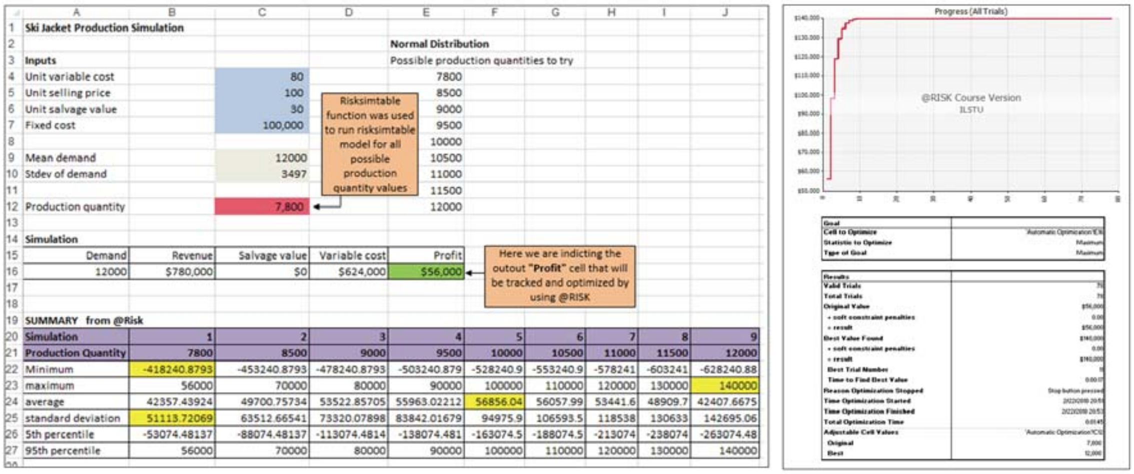


Figure 6. Discrete Event Simulation to Predict the System Behavior of Healthcare Provider

CURRICULUM CATEGORY IV: PRESCRIPTIVE ANALYTICS

In this category, the focus is to provide more definite directions on what parameters of the system need to be changed and how much they have to be modified to achieve the best overall quality. Since the simulation tools in the previous section have the capability to generate multiple scenarios, they are also capable to compare them and choose the best scenario for the users to adopt for optimum system performance. They often use the stochastic optimization techniques to provide more realistic solutions to users who operate the business system in risky and uncertain environments. This is another area where the conventional LSS tools do not provide convincing solutions.

Figure 7 below describe a Monte-Carlo simulation model (created by Palisade's @Risk) where an optimum production level is prescribed based on uncertainties.



	A	B	C	D	E	F	G	H	I	J
1	**Ski Jacket Production Simulation**									
2					**Normal Distribution**					
3	**Inputs**				Possible production quantities to try					
4	Unit variable cost		80		7800					
5	Unit selling price		100	Risksimtable function was used to run risksimtable model for all possible production quantity values	8500					
6	Unit salvage value		30		9000					
7	Fixed cost		100,000		9500					
8					10000					
9	Mean demand		12000		10500					
10	Stdev of demand		3497		11000					
11					11500					
12	Production quantity		7,800		12000					
13										
14	**Simulation**									
15	Demand	Revenue	Salvage value	Variable cost	Profit	Here we are indicting the outout **"Profit"** cell that will be tracked and optimized by using @RISK				
16	12000	$780,000	$0	$624,000	$56,000					
17										
18										
19	**SUMMARY from @Risk**									
20	**Simulation**	1	2	3	4	5	6	7	8	9
21	**Production Quantity**	7800	8500	9000	9500	10000	10500	11000	11500	12000
22	Minimum	-418240.8793	-453240.8793	-478240.8793	-503240.879	-528240.9	-553240.9	-578241	-603241	-628240.88
23	maximum	56000	70000	80000	90000	100000	110000	120000	130000	140000
24	average	42357.43924	49700.75734	53522.85705	55963.02212	56856.04	56057.99	53441.6	48909.7	42407.6675
25	standard deviation	51113.72069	63512.66541	73320.07898	83842.01679	94975.9	106593.5	118538	130633	142695.06
26	5th percentile	-53074.48137	-88074.48137	-113074.4814	-138074.481	-163074.5	-188074.5	-213074	-238074	-263074.48
27	95th percentile	56000	70000	80000	90000	100000	110000	120000	130000	140000

Figure 7. Prescriptive Analytics for Production Quantity

SUMMARY

The Quality Management Analytics (QMA) program is a new approach to change how the quality management practice can be done more effectively by leveraging the power of analytics capability to make quality improvement analysis and solutions robust and agile. In this article, various analytics applications are described to demonstrate how they can complement and improve the conventional DMAIC LSS tools. As analytics technologies evolve rapidly, professionals and academics seek to embed the analytics capability in their own domains. In QMA, the various analytics are actively used to deliver solutions to performance issues related to system variations; system wastes; system risks; and system predictions.

REFERENCE

1. Illinois State University, *Quality Management & Analytics (QMA) for Organizational Performance Improvement*, https://tec.illinoisstate.edu/downloads/qma.pdf

DECISION SCIENCES INSTITUTE

How Challenges and Successes of a Firm Depend on Its Input-Output Ecosystem

Jeffrey Yi-Lin Forrest
Slippery Rock University
Email: jeffrey.forrest@sru.edu

Yong Liu
Jiangnan University
Email: clly1985528@163.com

Theresa A. Wajda
Slippery Rock University
Email: theresa.wajda@sru.edu

Erkan Köse
Nuh Naci Yazgan University
Email: ekose@nny.edu.tr

Oğuzhan A. Arik
Nuh Naci Yazgan University
Email: oaarik@nny.edu.tr

ABSTRACT

We investigate holistically a business system, consisting of market invitations, a firm's supply-chain ecosystem, innovative recognition of market signals, and consequent challenges the ecosystem faces. We show that a firm's upstream component and downstream complement challenges place respectively a cap on the 'value' the firm can create and capture; a firm's capability of identifying challenges is positively correlated to its adeptness in creating and sustaining marketplace advantage and imposing challenges on its ecosystem. A firm's strong capability to learn and act accordingly is an advantage.

KEYWORDS: capability, market invitation, performance, systemic yoyo, value creation/ capture

INTRODUCTION

Systemically speaking, when a firm recognizes a market invitation in an epoch-making way, the firm has an opportunity to develop a significant competitive advantage over its competitors. However, other than internal reasons other players in the ecosystem of the firm's supply chain greatly affect the enhanced performance the firm consequently enjoys (Adner et al, 2013). For example, the difficult daily commuting problem, confronting those individuals who live in a suburban area and commute to work inside a major U.S. metropolitan, necessarily invites politicians and entrepreneurs alike to come up with a solution. One possible resolution to this challenge, for example, is to design, produce, and offer flying cars (Lemoussu et al, 2018) to these frustrated commuters. To manufacture realistically such cars, a firm has to have appropriate components supplied to it in its assembling of the imagined product. However, the value created for commuters is practically realizable until the appropriate road conditions and air traffic controls, as key complements, are constructed and installed in order for such cars to move and fly around freely and orderly. So, a natural question arises: After a firm receives a market invitation for innovation, how will the consequent success of the firm be dependent on the success of other players in its ecosystem?

This paper seeks to address this question, which is important both theoretically and practically, by randomly selecting a firm as our focal firm and by using the thinking logic and methodology of systems science. To accomplish this end, we first distinguish suppliers and complementors so that suppliers provide their outputs for the focal firm to integrate in its effort to offer a complete product to its customers and complementors facilitate necessary conditions for customers to fully utilize the product. Then we develop the following main conclusions amongst others:

- The upper limits of value the focal firm can create and capture are respectively constrained by its upstream component challenges and downstream complement challenges; and
- The more adept a focal firm is at learning and acting, the greater the challenges that will face the ecosystem of the firm's supply chain.

The remainder of this paper is organized as follows. The following section provides a literature review while showing how this work contributes to established knowledge. The next section lists the necessary background information to make this paper self-contained. Following this is the construction of the ecosystem of an arbitrarily chosen business firm. Then studied are the relationship between the firm's capability of learning and its consequent challenges imposed on the ecosystem. The final section concludes the presentation of this work with managerial recommendations and open problems for future research.

LITERATURE REVIEW

Due to the power of systems thinking and methodology, this paper contributes to an array of different literatures, simultaneously covering business ecosystems, value capture and creation, learning capabilities, competitive advantage, and customer adoption of innovation. In this section, we review relevant works in each of these discrete literatures.

In the literature of ecosystems, Adner et al (2013) publish a collection of papers addressing how system players collaborate and compete with each other. Refuting the common beliefs that knowledge ecosystems in technology hotspots leads to business ecosystems in which there are competitive advantages for each of the players in the ecosystem, Clarysse et al (2014) find a particular case where a structured knowledge ecosystem is concentrated around only a number of central players. The result is an imagined business ecosystem that is nearly non-existent. When a business ecosystem includes well-established companies and new ventures, Zahra and Nambisan (2012) claim that success of that ecosystem requires collaboration and competition that demands strategic thinking to leverage a firm's resources and capabilities. Kapoor and Agarwal (2017) study how platform firms in a business ecosystem orchestrate the ecosystem's functioning by providing platforms and setting the rules for participation by complementor firms. They develop a reason for how the ecosystem's structural and evolutionary features may shape the extent to which participating complementor firms can sustain their superior performance. Instead of using technical interdependencies, Hellström et al (2015) explore how firms in an energy business ecosystem collaborate by using a business model lens, pointing to the importance of locating the factors that drive the business models of collaborating firms to facilitate system transitions and change in the logic of an industry. Attour and Lazaric (2018) emphasize that transformation of a knowledge ecosystem can lead to the emergence of a technological platform that provides the resources required especially for firm startup. By conceptualizing the construct of ecosystems, Adner (2017) offers an examination of the relationship between ecosystems and a host of other constructs, such as business models, platforms, coopetition, multisided markets, networks, technology systems, supply chains, and value networks.

In the literature on value creation and capture, Chemmanur et al (2014) analyze how entrepreneurial firms differ when they are corporate venture capital (CVC) based or independent venture capital (IVC) based in terms of their innovativeness. What is found is that CVC-backed firms are more innovative, although they are younger, riskier, and less profitable than those backed by IVC. Kagermann (2015) discusses the impact, challenges and opportunities of digitization and finds that the two key instruments for greater value creation in the Age of Industry 4.0 are platform-based cooperation and a dual innovation strategy. By looking at how the usage of big data analytics (BDA) affects value creation, Chen et al (2015) conceptualize BDA use as an information processing capability that brings competitive advantage to organizations and identify paths through which factors influence the actual usage of BDA.

In terms of learning capability, Camps et al (2016) test hypotheses that link organizational learning capability (OLC) and individual employee performance, OLC and employee flexibility, and employee flexibility and individual performance, and show that employee flexibility fully mediates the relationship between OLC and individual performance under the influence of environmental turbulence. Altinay et al (2016) investigate the interface between OLC, entrepreneurial orientation (EO), and small business performance, and find a positive relationship between EO and sales and market share growth and OLC and EO, but not between EO and employment growth. Kalmuk and Acar (2015) emphasize the mediating role of OLC on the relationship between innovation and company's performance and attempt to demonstrate how multiple types of innovations could contribute to firm performance through OLC. Sung and Choi (2014) examine how training and development affect organizational innovation and suggest that investments in training and development do exert influence on firms' innovative performance through promotion of various learning practices.

Regarding performance and competitive advantages, Yadav et al (2017) look at the association between a firm's environmental efforts and the sustainability of its competitive advantage through analyzing how changes in the firm's environmental performance affect the persistence of its profitability growth. These authors find that environmental resources permit superior financial performing firms to maintain their competitive advantage, while complementing poorly performing firm's efforts to recover. This study empirically supports the proposition that economic value can be created through benefiting the environment. Prajogo and Oke (2016) examine the effect of human capital (HC) on service innovation advantage (SIA) and business performance (BP), and how external environmental factors play a role. These authors show that HC is positively correlated to the creation of value or SIA which in turn results in rent generation for firms, and that the effect of SIA on BP is influenced by environmental factors such that it is strengthened by dynamic environments and weakened by competitive environments. Chang et al (2016) consider the inconsistent relationship between supply chain integration (SCI) and firm performance by proposing that the inconsistency may be associated with selection bias, failure to consider the mediating routes of how SCI affects financial performance, and lack of investigation of moderators. Their findings confirm the early claim that each dimension of SCI indeed improves financial performance.

In terms of market adaptation of innovation, Oliveira et al (2016) aim to identify determinants of mobile payment adoption by proposing a new model. These researchers find that compatibility, perceived technology security, performance expectations, innovativeness, and social influence have significant effects on consumer adoption. This work provides a basis for further refinement of individual models of acceptance. Bilgicer et al (2015) study the role of social contagion in customer adoption of new sales channels by focusing on two aspects of social contagion (i.e., local contagion and homophily), as they influence a consumer'spreference for either an internet channel or a brick-and-mortar store. Bilgicer et al find that social contagion does indeed play a major role with long-term versus short-term customers being less inclined to consider alternate sales channels. By recognizing the importance of downstream activities of the innovation process, specifically marketing and commercialization, Brem and Viardot (2015) consider the adoption of innovations by the market by providing conceptual insights and manners that appear to stimulate and facilitate the adoption of every kind of innovation. Laukkanen (2016) posits that all innovations meet consumer resistance and overcoming this unfriendliness needs to occur prior to product adoption. He examines how five adoption barriers – usage, value, risk, tradition, and image – and three consumer demographics – gender, age, and income – influence consumer adoption versus rejection decisions in internet and mobile banking. Laukkanen finds that value is the strongest barrier, while image slows adoption; and that both gender and age significantly predict adoption and rejection decisions.

This paper contributes to the literature in several different ways. First, it establishes general propositions that are practically applicable with reliability and certainty. Second, it introduces a new methodology for investigating important issues relating to the distribution of challenges across the ecosystem of a focal firm. Third, this paper develops a general theory of business ecosystems without suffering from the constraints and limitations of anecdotes and data analyses (or data mining). In other words, the results established in this paper are universally true within the system of our cohesive theory constructed on top of a few basic assumptions and logical reasoning.

PREPARATION

To make this presentation self-contained, this section briefly introduce the relevant basics and facts of systems science and methodology, the dishpan experiments and the list necessary assumptions and terminology.

Basics of Systems Science and Methodology

Systems are everywhere, especially in investigations of economic issues and business decision making. For example, each person is a complex biological system that is made up of numerous smaller systems. Additionally, that person also belongs to many social and economic systems, such as families, neighborhoods, communities, etc., and that person interacts with a range of systems, such as a car, a smart device, retail stores, the organization the person works for, etc. Because of the existence of such individuals, these systems constantly interact with one another. Hence, other than using numbers and variables to investigate problems and issues regarding organizations, which is mostly reflected in the literature of social sciences, there is a need to employ the concept of systems and relevant methods to study how organizations evolve and how they interact with each other in order to obtain insightful understandings and conclusions that can potentially lead to tangible conclusions in real life.

Historically, the concept of systems has been directly or indirectly introduced by scholars from different disciplines. For example, in the area of economics Rostow (1960) wrote that: The classical theory of production is formulated under essentially static assumptions … to merge classical production theory with Keynesian income analysis … introduced the dynamic variables: population, technology, entrepreneurship, etc. But … do so in forms so rigid and general that their models cannot grip the essential phenomena of growth … We require a dynamic theory … which isolates not only the distribution of income between consumption, savings, and investment (and the balance of production between consumers and capital goods) but which focuses directly and in some detail on the composition of investment and on developments within particular sectors of the economy.

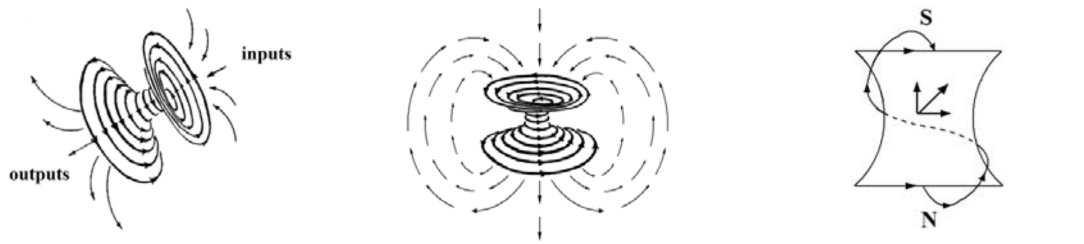
In the area of biology, von Bertalanffy (1924) points out that because the fundamental character of living things is their organization, the customary investigation of individual parts and processes cannot provide a complete explanation of the phenomenon of life. And others, such as Klir (1985), Lin (2009), Porter (1985), etc., also demonstrate how powerful holistic thinking and relevant methodology could be used in producing conclusions regarding organizations and their behaviors that are realistically reliable and practically usable. That is, business entities, economies, and markets, as organizations, do have their individually different internal structures and interact externally with each other. As a matter of fact, since the 1920s (von Bertalanffy, 1924), such a holistic view of nature, organizations, and social events has permeated the entire spectrum of knowledge (Lin, 2009).

The concepts of numbers and systems are respectively rooted in the physical world from two different points of view. For example, when an organization is seen as a set of elements, numbers come into play, such as *n* employees, *m* buildings, etc. However, when the organization is treated holistically, systems emerge, where elements such as employees, capital, properties, etc., form an organic whole through various relationships; without these relationships the organization does not exist. That is, studies in business related disciplines are essentially about relationships of organizations (or systems), be they small economic agents, firms (large or small), markets (local, national or international), industries, or economies.

To summarize, the major differences between the concepts of numbers and systems include: 1) the former is a small scale local concept, while the latter is a large-scale organizational concept (Lin, 1988; 1999); and 2) numbers exist only post existence, while systems emerge at the same time when physical or intellectual existence comes into being (Lin, 2009).

The first difference tells the reason why systems methodology is a more appropriate tool than the classical theories developed on numbers and variables for the investigation of economic entities when their internal structures are being investigated. The second difference explains why business scholars and practitioners still cannot successfully make advanced predictions for imminent economic disasters (Lin and OuYang, 2010).

As for systems science, it is simply the totality of studies of various kinds of systems. In the past century, the methods and thinking logic of systems science have been widely employed in scholarly works (Klir 2001). Similar to how the Cartesian coordinate system – consisting of the crossing of two or more number lines – plays its role in the development of traditional science and mathematics (Kline, 1972), in systems science the role is played by the systemic yoyo model (Lin, 2007), shown in Figure 1.



(a) Eddy motion model of the general system
(b) The meridian field of the yoyo model
(c) The typical trajectory of how matters return

Figure 1. The systemic yoyo model in 3-Dimensional Euclidean Space

Specifically, the blown-up theory (Wu & Lin, 2002), a general theory of development, and discussions on whether or not the world can be seen from the viewpoint of systems (Lin, 1988; Lin et al, 1990) have led to the development of this yoyo model for each object and every system imaginable. Speaking differently, each system can be theoretically seen as a multi-dimensional entity that spins about its axis. If such a spinning entity is fathomed in the 3-dimensional space within which we live, such a structure as artistically shown in Figure 1(a) appears. The input side pulls in *things*, such as human resources, materials, information, investment, profit, etc. After funneling through the "neck", *things*, such as products and services, are spit out from the output side. Some of the things, spit out as outputs, never return to the input side and some will (Figure 1(b)). Due to its general shape, such a structure is referred to as a yoyo.

What this systemic model says is that each physical or intellectual entity in the universe, be it a tangible or intangible object, a living being, an organization, a market, an economy, etc., can all be treated as a realization of a certain multi-dimensional spinning yoyo with an eddy and meridian field around. This yoyo remains in a spinning motion as depicted in Figure 1(a). When it stops its spinning motion, it will no longer be an identifiable system. What Figure 1(c) tells is that the interaction of the eddy field, which spins perpendicularly to the axis of spin, of the model, and the meridian field, which rotates parallel to axis of spin, makes *things* that are either new to the yoyo body or returning to the input side travel along a spiral trajectory.

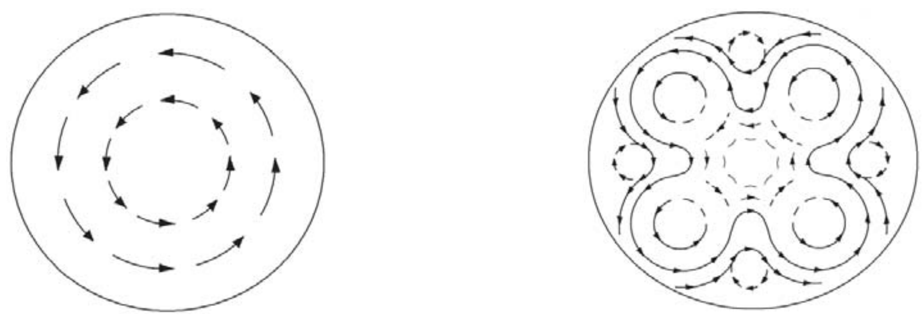
The Dishpan Experiment

In order to understand how the yoyo model (Figure 1) can be applied to the studies of evolution and development of systems, organizations, etc., this subsection will look at the well-known dishpan experiment.

Dave Fultz, University of Chicago, and his colleagues in late 1950s (Fultz et al, 1959) constructed the following experiment: they partially filled a cylindrical vessel with water, placed it on a rotating turntable, and subjected to heating near the periphery and cooling near the center. The bottom of the container is intended to simulate one hemisphere of the earth's surface, the water the air above this hemisphere, the rotation of the turntable simulates the earth's rotation, the heating and cooling simulate the excess external heating of the atmosphere in low latitudes and the excess cooling in high latitudes.

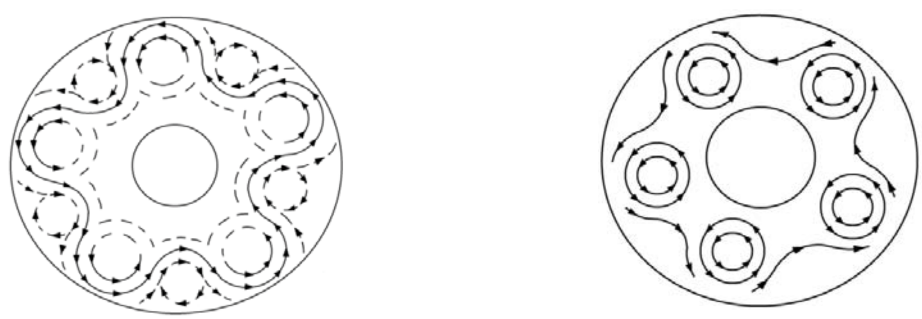
In order to observe flow patterns at the upper surface of the water, which was intended to simulate atmospheric motion at high elevations, Fultz and his colleagues sprinkled some aluminum powder. A special camera that effectively rotated with the turntable took time exposures so that a moving aluminum particle would appear as a streak and sometimes each exposure ended with a flash, which could add an arrowhead to the forward end of each streak. The turntable generally rotated counter clockwise, as does the earth when viewed from above the North Pole.

Even though everything in the experiment was arranged with perfect symmetry about the axis of rotation, such as no impurities added in the water, the bottom of the container was kept flat, etc., Fultz and his colleagues observed more than they had bargained for. First, both expected flow patterns (as shown in Figure 2) appeared, and the choice depended on the speed of the turntable's rotation and the intensity of the heating. Briefly, with fixed heating, a transition from circular symmetry (Figure 2(a)) would take place as the rotation increased past a critical value. With a sufficiently rapid, but fixed rate of rotation, a similar transition would occur when the heating reached a critical strength, while another transition back to symmetry would occur when the heating reached a still higher critical strength.



(a) Symmetric flow at the upper surface (b) Asymmetric flow at the upper surface

Figure 2. Patterns observed in Fultz's dishpan experiment



(a) One flow pattern (b) Another flow pattern

Figure 3. Streak patterns of the flow at the upper surface of Hide's dishpan experiment. (a) is followed by (b) after 8 rotations; and (b) is followed by (a) after another 8 rotations

A few years before the previous dishpan experiment was conducted in Chicago, Raymond Hide, Cambridge University, did a similar experiment in England (Hide, 1953). Instead of a dishpan, Hide used two concentric cylinders with a fluid placed in the ring-shaped region between the cylinders. He discovered similar transitions between symmetric and asymmetric flow patterns. In this experiment, the asymmetric flow was often regular and consisted of a chain of apparently identical waves, which would travel around the ring-shaped region without changing their shapes. What's more remarkable about what Hide found is that a chain of identical waves would appear. But, as they traveled along, they would alter their shapes in unison in a regular periodic fashion and after many rotations of the turntable, they would regain their original shape and then repeat the cycle (Figure 3).

What's important about these experiments is that structures, such as jet streams, traveling vortices, and fronts appear to be basic features in rotating heated fluids, and are not peculiar to atmospheres only.

Basic Assumptions and Terminology

To make our logical reasoning valid, we assume that each and every firm in the business world exists for the purpose of satisfying a particular market niche through generating a positive cash flow. The positive cash flow can be a result of profits from the marketplace, or investments from various investors, or both.

By a *value system*, we refer (Porter, 1985) to associated activities that jointly transform raw materials into products or services for end users. These activities are usually performed by various firms with end users known as *consumers*. In contrast, intermediate, B2B purchasers in a value system are referred to as *customers* instead of consumers. When a firm delivers its products and/or services to both customers and consumers simultaneously, for the sake of convenience of communication, we will just refer to both of these users as customers.

Assume that value creation is the principal objective of any business enterprise, where the outputs are more valuable than the sum of the inputs. It consists of two components: creating value for customers as well as creating value for shareholders. The former helps sell products and services, while the latter, in the form of a rising stock price, insures the future availability of investment capital to fund the operations or to help with the positive cash flow of the enterprise. The *value creation* of a value system is determined by the expected benefits that consumers will receive from their purchases, as represented by their willingness-to-pay (WTP). Due to the lack of perfect price discrimination, consumers generally maintain a consumer surplus, because consumers' spending with payments into the value system is mostly less than the aggregate WTP.

After value has been created, the business enterprise whose efforts led to the value creation needs to capture the value in order to achieve its financial goal of first surviving and then thriving in the economic world. For example, value is created when farmers plant and grow crops; however, these farmers have to capture their created value by harvesting and selling their crops. Value capture by an individual firm is usually determined by its ownership of resources and power status within the value system, as well as by the firm's cost structure and negotiating skill in relation to other firms in the value system (Bowman & Ambrosini, 2001; Priem, 2007).

THE ECOSYSTEM OF AN ARBITRARILY CHOSEN FIRM

This section looks at why market invitations can be at least partially answered and establishes some basic properties of an arbitrarily chosen focal firm in relation to its ecosystem.

The Ecosystem and a Systemic Modeling

The success of our focal firm primarily depends on how well other firms are doing in its input-output environment because of their economic interdependence, where the inputs of one firm are the outputs of some other firms. That is, by using the flow of inputs and outputs we can talk about the ecosystem within which the focal firm resides with other firms seen as either upstream components or downstream complements of the system (Adner & Kapoor, 2010). In particular, suppliers represent some of the upstream components of the focal firm; customers, supporters and assistants who help to make the product of the focal firm usable by consumers are the downstream complements. Supporters and assistants are called complementors because although they are outside the focal firm’s direct supply chain they need to invest and develop new infrastructure to make the focal firm’s product practically usable by the ultimate consumer. Governments, regulators, etc., are examples of complementors that help to build and to regularly maintain, for instance, the roads necessary for transporting products, or the specification of new safety procedures, etc. With this explanation in place, we can see that the set of components of the focal firm also includes upstream complementors that help to make the outputs of the suppliers usable by the focal firm.

Figure 4 depicts the systemic structure of this ecosystem, where the focal firm utilizes the inputs, which will be called components, from *n* supplies and delivers its outputs to customers with the support and assistance of *m* complementators, *m*, *n* = 1, 2, 3, … These support and assistance are called complements. As shown in Figure 4, only first-tier components and complements are given, although this structure in real life should be extended forward and backward along the chain of value creation to include actors of other tiers, such as suppliers' suppliers, customers' customers, complementors’ complementors, etc. And when the full range of all players of different tiers is included in the ecosystem, one actually develops a practical situation of linked systems (Lin & Ma, 1987). (To make the figure intuitively clear, all upstream complementors of the focal firm are omitted.) In other words, the ecosystem of the focal firm possesses the characteristics of the yoyo model in Figure 1(a), where the firm’s internal working is around how its employees with adequate organizational supports can innovatively integrate the components from suppliers in order to deliver its functional product(s) to its customers. Note that similar systemic constructs such as this one has also attracted the attention of academics and practitioners (Iansiti & Levien, 2004; Adner, 2006; Intel Corporation, 2004; SAP Corporation, 2006).

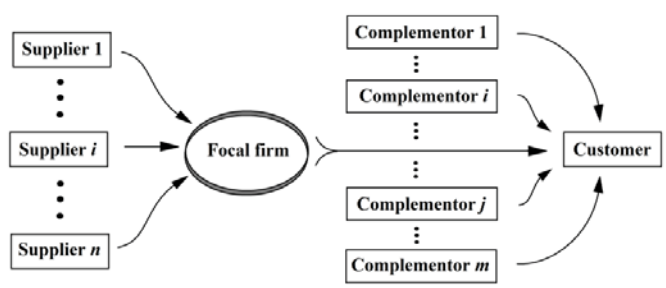


Figure 4. The ecosystem of any focal firm

By modeling each firm in the ecosystem of the focal firm and the market as a system, and a mapping ℓ_{ts} from an upstream supplier firm S_s to a downstream customer firm S_t as the linkage of how S_s’s component m_s is applied in S_t’s product m_t, we have the following result.

Theorem 1. Assume that *T* is a given partially ordered set of type α and that *S* is an α-type hierarchy of systems over *T*, satisfying that each state S_t is a nontrivial system, for $t \in T$. Then *S* can be made into a linked hierarchy by a family $\{\ell_{ts}: t, s \leq T, s \geq t\}$ of linkage mappings of *S*.

The proof of this result is very technical and is omitted.

One interpretation of this theorem of systems science can be given as follows: When the focal firm receives a market invitation for innovation, meaning that the partially ordered set *T* has a set of minimal elements (representing the various market demands related to the invitation), then each and every firm within the ecosystem of the focal firm will be challenged one way or another to provide a resolution to their respective challenges in order to jointly answer the market call. Although the initial market invitation might not be answered perfectly, some combinations of individual firms' resolutions to their respective challenges would provide at least a partial solution to the market demand.

Success and Challenge of the Focal Firm

Based on the ecosystem of the focal firm constructed previously, this subsection establishes a few fundamental conclusions about the focal firm in terms of its market success, the values it creates and captures, and the magnitude of challenges it faces in the marketplace.

Proposition 1. The success for the focal firm to meet its challenges that appear when the firm decides to answer a market invitation is often dependent on the successes of the other firms within the ecosystem to meet their respective challenges that arise out of the focal firm's desire to answer the market invitation.

In fact, this result follows from how the ecosystem of the focal firm is constructed, as shown in Figure 4. In particular, this ecosystem organically relates the firm's performance with not only the challenges the firm faces (Henderson & Clark, 1990; Christensen, 1997), but also those that confront the external partners. For instance, to meet a challenge, the focal firm may very well need to acquire certain special upstream components, which may in turn represent challenges to the relevant suppliers and their corresponding complementors. At the same time, a solution to the challenge, as identified by the focal firm, may likely require downstream complementors to successfully meet their corresponding challenges in order for the focal firm's new product to be fully usable by the customers. In other words, this ecosystem, as a systemic model of the environment of the focal firm, explicitly explains how the network of challenges centered on that particular challenge the focal firm faces needs to be successfully resolved by different actors in the ecosystem in order for the ecosystem to create value.

On the practical front, there are many real-life instances that support the conclusion of Proposition 1. For example, direct current (DC) generation technologies declined because of unconquerable bottlenecks in the development of DC distribution technology (Hughes, 1983), while in the semiconductor lithography industry optical lithography dominates all non-optical approaches as a consequence of the fact that suppliers, customers, and complementors played roles in meeting their respective challenges in optical lithography technology (Henderson, 1995). On a more personal level, when one of the co-authors of this paper had to convert her house to gas heating from its original 100% electric home, although the conversion was possible and done successfully without any trouble, the system became louder than before. Apparently, the duct work in electrical homes is not directly (1 to 1) compatible with that of gas-heated homes.

The duct work of the former homes is narrow so that the newly installed gas system pushes more air through a smaller space, and hence, lending itself to a noisier output.

Proposition 2. The upper limit of the value the focal firm can *create* when it answers a market invitation is constrained by the upstream component challenges the suppliers consequently confront, while the upper limit of the value the focal firm can *capture* from its effort to answer the market invitation is constrained by the downstream complement challenges.

In fact, the first conclusion follows from the fact that the focal firm's ability to actually make a product that is expected to be competitive in the marketplace is constrained by the challenges the upstream components faces. The second conclusion is a consequence of the fact that how much of the full benefit of the focal firm's offering the customer can enjoy from consumption is influenced by the downstream complements.

By the magnitude of a challenge the focal firm faces, it means how much the challenge requires the current approach of problem solving of the firm to change. In other words, the magnitude of a challenge to the focal firm corresponds to the degree to which this firm is charged with changing its approach to problem-solving.

Proposition 3. The greater the magnitude of a challenge the focal firm faces in the marketplace, the greater change the entire ecosystem of the firm will be required to undergo in order for the focal firm to resolve the challenge.

When the focal firm confronts a challenge in the marketplace, it has to make certain necessary changes in its adopted way of problem-solving in order to come up with a resolution beneficially. Now, the required changes necessary to produce the desired resolution could be in:

- The discovery of market trends, the consequent design of brand new ideas, and the development of competitive products (Tushman & Anderson, 1986; Henderson & Clark, 1990), or
- The identification of how some environmental elements can be integrated into firms' internal structures (McGrath, 2013; Takeishi, 2002; Brusoni, Prencipe, & Pavitt, 2001), or
- The determination of how to scale up the production and delivery of identified solutions (Argote, 1999; Hatch & Mowery, 1998).

However, for the ecosystem's purpose of creating value, the focal firm's current approach of problem-solving has achieved a dynamic equilibrium with those of the other players in the ecosystem. In other words, in their effort of achieving the common goal of value creation, all players in the ecosystem are in a state of mutual forbearance. That is, the players, be they firms, complementors or consumers, mitigate the overall challenge of creating and capturing value by dividing the series of necessary activities into bundles, each of which is manageable by one player according to its strength (Bernheim & Whinston, 1990). They cede efforts in less efficient areas to other more capable players, while in exchange the latter do the same in areas where the former are more efficient (Li & Greenwood, 2004). The players' codependence gradually makes the ecosystem stable in their concerted effort of value creation and value capture (Yu & Cannella, 2012). Eventually, the association of the players becomes strengthened (Fuentelsaz & Gómez, 2006) and the network of adopted approaches of problem-solving by different players reaches equilibrium (Haveman & Nonnemaker, 2000). Therefore, when the focal firm is forced to make changes in its adopted approach of problem solving, the stability and equilibrium of this network has to be broken; and at the same time, the greater the change the focal firm has to make, the greater alteration the entire ecosystem will have to make in order for the system to reach another equilibrium.

In terms of the systemic yoyo model, the situation implied by the setup of Proposition 3 can be depicted in Figure 5, where one linear branch of the supply chain is shown as the horizontal curve with the arrows indicating the direction of movement of inputs and outputs of the firms along the chain, and the vertical arrowed curves stand for the combined eddy field of the individual eddy fields of the firms along the supply chain. When a certain amount of value is created and captured by the firms, an equilibrium and stability between the vertical and horizontal fields are established. However, when the focal firm faces a challenge from the marketplace, it means that its outputs need to be altered, which in turn implies that components from suppliers different from before will be needed. That consequently means that the suppliers' outputs have to be different from before. At the same time, the outputs of the downstream firms will in turn be changed accordingly. Such a domino effect quickly rises upward affecting all the suppliers and the suppliers' suppliers as well as downward exerting demands for downstream firms to alter their offers accordingly. So, once again, this analysis indicates that the more a change is required of the focal firm, the greater change the entire ecosystem will have to go through.

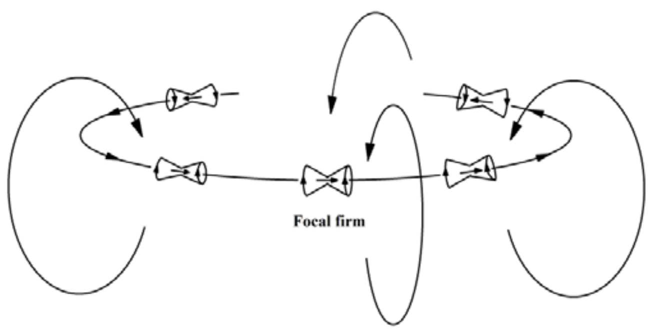


Figure 5. One linear branch of the supply chain within the ecosystem of the focal firm

CAPABILITY OF LEARNING AND CHALLENGE ON THE ECOSYSTEM

As the title suggests, this section studies how the market signals for additional competition, new innovation and how the focal firm's capability to learn and act accordingly leads to challenges to its ecosystem other than improved advantages. In particular, as shown in Figure 6, the systemic yoyo field of the market includes inputs, which are reflected in such terms as new kinds of products and/or services supplied into the market, and outputs, which could be conceived as demands for improved products and/or services. Although various upstream supplies are available to the focal firm, when the firm learns what the market is calling for, it designs innovative products, which in turn requires the upstream suppliers to provide different kinds of innovative components.

Theorem 2 (Forrest et al, 2017). In the coordinately monopolized market described below, if the magnitude of the market segment of switchers – those consumers who make purchases based on which firm's price is most competitive – is greater than α, then at least one new enterprise would enter the market profitably, as a competitor of the existing firms. The said market consists of m incumbent firms, m = 1, 2…, which provide consumers with mutually substitutable products such that

- Each of the incumbent firms has developed its respective share α of loyal consumers who purchase the products from their respective firms only as long as the price is not more than their reservation price, which is set to be 1;
- The incumbent firms compete over the switchers with adjustable prices charged to their consumers; and
- The managements of these *m* firms are well aware of the pricing strategies employed by the firms and have established their best responses by playing the Nash equilibrium through pure self-analyses.

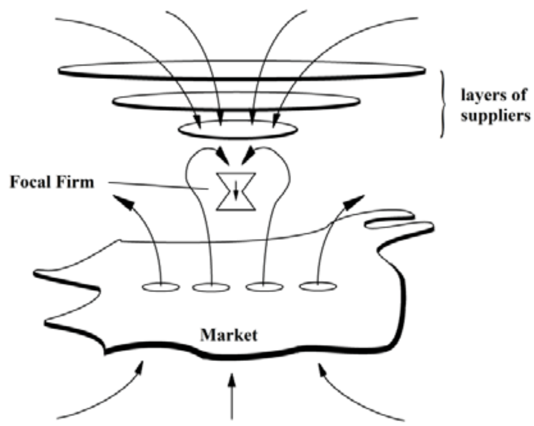


Figure 6. The interaction between what the focal firm learns and what is demanded by the firm from the upstream suppliers

In particular, the condition of this theorem – the magnitude of the market segment of switchers is greater than the average market share of the incumbent firms – is such a key that it is subject to different understandings of the market players. For instance, other than frugality or other thousands of reasons, one explanation for the existence of these switchers is that they are not totally satisfied with the features of the existing products. Aiming at such an understanding of the market cue, our focal firm could potentially come up with an innovative idea on how to improve the functionalities and features of the existing products. However, such an idea could most likely post challenges to the upstream suppliers for them to design and produce relevant components for the focal firm to integrate in order to make its imagined offer to the marketplace. Additionally, at the same time, the idea could potentially post challenges to the downstream complementors, too, in order for customers to enjoy the full benefit of the revolutionary offer of the focal firm. That is, the entire ecosystem now confronts a challenge, the magnitude and location of which depends on the first innovative idea of the focal firm. At this junction, a natural question arises: Would being an early mover be advantageous for the focal firm?

Intuitively, a leader (or first mover), by virtue of having beaten all rivals in the marketplace, should be able to enjoy non-existing or reduced competition when presenting its offer to the market, as documented by Lieberman and Montgomery (1988). That would naturally provide the leader with an advantageous exclusivity although it might be only temporary. However, from Theorem 2, this intuitive understanding, even so it is empirically supported, suffers from one practical problem – the existence of the desired market. In other words, if the imagined offer is too much ahead of its time, there might not be any readily available market for the product. So, first movers/leaders or pioneers in this sense will have to spend their time, efforts, and resources to first explore and define the market for their products. This end explains why studies, see (Lieberman and Montgomery, 2013) and references there, suggest that time can also benefit later entrants by helping to improve risk estimations, clear uncertainties, improve capability gaps, and facilitate learning from the costly mistakes of earlier entrants.

So, in terms of market entry timing as defined by (Zachary et al, 2015), Theorem 2 implies that market characteristics (or invitation) beget the entry of new comers with the potential of making more profits than any of the earlier movers. Therefore, the general symmetry assumed in Theorem 2 indicates that new entrants have at least the same amounts of capabilities and resources as the earlier movers. That simply implies that established firms should not strive to be early movers in any new market until the early movers have established the market with sufficient depth. This end theoretically confirms what Golder and Tellis (1993) find empirically: pioneers enjoy an average of 10% of the market share with nearly half of them failing outright, while early followers take roughly 28% of the market share with failure rates around 8%.

Practically speaking, Theorem 2 explains why large, multi-business enterprises generally exploit slack resources and complementary capabilities to scale up their operation and bypass early entrants (Markides & Geroski, 2005). In other words, size advantage, be it small or large, can make a firm more fitting for a later entry, as noted by Zachary et al (2015).

Systemically, Theorem 2 can be seen readily from the systemic yoyo model. For example, the entire market of concern can be modelled as the spinning dish in Figure 2(b), and each of the *m* incumbent firms aa a regional eddy pool. As such, the assumed symmetry in the market place, just like that that exists in the dishpan experiment, suggests that it is impossible that a blank space, which is at least as big as the area occupied by one of the local eddy leaves, would appear within the circular chain of the local eddy pools. Furthermore, the total area between adjacent local eddy leaves and between the periphery of the dish and circular chain of the local eddy leaves will be too large due to the fact that the local eddy leaves are created by uneven distributions of forces that act on the fluid particles located at different distances from the center of the dish.

What has just been discussed in the previous paragraphs actually explains theoretically why the key challenge underlying the early mover advantage is the opportunity for the focal firm to experiment innovative conceptual products, search for potential customers, and eventually form a market of sufficient depth. If successful, the firm would have acquired very valuable knowledge on production exploitation and market experience, while increasing its added value by improving its offer’s performance and/or costs (Spence, 1981; Lieberman, 1984, 1989; Argote, 1999). At the same time, it also illustrates the reason why pioneers tend to occupy small market share with nearly half of them failing outright (Golder & Tellis, 1993) because of the prohibitive difficulties existing first with the design and production of innovative conceptual products and then with the formation of the relevant market of sufficient depth.

Proposition 4. The more capable the focal firm is to learn and to discover, the greater advantage the firm is able to create and the greater the challenge the firm imposes on its ecosystem.

In fact, the previous analysis shows that the firm’s capability to learn and to discover is required in the understanding of market cues, in the formation of innovative concepts, and in the development and production of competitive products to satisfy the emerging market need. Therefore, the more capable the focal firm is in terms of learning and discovery, the more likely the firm will come up with innovative concepts and the consequent design and production of epoch-making products. Now, Proposition 3 implies that in this case the ecosystem will confront greater challenges in order to supply for the focal firm's design and production of new products and facilitate the delivery of the firm's products to customers.

Speaking systemically, when the focal firm is very capable of learning and discovering, it means that it is more convenient and ready for the firm to sense opportunities and emerging demands in the marketplace. That sensitivity in real life generally means that the firm has an established mechanism for learning and discovering and for how to benefit from what is learned and what is discovered, otherwise any such sensitivity would be dulled either consciously or unconsciously because resources are needed to maintain the said sensitivity. That implies that the firm's output will be replaced with improved or brand new products based on new knowledge and discovery, which consequently imposes challenges, Figure 5, on the upstream suppliers to provide different and, more likely, innovative outputs from before and on the downstream complementors to develop appropriate supports and assistance in order for customers to fully enjoy the new focal offers. Such chain effects rise quickly upward affecting all the suppliers and the suppliers' suppliers and downward disturbing all the complementors and the complementors' complementors. So, once again, this analysis indicates that the more a change is required of the focal firm by new knowledge and discovery, the greater advantage the firm is able to create and the greater change the entire ecosystem will have to go through.

Proposition 5. For the focal firm, its strong capability to learn and to act accordingly is an advantage of the firm over its competitors.

In fact, beyond what is given in the reasoning of Proposition 4, the level of learning capability depends on both the magnitude and the scope the focal firm needs to make changes. It is the emerging mastery of new routines of operation, different ways of thinking and forever changing processes of production that underpins the firm's success with its rides on the waves of transient competitive advantages in the increasingly turbulent economic market (McGrath, 2013). If little or no change from the status quo is necessary, as one of those incumbent firms described in Theorem 2, then there is not much for the firm to learn and to discover so that the motivation for learning and the relative advantage gained from learning and corresponding actions will be either low or nonexistent. In contrast, when the focal firm needs to bring its constantly different and supposedly better offers to the market, it has to overcome related uncertainties and complexities. That explains why a strong capability for the firm to learn and to act accordingly represents a unique advantage of the focal firm.

CONCLUSION

One important characteristic of the business landscape of the modern world is that successful firms have to ride waves of transient competitive advantages (McGrath, 2013). To do so, a firm faces challenges first in learning what the market is signaling, second in deciphering what innovative solution the firm can come up with to answer the market call by mobilizing its resources, and third, in how to deal with considerable challenges in resolving uncertainties appearing with upstream suppliers and downstream complementors.

The literature has provided a rich collection of factors, each of which represents a challenge facing business enterprises with respect to innovativeness and good performance outcomes for a firm (Christensen, 1997; Marzi et al, 2017). However, as pointed out by (Adner & Kapoor, 2010) and as shown in this paper, challenges that surround an innovative design, consequent production and movement of the product out to the market are often situated both inside a focal firm and within the firm's ecosystem. This paper propels us beyond the traditional analysis of the literature of innovative firms' internal and environmental challenges by theoretically establishing a series of generally true conclusions on systemic logic reasoning and relevant intuition.

By holistically looking at the ecosystem of a focal firm that recognizes an invitation from the market, this paper provides a systemic framework by focusing on input-output threads located throughout a supply chain. Within this framework, we show that the focal firm's innovative idea introduced to answer the market invitation generally posts challenges to upstream suppliers and downstream complementors. We illustrate that suppliers' resolutions to component challenges help the focal firm to produce its ideal product, while complementors' success with their challenges facilitates the necessary conditions for the focal offer to reach customers with full potential of value creation. We examine how an epoch-making understanding of a market invitation by a focal firm can simultaneously challenge both suppliers and complementors with asymmetrical levels of effects on the performance of the focal firm. Specially, we establish the following main results among others:

- Upstream component challenges and downstream complement difficulties can respectively constrain the values a focal firm can create and capture; and
- When a focal firm is more capable to learn and to act accordingly, that firm's ecosystem will experience greater challenges.

Other than developing a series of practically useful conclusions, this paper expands the traditional toolbox of reasoning and analysis through demonstrating the power and usefulness of systems science and methodology. By utilizing holistic thinking and logic reasoning, we are able to investigate the dynamics of the ecosystem of a focal firm, and produce different while more insightful understandings than those in the literature on the distribution of challenges, appearance of uncertainties, evolution of technologies, and performance of a focal firm, (e.g. Utterback & Abernathy, 1975; Anderson & Tushman, 1990; Cusumano et al, 1992).

The conclusions established in this paper provide a practical managerial guideline on how to manage investments in meeting challenges originated from market calls. In particular, we have the following recommendations:

- A firm's investment in speeding up the commercialization cycles when pursuing first mover advantage has to carefully consider whether other firms in the firm's ecosystem can promptly meet their respective challenges or not.
- When introducing an innovative product to answer a market invitation, the more upstream suppliers and the less downstream complementors are challenged, the greater value the product is expected to create and capture.
- If a perceived innovative product posts challenges, downstream complementors will likely have difficulty in meeting the market's need in a timely manner. Consequently, the innovative product needs to be modified to ease the expected difficulty confronting complementors.
- Any firm that desires to be succeed in its market domain needs to focus on improving its capability to learn and to discover. Potential performance advantages of the firm originate from such capability.

In other words, a firm can estimate the levels of relative challenges across its ecosystem well in advance of the commercialization of the firm's innovative idea.

As for the limitations of this work, the first one stems from the assumption of why a firm in the business world exists, because not all firms aim at generating positive cash flows. Firms, developed for reasons beyond generating positive cash flows, can behave in different ways than what is described in this paper. Second, although this paper looks at the distribution of challenges across the ecosystem of a focal firm, we did not consider the innovation challenges the focal firm faces in its core activities. These internal challenges might possibly alter some of the established results if they are somehow correlated with external challenges. Third, this study stays within the context of a single supply chain; however, in real life supply chains, as systems, also interact with each other. That interaction of a higher systemic level could very possibly affect the distribution of challenges within each of the interacting supply chains. In other words, this study only presents a baseline structure on top of which other factors and considerations need to be overlaid.

Corresponding to these listed and other unlisted limitations, future studies should incorporate additional contingencies and antecedents in order to produce more practical useful insights than what have been established in this paper.

REFERENCES

Adner, R. (2006). Match your innovation strategy to your innovation ecosystem. *Harvard Business Review*, *84*(4), 98-107.

Adner, R. (2017). Ecosystem as Structure: An Actionable Construct for Strategy. *Journal of Management*, *43*(1), 39-58.

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems: how the structure of technological interdependence affects firm performance in new technological generations. *Strategic Management Journal*, *31*(3): 306-333.

Adner, R., Oxley, J.E. and Silverman, B.S. (eds.) (2013). *Collaboration and Competition in Business Ecosystems*. Emerald: Bingley, UK.

Altinay, L. Madanoglu, M., Vita, G. D., Arasli, H., & Ekinci, Y. (2016). The interface between organizational learning capability, entrepreneurial orientation, and SME growth, *Journal of Small Business Management*, *54*(3), 871-891.

Anderson, P.C., & Tushman, M.L. (1990). Technological discontinuities and dominant designs. *Administrative Science Quarterly*, *35*(4), 604-633.

Argote, L. (1999). *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer: Norwell, MA.

Attour, A., & Lazaric, N. (2018). From knowledge to business ecosystems: emergence of an entrepreneurial activity during knowledge replication. *Small Business Economics*, https://doi.org/10.1007/s11187-018-0035-3.

Bernheim, B. D., & Whinston, M. D. 1990. Multimarket contact and collusive behavior. *RAND Journal of Economics, 21*: 1 – 26.

Bilgicer, T., Jedidi, K.,.Lehmann, D.R., & Neslin, S.A. (2015). Social contagion and customer adoption of new sales channels. *Journal of Retailing*, *91*(2), 254-271.

Bowman, C., & Ambrosini, V. (2001). “Value” in the resource-based view of the firm: A contribution to the debate. *Academy of Management Review*, *26*, 501-502.

Brem, A., & Viardot, É. (2015). Adoption of innovation: Balancing internal and external stakeholders in the marketing of innovation. In: Brem A., Viardot É. (eds) *Adoption of Innovation*. Springer, Cham, 1-10.

Brusoni, S., Prencipe, A., & Pavitt, K. (2001). Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make? *Administrative Science Quarterly*, *46*, 597-621.

Camps, J., Oltra, V., Aldás-Manzano, J., Buenaventura-Vera, G., and Torres-Carballo, F. (2016). Individual Performance in Turbulent Environments: The Role of Organizational Learning Capability and Employee Flexibility. *Human Resources Management*, *55*(3), 363-383.

Chang, W., Ellinger, A.E., Kim, K., & Franke, G.R. (2016). Supply chain integration and firm financial performance: A meta-analysis of positional advantage mediation and moderating factors. *European Management Journal*, *34*(3), 282-295.

Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate venture capital, value creation, and innovation. *The Review of Financial Studies*, *27*(8), 2434-2473.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, *32*(4), 4-39.

Christensen, C.M. (1997). *Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press: Boston,

Clarysse, B., Wright, M., Bruneel, J., & Mahajan, A. (2014). Creating value in ecosystems: Crossing the chasm between knowledge and business ecosystems. *Research Policy*, *43*(7), 1164-1176.

Cusumano, M., Mylonadis, Y., & Rosenbloom, R. (1992). Strategic maneuvering and mass market dynamics: the triumph of VHS over Beta. *Business History Review*, *66*(1). 51-94.

Forrest, YL. J., Buttermore, J., & Wajda, T. A. (2017). At Nash equilibrium when new market competitions appear? *Kybernetes, The International of Cybernetics, Systems and Management Science*, *46*(2), 256-271.

Fuentelsaz, L., & Gómez, J. (2006). Multipoint competition, strategic similarity and entry into geographic markets. *Strategic Management Journal, 27*(5), 477-499.

Fultz, D., Long, R.R., Owens, G.V., Bohan, W., Kaylor, R., & Weil, J. (1959). *Studies of Thermal Convection in a Rotating Cylinder with Some Implications for Large-Scale Atmospheric Motion.* Meteorol. Monographs (American Meteorological Society) vol. 21, no. 4.

Golder, P. N., & Tellis, G. J. (1993). Pioneer advantage: Marketing logic or marketing legend? *Journal of Marketing Research, 30*, 158-170.

Hatch, N.W., & Mowery, D.C. (1998). Process innovation and learning by doing in semiconductor manufacturing. *Management Science*, *44*(11), 1461-1477.

Haveman, H. A., & Nonnemaker, L. (2000). Competition in multiple geographic markets: The impact on growth and market entry. *Administrative Science Quarterly, 45*(2), 232-267.

Hellström, M., Tsvetkova, A., Gustafsson, M., & Wikström, K. (2015). Collaboration mechanisms for business models in distributed energy ecosystems. *Journal of Cleaner Production*, *102*(1), 226-236.

Henderson, R.M. (1995). Of life cycles real and imaginary: the unexpectedly long old age of optical lithography. *Research Policy*, *24*(4), 631-643.

Henderson, R.M., & Clark, K.B. (1990). Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, *35*(1) 9-30.

Hide, R. (1953). Some experiments on thermal convection in a rotating liquid. *Quart. J. Roy. Meteorol. Soc*., *79*, 161.

Hughes, T.P. (1983). *Networks of Power: Electrification in Western Society 1880-1930*. Johns Hopkins University Press: Baltimore.

Iansiti, M., & Levien, R. (2004). *The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*. Harvard Business School Press: Boston, MA.

Intel Corporation (2004). Intel sees unified platform and ecosystem as key to enabling the digital home. Press release 17 February.

Kagermann, H. (2015). Change through digitization: Value creation in the age of industry 4.0. In: Albach, H., Meffert, H., Pinkwart, A., & Reichwald, R. (eds) *Management of Permanent Change*. Springer Gabler, Wiesbaden, 23-48.

Kalmuk, G., & Acar, A.Z. (2015). The mediating role of organizational learning capability on the relationship between innovation and firm's performance: A conceptual framework. *Procedia - Social and Behavioral Sciences, 210* (2), 164-169.

Kapoor, R., & Agarwal, S. (2017). Sustaining Superior Performance in Business Ecosystems: Evidence from Application Software Developers in the iOS and Android Smartphone Ecosystems. *Organization Science*, *28*(3), 379-595.

Kline, M. (1972). *Mathematical Thought from Ancient to Modern Times*, Oxford: Oxford University Press.

Klir, G. (1985). *Architecture of Systems Problem Solving*, New York, NY: Plenum Press.

Klir, G. (2001). *Facets of Systems Science*, New York: Springer.

Laukkanen, T. (2016). Consumer adoption versus rejection decisions in seemingly similar service innovations: The case of the Internet and mobile banking. *Journal of Business Research*, *69*(7), 2432-2439.

Lemoussu, S., Chaudemar, J. C., and Vingerhoeds, R. A. (2018). Systems Engineering and Project Management Process Modeling in the Aeronautics Context: Case Study of SMEs. *International Journal of Mechanical and Mechatronics Engineering*, 12 (2). 88-96.

Li, S. X., and Greenwood, R. (2004). The effect of within industry diversification on firm performance: Synergy creation, multi-market contact and market structuration. *Strategic Management Journal, 25*(12), 1131-1153.

Lieberman, M.B. (1984). The learning curve and pricing in the chemical processing industries. *RAND Journal of Economics*, *15*(2), 213-228.

Lieberman, M.B. (1989). The learning curve, technological barriers to entry, and competitive survival in the chemical processing industries. *Strategic Management Journal*, *10*(5), 431-447.

Lieberman, M.B., & Montgomery, D.B. (1988). First-mover advantages. *Strategic Management Journal, 9,* 41-58.

Lieberman, M.B., & Montgomery, D.B. (2013). Conundra and progress: Research on entry order and performance. *Long Range Planning, 44*, 312-324.

Lin, Y. (1988). Can the World be Studied in the Viewpoint of Systems?? *Mathl. Comput. Modeling*, *11*, 738-742.

Lin, Y. (1999). *General Systems Theory: A Mathematical Approach*. New York: Plenum and Kluwer Academic Publishers.

Lin, Y. (2007). Systemic yoyo model and applications in Newton's, Kepler's laws, etc. *Kybernetes: The International Journal of Cybernetics, Systems and Management Science*, *36*(3-4), 484-516.

Lin, Y. (2009). *Systemic Yoyos: Some Impacts of the Second Dimension*. New York: Auerbach Publication, an imprint of Taylor and Francis.

Lin, Y., and Ma, Y. H. (1987). Some properties of linked time systems. *International Journal of General Systems*, *13*(2), 125-134.

Lin, Y., Ma, Y., & Port, R. (1990). Several epistemological problems related to the concept of systems. *Math. Comput. Modeling*, *14*, 52–57.

Lin, Y., & OuYang, S.C. (2010). *Irregularities and Prediction of Major Disasters*. New York: CRC Press, an imprint of Taylor and Francis.

Markides, C.C., & Geroski, P.A. (2005). *Fast second: How smart companies bypass radical innovation to enter and dominate new markets.* San Francisco: Jossey Bass/Wiley.

Marzi, G., Dabic, M., Daim, T., & Garces, E. (2017). Product and process innovation in manufacturing firms: A 30-year bibliometric analysis. *Scientometrics*, DOI 10.1007/s11192-017-2500-1.

McGrath, R.G. (2013). *The end of competitive advantage: How to keep your strategy moving as fast as your business*. Boston: Harvard Business Review Press.

Oliveira, T., Thomas, M., Baptista, G., & Campos, F. (2016). Mobile payment: Understanding the determinants of customer adoption and intention to recommend the technology. *Computers in Human Behavior*, *61*, 404-414.

Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press: New York.

Prajogo, D. I., & Oke, A. (2016). Human capital, service innovation advantage, and business performance: The moderating roles of dynamic and competitive environments. *International Journal of Operations & Production Management*, 36(9), 974-994.

Priem, R. L. (2007). A consumer perspective on value creation. *Academy of Management Review*, *32*, 219-235.

Rostow, W.W. (1960). *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge University Press, Cambridge, UK.

SAP Corporation. (2006). SAP ecosystem unites to bring innovation to chemicals industry. Press release, 8 May.

Spence, M. (1981). The learning curve and competition. *Bell Journal of Economics*, *12*(1), 49-70.

Sung, S. Y. & Choi, J. N., (2014). Do organizations spend wisely on employees? Effects of training and development investments on learning and innovation in organizations. *Journal of Organizational Behavior, 35*(3), 393-412.

Takeishi, A. (2002). Knowledge partitioning in the interfirm division of labor: the case of automotive product development. *Organization Science*, *13*(3), 321-338.

Tushman, M., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, *31*(3), 439-465.

Utterback, J.M., & Abernathy, W.J. (1975). A dynamic model of process and product innovation. *Omega*, *3*(6), 639-656.

von Bertalanffy, L. (1924). Einfuhrung in Spengler's werk. *Literaturblatt Kolnische Zeitung*, May.

Wu, Y., & Lin, Y. (2002). *Beyond Nonstructural Quantitative Analysis: Blown-ups, Spinning Currents and Modern Science*. River Edge NJ: World Scientific.

Yadav, P. L., Han, S. H., & Kim, H. (2017). Sustaining competitive advantage through corporate environmental performance. *Business Strategy and the Environment, 26*(3), 345-357.

Yu, T., & Cannella, A.A. Jr. (2012). A comprehensive review of multimarket competition research. *Journal of Management, 39*, 76-109.

Zahra, S.A., & Nambisan, S. (2012). Entrepreneurship and strategic thinking in business ecosystems. *Business Horizons*, *55*(3), 219-229.

Zachary, M.A., Gianiodis, P.T., Payne, G.T., & Markman, G.D. (2015). Entry timing: Enduring lessons and future directions. *Journal of Management*, *41*(5), 1388-1415.

DECISION SCIENCES INSTITUTE

Accelerating innovation by leveraging external knowledge: The impact of resource constraint

Tzy-wen Tang
National Chung Hsing University
Email: twtang@dragon.nchu.edu.tw

Rui-Ting Huang
National Chung Hsing University
Email: rthuang0324@dragon.nchu.edu.tw

Sheng-Chun Chang
National Chung Hsing University
Email: d104023005@mail.nchu.edu.tw

Ting-Yu Huang
National Chung Hsing University
Email: dot801108@gmail.com

ABSTRACT

This paper aims to study the impact of resource Constraint on accelerating innovation by leveraging external knowledge. We find the effects of resource constraints on innovation performance are varied on the types of resource constraint. This result differs from the results in some of the previous literature.

KEYWORDS: Open Innovation, Resource constraints, External knowledge search, and Knowledge Management

INTRODUCTION

The 21st century is the era of knowledge-based economies, where innovation is the principal focus of national policy. While companies attempt to enhance competitiveness and expand their market size, innovation can promote continuous business development for companies. It ensures the survival and development of companies that survive in viciously competitive environments. In recent years, innovation has become widespread due to its simplicity and the need to reduce costs. However, most innovations, at least at the grassroots level, are temporary devices created by indigenous creativity that achieve their goals under the constraints of various resources (Balkrishna C. Rao, 2018)

In knowledge-based theory, knowledge is characterized as a key corporate resource and a source of competitive advantage with great strategic value as a tool for creating value (Spender & Grant, 1996; Teece, Pisano, & Shuen, 1997). The continuous pursuit of new knowledge is essential for innovation and to achieve long-term profitability and competitiveness. Open innovation advocates that the strategic application of external resources outweighs the benefits of the confined use of internal knowledge in facilitating innovation (Chesbrough, 2003). Smith et al. (2005) found that the knowledge acquired by dedicated units within an organization is usually authentic, valuable, and difficult to imitate. Internal knowledge fundamentally helps companies discard unimportant information during the knowledge searching process. However, internal research and development (R&D) requires immense time and monetary investments and contains considerable risks. Therefore, companies can use external knowledge searches (EKS) to complement internal knowledge.

EKS strategies can help companies in their innovation efforts (Laursen & Salter, 2006; Levinthal & March 1981; Zhang & Li, 2008). Katila and Ahuja (2002) expressed that during technical and market uncertainty, the most effective action that companies can take is EKS to adjust and restructure internal organizations. The acquisition of external information and technical resources reduce the cost and complexity of technical company is insufficient to support rapid responses to market changes. Companies must appropriately utilize external sources or channels to acquire the latest resources and innovation information. Laursen & Salter (2006) also pointed out in his research that the more open and breadth or depth of search an organization has, the more innovative its ability is. Because openness plays an extremely important role in the innovation performance of an enterprise, There should be a focus on openness, good connections with Communities and Networks, but too much breadth or depth of EKS can actually be counter-productive to spend too much time and money. Different forms of organizational slack divergently influence a firm's open search strategy, contributing to the understanding of the relationship between organizational slack and knowledge search behavior in a broader context, as well as the understanding of the moderating effect of absorptive capacity. Many scholars have suggested that different external knowledge search strategies may produce different results (Chen et al., 2011; Laursen and Salter, 2006), which may be the cause of inconsistencies in past empirical results. In the depth and breadth of external knowledge search, past empirical results show that the breadth and depth of external knowledge search has no significant impact on innovation performance (Ferreras-Méndez et al., 2015), and some research results show that external knowledge breadth and radical Innovation has an inverted U-shaped relationship (Ardito & Petruzzelli, 2017), and some results show that breadth of external knowledge has only an inverted U-shaped relationship with incremental innovation, and has no significant impact on radical innovation (Hwang & Lee, 2010).

The objectives of this paper are: a) to analyze the relationships of the EKS strategies companies adopt on their innovation performance, and b) to evaluate the influences that companies' resource constraints have on their EKS strategies and innovation performance. In collation of the preceding research motivation, this paper examines the following issues: (1) Relationships of EKS on innovation performance. (2) Relationships of resource constraints on innovation performance. (3) Relationships of resource constraints on EKS.

LITERATURE REVIEW AND HYPOTHESES

External Knowledge Search and Innovation

Knowledge-based theory emphasizes that enterprises must demonstrate sufficient internal competency to manage knowledge effectively. It also focuses on the effects of knowledge acquisition models on the learning performance of organizations (Grant, 1996; Tidd & Trewhella, 1997; Poppo & Zenger, 1998).
Abeson and Taku (2009) categorized knowledge into internally generated knowledge (R&D) and externally acquired knowledge (cooperation and purchase). Probst et al. (2000) proposed the knowledge gap model and asserted that the key objectives of knowledge management are to bridge the knowledge gap in strategies and determine the internal and external knowledge required by enterprises, the knowledge enterprises currently have, the knowledge enterprises lack, and whether enterprises are able to self-generate the knowledge that they lack. When enterprises are unable to generate the necessary knowledge through R&D, they are required to acquire the knowledge externally. Innovation processes consist of actors searching for knowledge, and of combination and integration processes of that knowledge (Savino et al., 2015). Martini et al. (2016) .Laursen & Salter (2006) combines the ways in which researchers access knowledge and works on the relationship between external knowledge search and innovation performance and establishes a solid foundation for an open innovation strategy. Its strategy of depth and breadth is proposed for external knowledge search. "External search of breadth" refers to the number of enterprises 'innovation activities relying on external sources or search pipelines, while "deep external search" refers to enterprises' diversification from different external sources or Laursen & Salter also proposed 16 sources of knowledge (search pipeline) such as upstream suppliers, customers, competitors, governments, academic institutions, research institutes and so on.Open innovation advocates that the strategic application of external resources outweighs the benefits of the confined use of internal knowledge in facilitating innovation (Chesbrough, 2003). However, a number of scholars argue that knowledge-rich environments are able to reinforce the innovation capability of enterprises, even if they do not actively employ search strategies (Garriga, von Krogh, & Spaeth, 2013).
Innovation, defined as the application of unique methods to incorporate novel concepts into new products, services, or manufacturing processes (Urabe, 1988), involves a linear conversion of scientific knowledge into technical knowledge to improve product development outcomes (Gomory & Schmitt, 1988). Hence, innovation transforms creative ideas into useful products, services, or operating methods that are profitable for enterprises.
Enterprises can eliminate or change the current organizational protocols and routines through knowledge acquisition. Knowledge enables enterprises to make different decisions and achieve innovation by redefining existing markets or creating new markets (March, 1991; Miner et al., 2001). Enterprises engage in external searches to acquire external knowledge. They then use external knowledge sources to achieve innovation and reap immense profit (Menon & Pfeffer, 2003). EKS enhance enterprise innovation. However, such searches incur costs. Excessive knowledge searches (including excessive broad and in-depth knowledge searches and excessive outward exploration and exploitation knowledge searches) may cause an enterprise to waste unnecessary energy, time, and money, hindering enterprise performance.
Companies are increasingly competing in an open innovation environment. Therefore, the search strategy of external knowledge is decisive for the success of the company. Existing research distinguishes the breadth (diversity) and depth (strength) of firms dealing with external sources of knowledge (Iferd, Younes; Plötz, Patrick,2018). Laursen & Salter (2006) examines UK manufacturing knowledge search depth and breadth. The two search strategies for innovation performance (new product revenue share of business turnover), found that both exist inverted U-shaped relationship. In other words, the external search strategy can improve the

innovation performance of an enterprise, but accumulate to an extent without any subtraction because the extensive knowledge source pipeline or the deep search of too many source pipelines not only leads to cost increase, but also complexity Increase, so that enterprises can not absorb the huge external information. In addition, Laursen & Salter believes that in the general environment, the appropriate depth and breadth search lead the enterprise to achieve the highest level of innovation performance. The depth and breadth of external knowledge have different effects on innovation. The depth of external knowledge helps to generate groundbreaking ideas, novel ideas, and different past solutions and is therefore particularly helpful for radical innovation (Cohen & Levinthal, 1990; Schilling & Green, 2011), while a breadth of external knowledge Are more conducive to the gradual and progressive progress of the organization. However, it is also more difficult to solve the fundamental problem of breakthrough than the high possibility of novelty (Schilling & Green, 2011)

Although EKS helps to improve the innovation performance of enterprises, some costs still exist. With the improvement of EKS, enterprises may have a negative correlation between EKS and performance due to the increase of pipeline complexity, the increase of transaction costs and the increase of information needs and information processing costs. Therefore, "excessive" knowledge search, including excessive depth and breadth of knowledge search, may cause enterprises to waste too much energy and time costs, resulting in the Performance and performance reduction. In summary of the aforementioned discussion, the following hypotheses were proposed:

H1a: Depth EKS has inverse U-shaped relationship with radical innovation.
H1b: Depth EKS has inverse U-shaped relationship with incremental innovation.
H2a: Breadth EKS has inverse U-shaped relationship with radical innovation
H2b: Breadth EKS has inverse U-shaped relationship with incremental innovation

Resource Constraints and Innovation

The daily routines of enterprises, such as production, marketing, labor allocation, R&D, and finance, require various tangible and intangible resources to sustain normal operations. Previous studies concerning the resource constraints of enterprises proposed a variety of descriptive, explorative, conceptual, and qualitative variables. Hewitt-Dundas (2006) indicated that a type of interrelation exists between the constraint variables that affect innovation.The intensity of innovation in an enterprise depends on the availability of resources, primarily financial and human resources constraints (Trajkovski & Jovan, 2018).

However, studies on cognitive psychology have found that when resources are restricted, individuals are more able to be creative and resolve innovation-related problems (Durham et al., 2000; Moreau & Dahl, 2005). Iammarino et al. (2009) found a positive correlation with the cognitive constraints and innovation tendencies of enterprises. These findings suggest that resource constraints may facilitate enterprises’ innovation activities. From a macro-environment perspective, the application constraints of corporate resources refer to intrinsic conditional constraints, external regulatory constraints (e.g., the human genome research in the field of life science), and country backgrounds (e.g., lack of professionals or lack the requirements to enter key markets). These forms of constraints may affect the enterprises’ innovation behavior (Shan, 1990). Garriga, von Krogh, and Spaeth (2013) found a negative correlation between resource application constraints and enterprise innovation. Moreover, national and regional regulatory restrictions are also factors obstructing innovation.

A consensus has yet to be reached concerning the impact of resource constraints on innovation. However, the findings of most studies indicated negative correlations between resource constraints and innovation. Vossen (1998) asserted that a number of enterprises

encounter personnel constraints when engaging in external searches. Large companies are able to attract highly specialized professionals and are therefore more capable of innovation. For SMEs, Resources are often limited and these constraints can interfere with their ability to innovate. Therefore, in order to meet research objectives, it is crucial to analyze the theory of the impact of corporate resources on innovation(Woschke, Tino; Haase, Heiko; Kratzer, Jan, 2017). Barber et al. (1989) asserted that it can be a challenge for smaller companies to recruit experienced, highly skilled employees and to have to continuously invest in personnel training and technology development, thus restricting innovation performance (Brown et al., 1990). Moreover, the lack of R&D professionals and/or experienced professionals may impact enterprises' innovation behaviors (Shan, 1990). The execution of EKS and the acquisition of external knowledge require R&D teams and innovation professionals. Therefore, enterprises that lack such teams and professionals are less able to engage in innovation. Hence, the following hypothesis was proposed:
H3: Personnel constraints has a negative effect on radical innovation and incremental innovation.
Knowledge constraints are also an issue that organizations cannot easily overcome and solve because innovation-related knowledge is often tricky and not readily available and not easily understood, therefore, it will not be conducive to the radical innovation activities of enterprises, but will promote some systems and processes depending on incremental innovation of the previously-used search channels and methods (Ahuja & Lampert, 2001; Carlile, 2004). Keupp & Gassmann (2013) assumes that there is a negative impact between knowledge constraints and radical innovation. The following hypothesis was proposed:
H4a: Knowledge constraints has a negative effect on radical innovation.
H4b: Knowledge constraints has a positive effect on incremental innovation.
Capital adequacy helps companies accelerate the introduction of new products without missing out on opportunities and creating maximum benefits (Patzelt et al., 2008). Therefore, companies without adequate funding will not be able to support relevant organizational activities, improve their products, conduct small-scale key technological improvements, introduce attractive products, and gain solid market share (Hsu-Hua et al., 2011) . From a financial resource point of view, the source of funding is a significant consideration for firms before they engage in innovation activities. Keupp & Gassmann (2013) assumes that financial constraints have a positive impact on radical innovation. Hewitt-Dundas (2006) also stated their relationship in the findings, arguing that there is a gap between financial constraints and innovation positive relationship, companies that face financial constraints or have experienced financial distress find themselves under pressure and enable organizations to engage in innovation-related activities, both in product development and business strategy, and companies can take advantage of this The chance of success of its innovation.
H5a: Financial constraints has a positive effect on radical innovation.
H5b: Financial constraints has a negative effect on incremental innovation.

Resource Constraints and External Knowledge Search

Laursen (2012) asserted that enterprises are met with resource constraints when they engage in EKS. Therefore, enterprises must take into account limited organizational resources and current conditions when engaging in broad and in-depth external searches in order to make accurate fit and trade-off decisions (Liu & Chen, 2013). Moreover, the majority of scholars believe that enterprises with resource constraints do not have adequate support and are less able to engage in in-depth EKS. The reason is because such searches are costly. However, internal resources also affect trade-offs and dilemmas concerning exploration and exploitation

(Argote & Ingram, 2000). Therefore, resource application constraints may form different search strategies and modes (Argote & Ingram, 2000). In terms of resource constraints, personnel and knowledge constraints are key aspects of knowledge acquisition. Hence, negative innovation resource constraints may hinder enterprises' EKS.

For many businesses, knowledge constraints are a problem that cannot easily be overcome because innovation-related knowledge is often not readily available and not easily understood (Ahuja & Lampert 2001; Carlile 2004) In the absence of knowledge, innovation-related people may not know how to obtain market information, technology, etc. Zhang et al. (2008) indicated that enterprises may be detrimental to finding out-of-industry knowledge when limited by the use of resources. Hewitt-Dundas (2006) argues that there is a correlation between the limiting variables that affect innovation. From a knowledge point of view, for example, Roberts (1999) finds that an organization's R & D-related resources will affect knowledge restrictions and their innovations, ie, not enough Of the research funding, cannot support the R & D team to get more technology and knowledge and further breakthroughs, because R & D always requires constant trial and error, and the opportunity cost is high, therefore, knowledge restrictions will not be conducive to external knowledge search activities .The following hypothesis was proposed:

H6: Personnel constraints has a negative effect on depth EKS and breadth EKS.

H7: Knowledge constraints has a negative effect on depth EKS and breadth EKS.

Studies on cognitive psychology have found that when resources are restricted, individuals are more able to be creative and resolve innovation-related problems (Durham et al., 2000; Moreau & Dahl, 2005). Iammarino et al. (2009) found a positive correlation with the cognitive constraints and innovation tendencies of enterprises. These findings suggest that resource constraints may facilitate enterprises' innovation activities. From a macro-environment perspective, the application constraints of corporate resources refer to intrinsic conditional constraints, external regulatory constraints (e.g., the human genome research in the field of life science), and country backgrounds (e.g., lack of professionals or lack the requirements to enter key markets). These forms of constraints may affect the enterprises' innovation behavior (Shan, 1990). Garriga, von Krogh, and Spaeth (2013) found a negative correlation between resource application constraints and enterprise innovation. Moreover, national and regional regulatory restrictions are also factors obstructing innovation. In the absence of internal finance, firms are more inclined to conduct knowledge search activities abroad because organizations need financial support for many activities, such as research and development, marketing or innovation activities. Therefore, Survival and competitiveness are often accelerated by the introduction of new products as the faster the introduction of new products, the more profit can be derived from them, thereby reducing the overall risk and stress (Keupp & Gassmann, 2013).Although financial limitations contribute to external knowledge search, personnel limitations and knowledge restrictions are an important part of knowledge-related knowledge. Therefore, the negative effect will lead to the overall resource limitation is not conducive to external knowledge search. The following hypothesis was proposed:

H8: Financial constraints has a positive effect on depth EKS and breadth EKS.

Based on the concept of Laursen & Salter (2006), this study divides external knowledge search into depth and breadth and discusses its impact on innovation performance. In addition, this study attempts to explore the impact of resource constraints on external knowledge search and innovation performance of enterprises. Including the impact of financial resource constraints, knowledge resource constraints and personnel resource constraints on innovation performance and the depth and breadth of external knowledge search. The research frame is illustrated as figure 1.

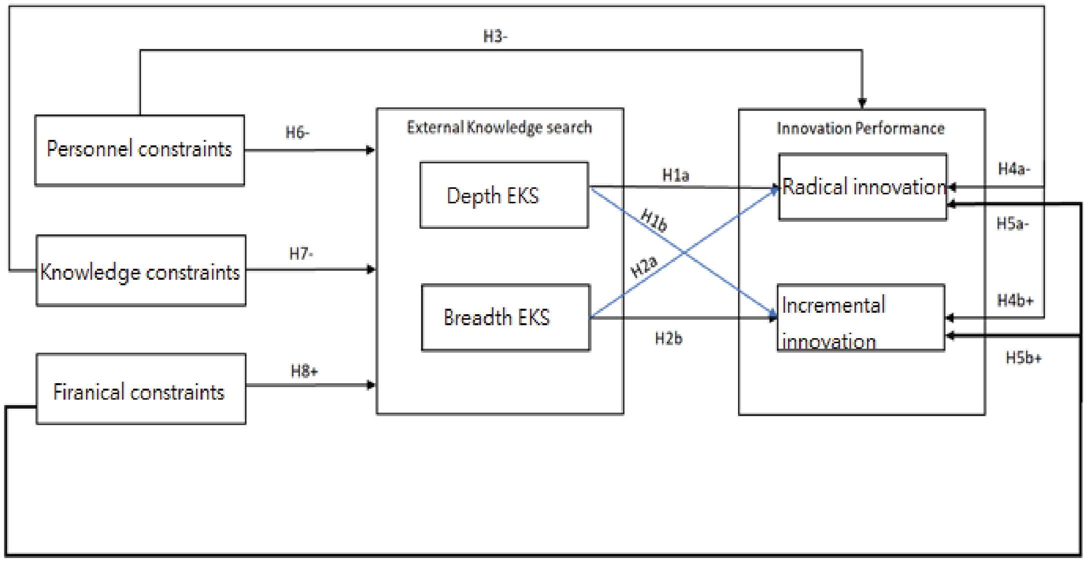


Figure 1. Research Framework

METHODOLOGY

Sampling

In this paper, the top one-thousand Taiwanese manufacturers reported in *Common Wealth Magazine* in 2015 were selected as the targets for a questionnaire survey. The reasons for selecting manufacturers are as follows: (1) Taiwan's economic growth is largely based on manufacturing, and the development of the manufacturing industry is a key aspect of future innovation and investment (*Common Wealth*, 2015). (2) The manufacturing industry centers heavily on developing new techniques and products. (3) Laursen and Salter (2006) selected manufacturers in England as the sample population. The questionnaires were administered via surface mail. One thousand questionnaires were sent and 118 were received, among which one invalid questionnaire was omitted. The final number of valid questionnaires was 117, with a recovery rate of 11.7%. In the second round of the survey, a convenient sampling approach was adopted in the form of an online questionnaire. A total of 65 questionnaires were received. Therefore, a total of 182 valid questionnaires were collected for the second sampling round.
A test for non-responsive bias was performed to identify the potential non-responsive biases in the questionnaire recovery process. The recovery channels of the questionnaires were first classified in chronological order. An independent *t*-test and chi-square test were performed to make a pair-wise comparison of the samples and determine whether biases existed in the research variables and demographics. The samples collected in the first sampling round, that is, those collected from the top one-thousand Taiwanese manufacturers via surface mail were characterized as the early recovery group, which comprised 117 samples. Those collected in the second sampling round, that is, those administered electronically to friends and relatives in the manufacturing industry, were characterized as the late recovery group, which comprised 65 samples. Significant effects confirmed in the *t*-test and chi-square test outcomes denoted that

the significant differences existed between the early and late recovery groups. Actual test results indicated no significant differences between the various variables and demographics in the two groups.
To prevent the subjectivity of a single respondent from creating common source biases in the questionnaires, the questionnaire was evaluated using research program design theory before administration (i.e., eliminate underlying preferences in the item descriptions) to reduce common source bias. Once the questionnaires were recovered, Harman's one-factor test was performed to evaluate the severity of common source bias (Podsakoff & Organ, 1986). A principal component analysis was adopted as the factor analysis without rotation. A total of 15 eigenvalues greater than 1 were extracted and examined. No one-factor conditions were observed. The explanatory variance of the first principal component was 23.121%. Hence, it was unable to explain the majority of the variance, implying that the research data did not contain significant common source biases

Sample Characteristics

Taiwanese manufacturers were the research participants in this paper. In terms of industry distribution, the majority were metal production and processing vendors (N=26; 14.3%). In terms of company age, almost half of the manufacturers (48.9%) have been in operation for over 31 years, suggesting that most of the samples were from sustainable enterprises. In terms of scale, most of the samples (51.6%) were large companies with more than 250 employees, suggesting that most of the companies engaged in large-scale operations. In terms of position, most of the respondents were managers or deputy managers (32.4%), suggesting that the majority of the respondents were in management roles and demonstrated adequate managerial experience

Variable Measurement

The EKS scale was developed based on the findings of Laursen and Salter (2006). The scale comprised 11 knowledge sources, including upstream suppliers (e.g., raw material, equipment, and software suppliers), clients, competitors, consultants, exhibition/display organizers, trade unions (e.g., export and fair trade unions), government-owned research institutes (e.g., Academia Sinica, Industrial Technology Research Institute, Institute for Information Industry, and National Health Research Institutes), other public/government departments (e.g., Ministry of Economic Affairs, Ministry of Labor, and Directorate-General of Budget, Accounting, and Statistics), professional conferences (e.g., domestic/foreign conference and international forums), university research institutes, and private research institutes. The respondents measured the innovation contribution of the knowledge and information acquired by the 11 knowledge sources to their companies. A four-point scale was used for scoring, where 0 denotes that the knowledge source made no innovation contribution to the company, 1 denotes a slight innovation contribution, 2 denotes some innovation contribution, and 3 denotes a large innovation contribution. A higher score signified a more extensive knowledge search.
The resource constraints scale was developed based on the findings of Keupp and Gassmann (2013), Hewitt-Dundas (2006), and Garriga and Spaeth (2013). The scale comprised six items to measure personnel constraints, financial constraints, and knowledge constraints. A Likert five-point scale was used to measure the items (1=very unsatisfied and 5=very satisfied). A higher score signified higher resource constraints perception.
Managerial ties refer to the individuals' external contact through their relationships and interpersonal networks. The managerial ties scale was developed based on the findings of Qin et al. (2010) and Li et al. (2014). The scale comprised nine items. A Likert five-point scale was

used to measure the items (1=very unsatisfied and 5=very satisfied). A higher score signified stronger ties between the respondents and external experts.
The innovation performance scale was developed based on the findings of Leal-Rodríguez et al. (2014) and Kim et al. (2012). The scale comprised nine items to measure radical product, service, and procedural innovation, and six items to measure incremental product, service, and procedural innovation. A Likert five-point scale was used to measure the items (1=very unsatisfied and 5=very satisfied). A higher score signified a higher innovation performance perception.
The control variables comprised industry type, number of employees, technical uncertainty, and market uncertainty. Industry type was set using dummy variables and chategorized into high-technology industries and ultra-high-technology industries. Damanpour (1992) asserted that company size and employee size could affect innovation performance in one of two ways. One is the generation of laziness because the company is too large to change, reducing innovation performance. The other is the increase in innovation performance because large companies contain adequate resources. Atuahene-Gima and Li (2004) examined past innovation-related issues and found that technical uncertainty and market uncertainty are the two key factors influencing IP. The technical uncertainty scale was developed based on the findings of Atuahene-Gima and Li (2004). It comprised four items. The market uncertainty scale was developed based on the findings of Jaworski and Kohli (1993). It comprised three items.

Reliability and Validity

Reliability measures the consistency and stability of evaluation and measurement tools and tests the internal consistency of the scale composition. In this paper, the Cronbach's α coefficient was adopted as the reliability measure. An increased coefficient signified a higher internal consistency. Nunnally (2010) proposed that the reliability reaches an acceptable range when the Cronbach's α coefficient is greater than 0.7. Based on this standard, the Cronbach's α coefficients for the various variables are tabulated in Table 1. The table shows that the Cronbach's α coefficients for all the variables were higher than 0.749, suggesting that the scales used in this paper achieve excellent reliability.

Table 1. Reliability Analysis

Variables	Items	Factor loading	Cronbach's α	CR	AVE
Personnel constraints	1. The company lacks "R & D personnel"	0.84	0.749	0.799	0.665
	2. The company lacks "management professionals"	0.79			
Financial constraints	1. The Company lacks "funds (including self-financing or borrowing money from banks)"	0.83	0.847	0.892	0.805
	2. The Company lacks R & D funding	0.96			
Knowledge constraints	1. The company lacks "technical knowledge"	0.96	0.799	0.881	0.789
	2. The Company lacks "market knowledge"	0.81			
Radical Innovation	1. Our"technological competitiveness" is higher than that of our competitors	0.70	0.901	0.953	0.692

	2.	Our "speed of technological innovation" is faster than our competitors	0.85			
	3.	Our "process, method update faster than competitors	0.89			
	4.	Our products / services have the latest technology	0.86			
	5.	We can grasp the most innovative processes, methods	0.87			
	6.	Our "new level of product / service" is great	0.87			
	7.	Our "new product development speed" is very fast	0.83			
	8.	We often introduce new products to the market	0.80			
	9.	Our products / services are unique, novel	0.80			
Incremental Innovation	1.	We have made minor improvements in "technology and technology"	0.68	0.877	0.900	0.547
	2.	Our products / services are further improvements on the original products / services	0.70			
	3.	We introduce new products that are more incremental (ie, less innovative) than our peers	0.84			
	4.	The proportion of our incremental innovation products, higher than their peers	0.91			
	5.	The Company's revenue from incremental innovative products has been growing year by year	0.88			
	6.	For customers, we are famous for progressive product innovation	0.88			

Content validity measures the appropriateness and representation of the scale content, or rather whether the scale encompasses all desired research aspects. All variables measured in this paper contained theoretical bases and were formulated based on an extensive literature review and an empirical questionnaire survey. Prior to formal testing, an expert scholar and six industry experts were invited to review and revise the questionnaire. Therefore, the scales used in this paper achieved excellent content validity.

Moreover, a confirmatory factor analysis was performed to test the convergent validity and discriminant validity of the scales. For convergent validity, the standards proposed by Fornell and Larker (1981) were adopted. They were: (1) the standardized loading between the items and their latent variables must be greater than .5; (2) the average variance extracted (AVE) must be greater than .5; and (3) the composite reliability (CR) must be greater than .6. The CR values for all the variables were between .8 and .95, meeting the recommended value of .6, and the AVE values were between .5 and .87, meeting the recommended value of .5. Hence, all variables converged at their respective factor dimensions, suggesting excellent convergence validity.

RESULTS

A regression analysis was performed to validate the proposed hypotheses and examine the associations among the variables. The Durbin-Watson (DW) value was adopted to determine whether the errors were auto-correlated. The DW values for the models ranged between 1.746 and 2.294, suggesting no severe autocorrelation in the errors. A regression analysis was performed on the variables to validate the hypotheses proposed in this paper.
H1a claims that depth EKS has an inverted U-curve relationship with radical innovation. The regression results are tabulated in Table 2. The results in the Model 2 column indicated no inverted U-curve for the relationship between depth EKS and radical innovation (β =0.141). Therefore, H1a is rejected. The results in the Model 4 column indicated a U-shaped curve for the relationship between depth EKS and incremental innovation (β =0.277). Therefore, H1b is supported the results in the Model 2 column indicated an inverted U-curve for the relationship between breadth EKS and incremental innovation (β =1.420; P<0.05). Therefore, H2a is supported. The results of the Model 4 column in Table 2 indicated Breadth EKS has inverse U-shaped relationship with incremental innovation (β =0.500; P<0.001). Therefore, H2b is supported. H3 claims that personnel constraints has a negative relationship with radical innovation and incremental innovation (β = -0.46,-0.076). Therefore, H3 is rejected. H4a claims that knowledge constraints have a negative relationship with radical innovation The results of the Model 2 column in Table 2 indicated a significant and negative correlation between personel constraints and innovation performance (β =-0.259; P<0.001). Therefore, H4a is supported. H4b claims that knowledge constraints have a positive relationship with incremental innovation (β =0.080). Therefore, H4b is supported. H5a claims that financial constraints has a postive relationship with innovation (β =0.0100). Therefore, H5a is supported. H5b claims that financial constraints have a negative relationship with innovation (β =-0.132). Therefore, H5b is supported.

Table 2. Effects of Resource Constraints and External Knowledge Search on Innovation Performance

	Radical Innovation		Incremental Innovation	
	Model 1	Model 2	Model 3	Model 4
Industry type	-0.64*	.056*	-0.15	-0.017
Number of employees	-.094**	-.117***	.016	-0.012
Technical uncertainty	.345***	.243***	.473***	.424***
Market uncertainty	.091**	.114***	-.028	-0.035
Personnel constraints		-0.46		-0.076***
Knowledge constraints		-0.259***		0.080
Financial constraints		0.0100***		-0.132***
EKS				
Depth EKS		0.141*		0.277***
Depth EKS^2		-0.154*		-0.213**

Breadth EKS		1.420***		0.500*
Breadth EKS^2		-0.204		-0.392*
F-value	18.458	36.853	16.914	41.781
R^2	.169	.258	.209	.253
Adj R^2	.167	.252	.207	.246
$\triangle R^2$	0	.085	0	.039

n=182, *p<0.05, **p<0.01, ***p<0.001

H6 claims that personnel constraints has a negative relationship with depth EKS and breadth EKS.. The results of the Model 7 column in Table 3 indicated personnel constraints has a positive relationship with depth EKS(β =0.218) . The results of the Model 10 column in Table 3 indicated personnel constraints has a positive relationship with breadth EKS(β =-0.161) Therefore, H6 is partially Supported . H7 claims that knowledge constraints have a negative relationship with depth EKS and breadth EKS. The results of the Model 7 column in Table 3 indicated knowledge constraints has a negative relationship with depth EKS (β =-0.212). The results of the Model 10 column in Table 3 indicated knowledge constraints has a positive relationship with breadth EKS (β =0.042) Therefore, H7 is partially Supported. H8 claims that financial constraints have a positive relationship with depth EKS and breadth EKS. The results indicated no significant correlation between financial constraints and depth EKS (β =0.002). The results of the Model 10 column in Table 3 indicated financial constraints has a positive relationship with breadth EKS (β =o.56). Therefore, H8 is partially supported.

Table 3. Effects of Resource Constraints on External Knowledge Search

	Depth EKS			Breadth EKS		
	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Industry type	0.019	0.019	0.027	0.081**	0.82**	-0.075***
Number of employees	0.095**	0.054*	0.094***	0.213***	0.144***	0.155***
Technical uncertainty	0.325***	0.246***	.0241***	-0.023	-0.146***	-0.163***
Market uncertainty	-0.060	-0.063*	-0.091**	0.058	0.050	0.037
Resource constraints		-0.024			0.33	
Personnel constraints			0.218***			-.161***
constraints			-.212***			0.042

Knowledge constraints			0.002			0.56***
Financial constraints						
-value	36.272	46.748	49.807	15.214	42.675	47.827
R^2	0.101	0.177	0.280	0.053	0.250	0.284
Adj R^2	0.099	0.173	0.275	0.050	0.246	0.279
$\triangle R^2$	0	0.074	0.102	0	0.196	0.033

n=182, *p<0.05, **p<0.01, ***p<0.001

In addition Figure 2 shows inverted U-shaped relationship between deep external knowledge search and innovation performance, with the highest point of innovation performance 4.14 at depth search level of 4。In other words, the external search strategy can improve the innovation performance of an enterprise. However, after accumulating to an extent, it will not be subtracted because the search of too many source channels will not only lead to cost increase, but also increase the complexity so that enterprises cannot absorb External huge information. Figure 3 shows that there is an inverse U-shaped relationship between breadth-of-external knowledge search and innovation performance, with the highest point of innovation performance 3.17 at a breadth-of-search level of 7。In other words, the external search strategy can improve the innovation performance of an enterprise. However, after accumulating to a certain extent, it will not be subtracted because the search for a pipeline of excessive breadth EKS not only leads to cost increase, but also increases the complexity and makes the enterprise unable to Absorption of huge external information.

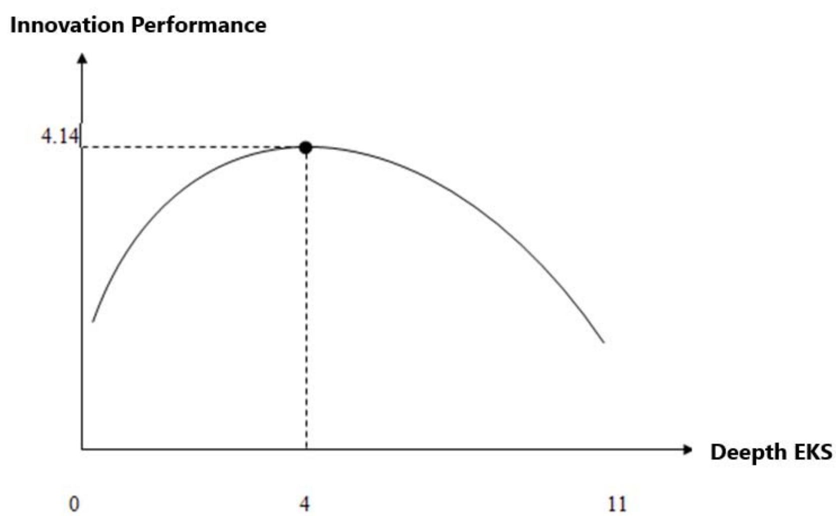


Figure 2. Depth of external knowledge search and innovative performance

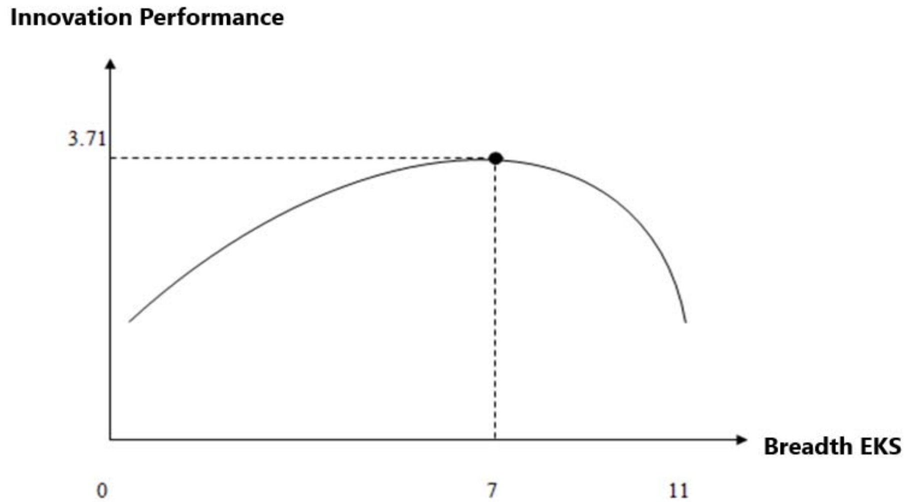


Figure 3. Breadth of external knowledge search and innovation performance

CONCLUSION AND FUTURE RESEARCH

Conclusion and Implications

This paper aims to elucidate the effects of resource constraints, breath and depth of EKS on innovation performance. The conclusion and theoretical contributions of this paper are presented in this section.

Empirical findings show that different EKS strategies affect radical and incremental innovation performance differently. This paper hypothesizes that depth and breadth of EKS has an inverted U-curve relationship with innovation performance; wherein an increase in EKS would initially improve innovation performance , but excessive EKS would reduce innovation performance. Empirical results showed that depth of EKS and innovation performance failed to has a significant inverted U-curve relationship (H1a was rejected); depth of EKS has an inverted U-curve relationship with incremental innovation (H1b was supported); and EKS and radical innovation failed to has an inverted U-curve relationship, but rather a U-shaped curve relationship (H1a was rejected). The inverted U-curve relationship exhibited between depth of EKS, and incremental innovation was similar to the findings of Laursen and Salter (2006). However, the relationship between depth of EKS and radical innovation was inconsistent with the results proposed by Laursen and Salter (2006). Rather, a U-shaped curve relationship was observed. The reason for the rejection of H1a may be the offset of contradicting effects between radical innovation and incremental innovation and EKS. Because EKS exhibited opposite effects on radical and incremental innovation, the offset of the two effects confirms that the effects of EKS on innovation performance are insignificant.

The most appropriate external search strategy for a business will depend on the type of business innovation (Garriga & Spaeth et al., 2013). When a company adopts incremental innovation, the use of external sources of knowledge will be more important than radical innovation (Moon, 2014). That is, incremental innovation relies more on the use of open EKS, which requires more breadth of EKS pipeline than depth of EKS. Although EKS helps to improve the innovation performance of enterprises, some costs still exist. With the improvement of EKS, enterprises may have a negative correlation between EKS and performance due to the increase of pipeline complexity, the increase of transaction costs and the increase of information needs and information processing costs. Therefore, "over-performing" knowledge search, including excessive depth and breadth of knowledge search, may cause enterprises to waste too much energy and time costs, resulting in lower business performance.

Although many scholars have previously investigated the success factors of corporate innovation, the competencies required to achieve innovation, and the importance of organizational learning, few have examined the internal resource constraints that could potentially threaten enterprises' innovation activities. In this paper, the empirical results personnel constraint rejected H3, indicated no significant correlation between personnel constraint and innovation performance. These results are inconsistent with a number of previous studies (Hewitt-Dundas, 2006; Garriga et al., 2013). From a resource-based perspective, enterprises are able to perform internal audits to determine the relationship between input resources and innovation performance (Hadjimanolis, 1999). Although large manufacturers have considerably more resources at their disposal, they may still be challenged with resources inadequacy. These constraints may affect organizations' innovation performance.

Few scholars have investigated the relationship between internal resource constraints and the knowledge activities of organizations. Based on the conclusion of Laursen et al. (2012), this paper argues that enterprises may be constrained by limited resources while engaging in EKS. In this paper, the empirical results resource constraints partially supported hypotheses, suggesting that resource constraints failed to significantly affect EKS. These results are prompted by one of two reasons. First, resource constraints may have a positive influence on EKS. A number of studies have indicated that certain resource constraints facilitate enterprises in implementing external search strategies (Gupta et al., 2006; Zhang Li, 2008; Schilling & Green, 2011). Because the effects of resource constraints on EKS activities contain positive and negative aspects, they offset each other. Second, different types of resource constraints produce different effects. For example, financial resource constraints may produce positive effects, while labor and knowledge resource constraints may produce negative effects. These effects offset one another, thereby partially supporting hypotheses.

Practical Implications

In terms of resource constraints, this research can help administrators to identify key internal resources and understand the effects of these factors on search strategies before they engage in depth and breadth of EKS or innovation activities. This paper also helps future scholars elucidate the effects of resource constraints on EKS strategies and highlight the relationships and contradictions between innovation activities and performance. For example, the provision of an open education environment to foster employees' perseverance and willingness to engage in lifelong learning, encourage further education, and/or participate in social exchanges, and suitably provide encouragement and incentives, can generate substantial knowledge sources and foster highly technical employees, thereby improving organizational performance.

Empirical results showed that EKS strategies facilitate innovation performance. Therefore, enterprises should utilize their information sources and channels from various levels, such as establishing R&D alliances and participating in conferences to help them overcome the problem of narrow search ranges and enhance innovation performance. However, excessive broad and in-depth searches could harm the enterprise by wasting the energy and time cost of the organization. Therefore, enterprises must seek to balance and trade-off searches. Before engaging in external searches or innovation activities, enterprises must evaluate their type of innovation they wish to achieve, such as incremental or radical innovation; wherein the former requires a shorter amount of time with relatively lower risk but less beneficial, or the latter, which requires more time and is riskier and costly but yields greater results. Moreover, successful innovation relies on the external searches of innovators. Once external knowledge sources are acquired, innovators can determine whether these sources can be internalized and their ideas commercialized. This process requires a series of market surveys and repeated trial and errors.

Limitation and Future Research

The design of this paper sought to meet the criteria of scientific research. However, several research limitations were apparent. First, a list of Taiwanese manufacturers proposed by *Common Wealth Magazine* in 2015 was used as the population for data collection. However, because the respondents were largely executives from the manufacturing industry, the research results may not apply to other industries. In addition, this paper adopts a cross-sectional research design, implying that the dynamic changes of resource constraints, depth and breadth of EKS, and innovation performance could not be determined. Future researchers can adopt a similar research framework and use a more diverse set of samples. Alternatively, they can also adopt a longitudinal research design coupled with qualitative analysis to examine the long-term development of the research variables.

Second, the questionnaire survey was administered via surface mail; wherein each enterprise was sent a copy of the questionnaire. The questionnaires were completed by executives or R&D teams. However, answers may be subjected to the subjectivity of the respondent. Therefore, the common source bias problem must be taken into account. Therefore, a statistical method was applied to elucidate whether common source bias existed in the recovered questionnaires. To prevent the problem of subjectivity, future researchers can survey a wider range of respondents, where one questionnaire is provided to different employees within the enterprise. This process can generate more objective results and reduce the effects of common source bias.

Third, the cause and effect variables influencing depth and bread of EKS and innovation are abundant and complex. This paper only adopts a resource-based perspective to analyze internal resource constraints such as personnel constraints, financial constraints, and knowledge constraints. Future researchers can collect a wider range of variables influencing resource constraints and categorize these variables into internal and external variables, tangible and intangible variables, and/or controllable and uncontrollable variables to facilitate their research.

REFERENCES

Abeson, F., & Taku, M. A. (2009). Knowledge source and small business competitiveness. *Competitiveness Review: An International Business Journal*, *19*(2), 88-95.

Ahuja, G., Lampert, C. M., & Tandon, V. (2008). Moving beyond Schumpeter: Management research on the determinants of technological innovation. *The Academy of Management Annals*, *2*(1), 1-98.

Argote, L. and Ingram, P. (2000), ‘‘Knowledge transfer: a basis for competitive advantage in firms’’, Organizational Behavior and Human Decision Processes, Vol. 82 No. 1, pp. 150-69

Balkrishna C. Rao. (2018). Science is indispensable to frugal innovations, *Technology Innovation Management Review*, 8(4), 49-56.

Brown, C. J., Abbas, P. J., & Gantz, B. (1990). Electrically evoked whole‐nerve action potentials: Data from human cochlear implant users. *The Journal of the Acoustical Society of America, 88*(3), 1385-1391.

Brown, S. L., & Eisenhardt, K. M. (1998). *Competing on the edge: Strategy as structured chaos*. Boston, MA: Harvard Business School Press.

Carlile, P., R. (2004). Translating, and Transforming: An Integrative Framework for Managing Knowledge across Boundaries,*Organization Science* 15(5), 555-568.

Chen, J., Chen, Y., & Vanhaverbeke, W. (2011). The influence of scope, depth, and orientation of external technology sources on the innovative performance of Chinese firms. *Technovation*, 31(8), 362–373.

Chesbrough, H. (2003). The logic of open innovation: managing intellectual property, *California Management Review, 45*(3), 33-58.

Chesbrough, H., Vanhaverbeke, W., & West, J. (2006). Open innovation: Researching a new paradigm. Oxford, England, Oxford University Press.

Cohen & Levinthal, (1990). **Absorptive capacity: a new perspective on learning and innovation,** *Administrative Science Quarterly*, 35 (1) , 128-152

Damanpour, F. (1987). The adoption of technological, administrative, and ancillary innovations: Impact of organizational factors. *Journal of Management,* 13(4), 675-688.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston, MA: Harvard Business School Press.

Ferreras-Méndez, J. L., Newell, S., Fernández-Mesa, A., & Alegre, J. (2015), Depth and breadth of external knowledge search and performance: The mediating role of absorptive capacity, *Industrial Marketing Management*, 47, 86-97.

Garriga, H., von Krogh, G., & Spaeth, S. (2013). How constraints and knowledge impact open innovation. *Strategic Management Journal*, *34*(9), 1134-1144.

Gomory, R. E., & Schmitt, R. W. (1988). Science and product. *Science, 240*, 1131-1132.

Grant, R. M. (1996). Toward a knowledge‐based theory of the firm. *Strategic management journal, 17*(S2), 109-122.

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The interplay between exploration and exploitation. *Academy of Management Journal*, 49(4), 693-706.

Hadjimanolis, A. (1999). Barriers to innovation for SMEs in a small less developed country. *Technovation, 19*(9), 561-570.

Hewitt-Dundas, N. (2006). Resource and capability constraints to innovation in small and large plants. *Small Business Economics, 26*(3), 257-277.

Hwang, J. & Lee, Y. (2010).External knowledge search, innovative performance and productivity in the Korean ICT sector, *Telecommunications Policy*, 34(10),.562–571.

Iferd, Younes; Plötz, Patrick(2018). **IDEAS Working Paper Series from RePEc .**

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal, 45*(6), 1183-1194.

Keupp, M. M., & Gassmann, O. (2013). Resource constraints as triggers of radical innovation: Longitudinal evidence from the manufacturing sector. *Research Policy, 42*(8), 1457-1468.

Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. *Strategic Management Journal, 27*(2), 131-150.

Laursen, K., Masciarelli, F., & Prencipe, A. (2012). Regions matter: how localized social capital affects innovation and external knowledge acquisition. *Organization Science*, 23(1), 177-193.

Liu, J., Chen, L., & Kittilaksanawong, W. (2013). External knowledge search strategies in China's technology ventures: The role of managerial interpretations and ties. *Management and Organization Review*, 9(3), 437-463.

Levinthal, D., & March, J. (1981). A model of adaptive organizational search. Journal of Economic Behavior and Organization, 2, 307–333.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*(1), 71–87.

Martini, A., Neirotti, P., & Appio, F. P. (2016). Knowledge searching, integrating and performing: Always a tuned trio for onnovation? Long Range Planning. 50 (2), 200-220.

Menon, T., & Pfeffer, J. (2003). Valuing Internal vs. External Knowledge: Explaining the preference for outsiders, *Management Science*, 49(4), 497- 513.

Moreau, C. Page and Darren W. Dahl (2005). Designing the Solution: The Impact of Constraints on Consumers' Creativity, *Journal of Consumer Research*, 32(1), 13-22.

Moon, S. (2014). How does the use of external knowledge influence innovative performance of service firm? An introductory study of openness and service innovation. *Seoul Journal of Business*, *20*(1), 35-62.

Patzelt, H., & Shepherd, D. A. (2008). The decision to persist with underperforming alliances: The role of trust and control, *Journal of Management Studies*, 45(7), 1217–1243.

Poppo, L., & Zenger, T. (1998). Testing alternative theories of the firm: transaction cost, knowledge‐based, and measurement explanations for make‐or‐buy decisions in information services. *Strategic Management Journal, 19*(9), 853-877.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects, *Journal of Management,* 12, 69–82

Probst, T. M. (2000). Wedded to the job: Moderating effects of job involvement on the consequences of job insecurity. *Journal of Occupational Health Psychology,* 5, 63—73

Qin, W., & Shanxing, G. (2010). Managerial ties and innovative performance: an open innovation perspective. *Proceedings of the 7th International Conference on Innovation & Management*, China, Wuhan University of Technology

Savino, T., Messeni Petruzzelli, A., & Albino, V. (2017). Search and recombination process to innovate: A review of the empirical evidence and a research agenda. *International Journal of Management Reviews*. 19 (1), 54-75

Schilling, M.A., Green, E.: Recombinant search and breakthrough idea generation: an analysis of high impact papers in

the social sciences. Res. Policy 40(10), 1321–1331 (2011)

Shan W. 1990. An empirical analysis of organizational strategies by entrepreneurial high-technology firms. Strategic Management Journal 11(2): 129–139.

Spender, J. C., & Grant, R. M. (1996). Knowledge and the firm: overview. *Strategic Management Journal, 17*(S2), 5-9.

Standifird, S. S., & Marshall, R. S. (2000). The transaction cost advantage of guanxi-based business practices. *Journal of World Business, 35*(1), 21-42.

Smith, K. G., Collins, C. J., & Clark, K. D. (2005). Existing knowledge, knowledge creation, capability, and the rate of new product introduction in high-technology firms. *Academy of Management Journal*, 48(2), 346-357.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509-533.

Tidd, J., & Trewhella, M. J. (1997). Organizational and technological antecedents for knowledge acquisition and learning. *R&D Management, 27*(4), 359-375.

Trajkovski, Jovan(2018).The oretical Framework for the Study of Intangible Investment Into Innovative Capital in Resource Limited Environment: A Case for Synchronous Innovations?, *Economic and Business Review for Central and South - Eastern Europe***,**20(1),51-83***.***

Urabe, K., Child, J., & Kagono, T. (1988). Innovation and management: International comparisons. Berlin, Germany: Walter de Gruyter.

Vanhaverbeke, W., Duysters, G., & Noorderhaven, N. (2002). External technology sourcing through alliances or acquisitions: An analysis of the application-specific integrated circuits industry. *Organization Science, 13*(6), 714-733.

Vossen, R. W. (1998). Relative strengths and weaknesses of small firms in innovation. *International Small Business Journal, 16*(3), 88-95.

Woschke, Tino; Haase, Heiko; Kratzer, Jan, (2017). Resource scarcity in SMEs: effects on incremental and radical innovations, *Management Research Review*, 40(2), 195-217.

Zhang, S. J., & Li, X. C. (2008). Managerial ties, firm resources, and performance of cluster firms. *Asia Pacific Journal of Management*, *25*(4), 615-633

DECISION SCIENCES INSTITUTE

Real Options-Based Strategizing for Competitive Advantage from IT Vendors' Perspective

Suvankar Ghosh
University of Toledo
2801 Bancroft Street
Toledo, OH 43606
Suvankar.Ghosh@utoledo.edu
(419) 530-4163

ABSTRACT

We develop a soft operational research heuristic called the CURVO Grid for IT vendors for competitive strategizing that combines real options logic with several theories of competitive advantage, including Porter's competitive forces model, the resource-based view (RBV), dynamic capabilities approach, core competence theory, and the theory of the ambidextrous firm.

KEYWORDS*:* Information technology strategy, real options, resource-based view, dynamic capabilities approach, core competency theory, theory of the ambidextrous firm

INTRODUCTION

Real Options (RO) is both a conceptual framework and a quantitative valuation methodology and it is here to stay. The chief difference between RO and traditional capital budgeting methodologies, such as the net present value (NPV), is that RO treats investment opportunities as options (Dixit and Pindyck 1995, Luehrman 1998). When investment opportunities are viewed as options, then timing is everything. One does not necessarily have to make the decision now. It is eminently reasonable to keep the option of investing in that opportunity open to gather more information on the value of the opportunity. This stands in stark contrast to the traditional NPV analysis, which takes a "now-or-never" approach to investing (Dixit and Pindyck 1995, Luehrman 1998). In the RO approach, the investment decision does not necessarily have to be made at time $t = 0$. If the uncertainty surrounding the investment opportunity is high, one could wait till time $t = 0 + \Delta t$ to make the decision and use the time Δt for more information to arrive, which resolves some of the uncertainty surrounding the investment opportunity. For example, new information could arrive about how some other firm fared with making a similar investment; or it could be about a technology standard related to the investment that emerged as the winner among several competing standards.

The time Δt need not necessarily be spent passively waiting for new information to arrive from the external environment that resolves some of the uncertainty. This time could well be spent in actively learning about new technologies associated with the investment. For example, if a firm is contemplating investment in some complex information technology (IT) such as the service-oriented architecture (SOA) and Web Services (Erl 2005, Krafzig et al. 2005), a firm without deep roots in this technology could use the time Δt in doing a small pilot project to acquire knowledge about the benefits and pitfalls of SOA. Hence, the RO approach to capital budgeting inherently has the element of timing when it is best to make the investment in much the same way that a holder of a financial call option is timing when to exercise the call option (Dixit and Pindyck 1995, Luehrman 1998).

As our literature review shows later, RO logic in the information systems (IS) discipline has been applied mainly from the standpoint of a user or buyer of technology to value the investment in the face of uncertainty and managerial flexibility to change course depending on how the future unfolds. RO logic has not been used from the perspective of the IT vendor on how to gain competitive advantage in the markets of their products and services. In this paper, we develop a software operational research heuristic called the CURVO Grid to help IT vendors with developing competitive strategy. Within an RO framework, we integrate various theories of competitive advantage including Porter's five competitive forces model (Porter 1980, 1985), Barney's resource-based view (RBV) (Barney 1991, 2001), the dynamic capabilities approach (Teece et al. 1997), the core competency theory (Prahalad and Hamel 1990), and ambidexterity theory (O'Reilly III and Tushman 2004). The CURVO Grid is a normative repertory grid (Eden and Jones 1984) that can be used to guide managerial decision-making on the right options and option chains to exercise depending on the state of the firm and the environmental uncertainties the firm faces. The CURVO Grid thus helps management avoid the pitfalls of pursuing misguided strategies.

LITERATURE REVIEW

The power of the RO approach is manifested in the wide range of studies conducted in diverse disciplines. In operations management, the RO approach has been used in different contexts such as switching between different modes of production or between different products as prices and demand change (Kulatilaka 1998, Triantis and Hodder 1990), expanding manufacturing capacity (Kumar 1995, Karsak and Özogul 2002), and modularizing production systems (Gamba and Fusari 2009). In research and development (R&D) in the pharmaceuticals industry, Cassimon et al. (2004) develop a complex multi-stage option to value investment in new drug discovery. Jensen and Warren (2001) have applied RO in evaluating R&D projects at British Telecommunications. In IS literature, RO has been used in various applications including Integrated Services Digital Networks (ISDN) (Dos Santos 1991), electronic banking (Benaroch and Kauffman 2000), multistage IT investments (Bardhan et al. 2004, Benaroch et al. 2006a), migrating from SAP R/2 to R/3 (Taudes 1998), and risk management (Benaroch 2002, Benaroch et al. 2006b, Benaroch et al. 2007).

Apart from studies performed by academics, RO is catching on in industry itself. Graham and Harvey (2001) found that, in a survey of 392 CFOs, approximately 27% of these CFOs are adopting real options analysis in capital budgeting. RO is being utilized in industries as diverse as bio-technology, manufacturing, and natural resources (Miller and Park 2002). As with any emerging paradigm with the potential for far-reaching ramifications for corporate finance, there will of course be dissenting views and conflicting evidence as to the popularity of the new paradigm. Baker et al. (2015) present a less rosy picture about the take-off of RO. In their study of 215 Canadian firms, they found that only 16.8% were using RO, and that RO actually ranked the last of the nine capital budgeting techniques they were using. In a similar vein, a study conducted by Bain and Co. in 2000 of 451 senior executives across more than 30 industries on their use of management tools, only 9% reported using RO (Teach 2003). Sobering as these statistics are for proponents of RO, it is perhaps more useful to probe the true source of the business world's problem with RO and focus efforts on correcting that problem.

The real problem that appears to be stymieing the embrace of RO by the practitioner world is that RO is too mathematical and complicated to implement (Baker et al. 2015, Teach 2003). Teach states, "The sophisticated mathematics of real options (such as partial differential equations) and the consequent lack of transparency and simplicity are real concerns." Baker et al. echo this sentiment and state, "Managers report that a lack of expertise and knowledge prevents them from using real options." Hence, the real problem to address about RO is devising ways to present the technique in a more transparent and similar fashion to practitioners that eschews gory details of partial differential equations and presents the insights provided by RO in a clear and compelling fashion. Two noted experts in RO, Martha Amram and Nalin Kulatilaka, do precisely that in their book on applying RO to manage strategic investments in an uncertain world (Amram and Kulatilaka 1998).

Following up on Amran and Kulatilaka's (1998) work, several efforts are on-going on presenting RO in a clearer fashion to practitioners. Feinstein and Lander (2002) attempt to make RO conceptually simpler by showing that RO can be recast as a form of NPV analysis by choosing an appropriate value for the discount rate. Ghosh and Troutt (2012) take the approach of open-sourcing the algorithms and the software they built for valuing complex compound options. Open-sourcing the software for evaluating options makes the valuation methodology of RO much less of a black-box, since other RO experts can look at the algorithms and use the associated software to verify if the algorithms for implementing the RO valuation models are indeed correct. This can generate trust among managers who may not have the mathematical background to directly understand the RO valuation models and algorithms but can see what the RO community at-large is saying about the open-source software for evaluating RO models. Apart from the open-sourcing approach for RO software tools, another approach is to develop heuristic frameworks for the application of RO in decision-making. Ghosh et al. (2013) develop a soft operational research (OR) heuristic based on the use of a repertory grid (Eden and Jones 1984) for guiding managerial action on investing in enterprise integration (EI) technology.

Many of the studies on the use of RO logic in the IS field have been done from the perspective of a user or buyer of IT. There is also not much discussion of competitive advantage in these studies. These studies of RO also do not tie in how the investment being made in IT results in competitive advantage in the product and service markets the firms operate in. The focus is mainly on building the quantitative options pricing model (OPM) to value the IT investment opportunity without demonstrating how competitive advantage flows from this investment. To make a case for competitive advantage, one would need to tie this investment to theories of competitive advantage such as Porter's five competitive forces model (Porter 1980, 1985), Barney's resource-based view (RBV) (Barney 1991, 2001), the dynamic capabilities approach (Teece et al. 1997), the core competency theory (Prahalad and Hamel 1990), and ambidexterity theory (O'Reilly III and Tushman 2004). Is the IT investment being made by the firm a unique resource not easily replicable by other firms (Barney 1991, 2001)? Does it renew or enhance a core competency of the firm (Prahalad and Hamel 1990)? Does it serve to create barriers to entry or serve to unbalance competitors thereby entrenching the firm in its market (Porter 1980, 1985)? These questions are generally not answered in RO studies in IS.

We review three key studies in RO in this section to demonstrate that their focus is mainly on building a quantitative OPM for valuing a cash flow stream with managerial flexibility with regard to future investments. There is little or no treatment of achieving competitive advantage in the product and service markets the firm competes in. Bardhan et al. (2004) apply RO to the case of a firm in the utilities industry that is evaluating 31 IT projects to be implemented in a phased fashion over three years where there are sequencing constraints across these projects. Bardhan et al. (2004) develop an OPM that determines the optimal sequencing of these projects given the project interdependencies and the annual budget constraints. Their OPM is a nested model where single-period Black-Scholes (1973) options pricing models corresponding to specific projects are nested within each other to compute the present value of the full, multi-period investment program. There is little discussion in the article of how these IT projects are expected to result in competitive advantage for the firm in the market for energy services that the firm serves. For example, some of these IT projects enable electronic billing and payment. But there is no discussion of what the competitors are doing and whether electronic billing and payment represents an innovation in this market. It could also be that the firm is playing catch-up with its competitors and electronic billing and payment is emerging as a requirement in this market; and the firm in question has to support this to prevent a loss of its customers to competitors. Hence, without a discussion of the competition and competitive advantage, it is not clear if the utilities firm in question is exercising a growth option or a survival option in implementing these IT projects. Some of these IT projects could indeed be for survival purposes, while others may be targeted at growth. Without framing the exercise of real options in the context of the competition makes it very difficult to assess if any competition advantage is accruing to the company from its actions.

Taudes (1998) builds an OPM for analyzing the benefits to a firm of deploying the client-server SAP/R3 version of an enterprise resource planning (ERP) system as compared to the SAP/R2 mainframe-based version of the system. The SAP/R3 deployment has the advantage that it could serve as a platform for value-added services such as electronic data interchange (EDI). Taudes (1998) develops an OPM for this scenario which demonstrates the value of having the flexibility to layer value-added services such as EDI at a later stage. Again, there is no discussion of competitive advantage in this article. The industry in which the firm operates is not even mentioned and it is unclear if the firm’s competitors are using EDI or not. Using EDI may result in some cost-savings for the firm as manual paper-based transactions are moved to an electronic format, but this may not be the source of any sustained competitive advantage as other firms may quickly adopt EDI if they are not already doing so. A reduction of internal inefficiencies is not tantamount to competitive advantage.

Benaroch and Kauffman (2000) develop an OPM for valuing the expansion of Yankee 24, a firm that provides electronic banking services, into point-of-sale (POS) debit services. This study is an RO analysis of whether it would be better for Yankee 24 to immediately enter the market for POS debit services in the New England region or to exercise the defer option and wait a while to enter the market. As it is waiting, it can see if POS debit services took off in other parts of the country, and also actively lobby for changes in Massachusetts regulations to create a more benign environment for POS debit services. The competition is addressed to some extent in this study as it factors in the market signal that Yankee 24’s principal competitor NYCE shows no signs of entering the POS debit services market. So while competition is addressed in this study, it should be noted that competitive advantage is not just about how to get there but also very much about how to stay there. Hence, the all-important aspect of sustained competitive advantage is not addressed in this study.

In this article, we look at how to use RO logic in creating sustained competitive advantage and we do so from the perspective of a vendor of IT rather than a user of IT. This article thus fills an important gap in the use of RO in IS literature. In IS literature, RO has been mainly applied from the standpoint of the user or buyer of IT and, furthermore, little attempt has been made in linking the decision to invest in some key IT to how this creates competitive advantage for the firm. We present in this article a qualitative framework in the form of a repertory grid that can guide IT vendors in leveraging RO logic to strengthen their strategies for achieving competitive advantage. As we argued earlier, the mathematical complexity of RO is proving to be too daunting for many practitioners (Baker et al. 2015, Teach 2003). However, there are clear conceptual benefits to using RO logic in managerial decision-making even if the quantitative OPM is difficult to build. Hence, in this article, we build a qualitative decision-making framework in the form of a repertory grid to guide IT vendors in applying RO logic for enhancing competitive advantage.

There have been efforts in the past to formulate qualitative frameworks incorporating the real options perspective to inform managerial thinking. Fichman who has advocated the use of RO in the IS field (Fichman 2004a) developed a four-perspective framework (Fichman 2004b) on the drivers of option value in IT platform adoption decisions. His framework sheds light on the positive and negative influences on option value in IT platform decisions and is based theories from the areas of strategy (Barney 1991, 2001; Wernerfelt 1984, 1989, 1995), organizational learning (Cohen and Levinthal 1990, Kogut and Zander 1992), managerial fashion and bandwagons (Abrahamson 1996, Abrahamson and Rosenkopf 1997), and technology adaptation (Tyre and Orlikowski 1994, DeSanctis and Poole 1994). Insights that his framework provides include: the option value of a technological innovation increases if it contributes to a firm's absorptive capacity, if the knowledge barriers to its absorption are not too high, if it is susceptible to network externalities, if it cannot be readily replicated by competitors, and if it is adaptable to a given firm's organizational context.

Fichman's (2004b) qualitative framework is at a more abstract level than the repertory grid that we describe in this article. Our repertory grid is more of a normative heuristic that can inform IT vendors on promising avenues to follow as well as pitfalls to avoid as they formulate their competitive strategies. Heuristics essentially allow practitioners to operationalize abstract theories. The AACSB has advocated with some urgency that academic research must become more practitioner relevant (AACSB International, 2008), and this issue has also been noted by researchers in various fields such as Benbasat and Zmud (1999) in IS and Gopinath and Hoffman (1995) in strategy. It is precisely in bridging this gap between abstract academic theory and what practitioners can operationalize that the field of soft operational research or "soft OR" with its focus on graphical techniques such as cognitive maps (Eden 2004) and repertory grids (Eden and Jones 1984) has much to offer.

An exemplar soft OR technique that has successfully bridged theory and practice is the widely used Strengths-Weaknesses-Opportunities-Threats (SWOT) matrix. Kua lists the SWOT matrix as among the six most widely utilized soft OR techniques (Kua 2016). It can be argued that it is the SWOT matrix that has brought the structure-conduct-performance theories of industrial organization (Mason 1949, Bain 1959, Schmalensee 1989, Schmalensee and Willig 1989) within easy reach of the practitioner. Our **C**apability-**U**ncertainty **R**eal **V**endor **O**ptions or CURVO Grid takes the SWOT matrix as its guiding light in developing a repertory grid that distills RO and associated theories of competitive advantage into a heuristic that vendors can use for developing competitive strategy. Ghosh et al. (2013) did something similar but, not surprisingly, their focus was on the user or buyer of IT and not a vendor of IT products and services. Indeed much of the work in the information systems (IS) discipline that has leveraged the RO approach in decision-making has been from the perspective of a buyer and user of IT rather than a vendor of IT. RO logic has not been applied directly from the standpoint of how RO-based decision-making can be used by IT vendors to gain competitive advantage in the IT product and service markets they serve. Hence, this study and the one by Ghosh et al. (2013) have fundamentally different perspectives.

THEORETICAL DEVELOPMENT

Our CURVO Grid is based on the soft OR technique of a repertory grid which seeks to classify a set of elements against a set of dimensions or constructs (Eden and Jones 1984). The elements that are being classified into the repertory grid are chains of real options, or a sequence of linked real options, that a vendor can exercise to improve revenues and enhance its competitive position. For example, a real options chain could be to bundle certain products and services into a package; if the market reaction to the package is favorable, then the package's price could be reduced to garner even more revenue; and this could be followed by enhancing the service and support available on the products in the package.

In an NPV valuation of a program of actions performed over time and the cash flows resulting from these actions, the future actions in the program are fixed *ex ante*. In an RO analysis, the action performed at some future time *t* depends on the state that one finds oneself in at *t*. For example, a reduction in the price of a package to ignite demand depends on whether the package got a favorable market reaction. Hence, the value of an options chain, which some authors refer to as the strategic net present value (SNPV) (Park and Herath 2000) differs from that of a fixed chain of actions because of the managerial flexibility that exists in changing any future action among a set of actions available at a future time. SNPV of an options chain is thus:

$$SNPV_{options\ chain} = NPV_{fixed\ action\ program} + Value\ of\ Managerial\ Flexibility \quad (1)$$

Although Equation (1) looks conceptually simply, the second term, or the value of managerial flexibility, can be notoriously hard to calculate as there can be many future states and many different actions that the firm can take depending on what the future brings. However, this is precisely where a qualitative framework is useful because it allows us to have a discourse on the real options and option chains that are attractive relative to others without actually building an OPM to determine the value of the second term.

Elements of CURVO Grid

The two dimensions of the CURVO Grid are uncertainty and firm capability and the elements being classified into the appropriate cells of the grid are real option chains. An individual real option exercised by the vendor can be any of the following:

- Alter price up/down – the vendor can change the price of the offering up or down. Multiple such actions can be taken over time. To learn about the price elasticity of the product, the vendor could lower the price in a geographically limited pilot, before rolling out the price change to the entire market.
- Introduce new features in existing offering – From introducing high-quality cameras in smartphones to sophisticated cluster analysis functionality in business analytics software, introducing new features is a must for IT vendors. This is a requirement for product differentiation players, which is one of the market strategies that Porter (Porter 1980, 1985) identified as resulting in sustained competitive advantage.
- Introduce new and innovative offering – Harvesting electricity from the air to power communication devices is an example of high-impact innovation (Collins 2018). Innovation can of course be less dramatic than that. There are many innovative apps being developed for smartphones and tablets. The kiwi.com app helps consumers book the cheapest flight by searching 18 billion cheap flight combinations hourly to get the cheapest fare (Kiwi.com 2019).
- Increase research and development (R&D) investment – Developing new and innovative products needs increased investment in R&D. R&D expenditure could be of two types:
    - Pursuit of incremental innovation – Incremental innovation is about making improvements to a firm's existing line of products and services, such as adding new features to the firm's offerings.
    - Pursuit of breakthrough innovation – Breakthrough innovation is about coming up with an entirely new and innovative product. The replacement of analog photography by digital photography which Kodak failed to do is example of breakthrough innovation (O'Reilly and Tushman 2004).
- Enter into technology partnerships (incremental or breakthrough) – Technology partnerships in the IT industry are done quite frequently to acquire necessary technology that the internal research and development (R&D) organization of the firm has not been able to deliver. Again, these technology partnerships could be in support of incremental changes in the firm's existing products or to build entirely new and innovative products.
- Acquire companies to get new technologies (incremental or breakthrough) - Sometimes the technology may be crucial enough for the IT vendor to simply acquire the company that has the technology rather than enter into a technology partnership arrangement to license the technology. Often this path may be followed if the acquiring company is a large company and the target company is much smaller in size. As is the case with technology partnerships, these corporate acquisitions could be in support of incremental changes in the firm's existing products or for building entirely new and innovative products.
- Diversify into adjacent markets with new yet related product – Technology acquisitions can enable an IT vendor to get into a related market. A PC manufacturer acquiring a tablet company would be an example of this.
- Improve the quality of the product – The IT offering could be made more resistant to breakage if it is a hardware product, or more resistant to crashing if it is a software product via rigorous quality assurance (QA) processes In fact, making

smartphones water-resistant which is essentially a quality issue is being seen as a key capability of the smartphones introduced in 2019 (TelcoWorld 2019)

- Step up marketing and advertising – Sometimes stepping up investing in marketing and advertising is the solution for a flagging product. Here, the firm can leverage its knowledge about the impact that previous advertising campaigns have had on sales.
- Open-source the software – Open-sourcing a software product can be used to build a following behind that product. There are many successful open-source products such as Linux and MySQL. The loss of revenue from software licenses can be made up by revenue from selling service and support on the open-sourced software.
- Enhance service and support – Sometimes the product, be it hardware or software, might be great but the service and support is found to be lagging. In such cases, a number of measures can be taken such as offering a variety of service contracts, offering more effective help desk services, investing in customer relationship management (CRM) systems to improve customer satisfaction with service delivery.
- Introduce packages bundling products and services – Create bundles of products and services and price the package attractively. This can help ignite demand for the products and services in the bundle.
- Diversify geographically with existing product – This could involve diversifying geographically within the same country or internationally.
- Do a limited pilot to test market response – Before making a full-scale commitment to sell in a new geographical region, the IT vendor may do a small pilot to test the market and understand differences in consumer preferences between the new region and the ones that it currently sells to.
- Enter into reseller partnerships – Reseller partnerships allow IT vendors to sell their offerings in geographical regions that is not reachable by their own sales force or their online selling channels. This is often the case for diversifying internationally. If the IT vendor does not have a presence in a certain country, then it may enter into a joint venture with a partner in that country to sell their products and services.
- Acquire companies for geographical distribution – A reseller relationship with a partner in a different geographical region, either within the country or abroad, could turn into a corporate acquisition later if the IT vendor determines that that geographical region is crucial to their business.
- Partner with third parties for value-added functionality – Partnering with third-parties is crucial for IT vendors to bring a full solution to their customers. For example, the apps on a smartphone are what brings useful functionality to the users. Yet most of these apps are not built by Google which builds the Android operating system but are built by third-parties.
- Build an eco-system of third-parties – Success in markets for IT offerings often involves building an effective eco-system of various players working collaboratively. For example, in the smartphone market, Google builds the Android operating system, the phones themselves are built by companies such as Motorola, Samsung, and HTC, the apps are built by a variety of third-party players, and the full smartphone system has to then be tested against carriers such as Verizon, Spring, and T-Mobile and their 4G and 5G services. This eco-system does not magically create itself but is something that has to be assiduously nurtured and developed.

- Exit from the market for the offering – Although many of the real options described above strengthen an IT vendor's position in a given market, the vendor may choose to exercise the option of exiting from a given market if it sees that its position in that market is not salvageable. An exit decision could also be triggered if the market is no longer seen as central to the firm's business strategy.
- Defer taking any action – The IT vendor could at any point in time simply defer taking any of the actions described above and choose to wait to gather more information before making any further moves. Depending on the information that arrives in the waiting period, it could resume the option chain it was executing or modify it suitably depending on the information that arrived and then proceed with the modified chain.

The real options defined above from a vendor perspective stand in contrast to the real options that are appropriate from the standpoint of a firm investing in IT as a user of IT (Benaroch 2002, Ghosh et al. 2013) . User real options are:

- Defer – postpone making any investment at this time until more is known about the opportunity
- Stage – Implement one facet of the proposed IT investment. Cash flows generally occur after all facets have been completed.
- Pilot – Implement a small-scale version of the full investment.
- Invest organization-wide – Implement on a large scale, such as investing in an IT enterprise-wide.
- Alter scale – Change the scale of the investment which could include either scaling up or scaling down.
- Alter scope/switch – Change the scope of the investment which could include adding new functionality, changing inputs and outputs, or changing processes for transforming inputs into outputs.
- Strategic growth – Link to a related but new opportunity.

From Options to Option Chains

As we noted earlier, firms don't simply exercise a single option, but they exercise a number of related options over time in an option chain. Exercising a chain of related options over time to increase product sales, build customer satisfaction and loyalty, increase market share, diversify to new markets, and introduce innovative follow-on products is essentially a strategy. As Bowman and Hurry (2003) state, "Strategies are produced by the sequential striking of this options chain." We use this notion of the equivalence between an option chain and a strategy in our CURVO Grid. The grid provides insights into the types of option chains or strategies that are best pursued depending on the state of the firm and the environment. But first we classify different types of option chains available to management.

- Operational Option Chains – These chains are part and parcel of the normal routine of IT vendors such as creating a package bundling in some existing products and services, waiting to see the market reaction to the package, reducing the price of the package if the market reaction is positive to ignite demand, and then enhancing service and support on the package to build customer satisfaction and loyalty. Increasing R&D expenditure for incremental innovation that introduces new features to the firm's existing offerings can also be part of an operational option chain. Management could decide to increase

R&D expenditure in a given product to introduce some innovative features in the next version of the product, launch the new version, gauge market reaction, and if the reaction is positive then step up marketing and advertising investment in the new version of the product.

- Learning Option Chains – Learning option chains are about the pursuit of the next big thing. Every firm has to keep its antenna raised and continue to explore turns in the technology and the market in order not to be blind-sided by them as Kodak was in the paradigm change from analog to digital photography. Learning option chains may involve not just exploring paradigm-shifting technologies but also major new markets, as is the case when a solely US-based business tries to go global. In such a case, a firm could decide to explore certain foreign markets by making a small investment such as a joint venture with a foreign partner to sell its offerings there before jumping in with investing in a foreign operation with its own offices and sales and marketing personnel. Learning option chains are also experimental in that if the prognosis turns out to be not very good, they can be abandoned easily since they don’t represent a major sunk cost. A learning option chain could be initiated by an internal R&D project to build a prototype followed by beta testing of the prototype in a limit market pilot to gauge the market reaction. If the market reaction is negative, the firm may simply exercise the abandon option as the last option in the chain.
- Strategic Growth Option Chains – When the firm decides to make a significant commitment to a new technology, such as a car company in the business of selling traditional cars invests in the manufacture and sales of hybrid cars, then that constitutes exercising a strategic growth option. A major international diversification move for a US-based company would also be considered as exercising a strategic growth option. The casting of a strategic growth option represents the culmination of a learning option chain and the committed pursuit of the growth opportunity. In other words, the experimenting with the new technology in the learning phase has been sufficiently fruitful for management to decide to jump into the new opportunity with full commitment. Investing in the manufacture of the product, and in what volume, is of course one of the real options to be exercised in the strategic growth option chain. Many subsequent options would need to be cast in terms of the investment to be made in the sales, marketing, advertising, and servicing of the new product. In other words, the strategic growth option chain starts when the learning option chain ends.

Dimensions of the CURVO Grid

Uncertainty

The value of an option or option chain depends on the level of uncertainty in the context of the decision. In fact, the more uncertain the environment is, the more is the value of having real options that could be exercised depending on how the future unfolds. There are several facets to the uncertainty facing the vendor that influence the value of the real options held by the vendor. These are:

- Technology – Technological uncertainty is a key factor in formulating market strategies of IT vendors. The pace of technological change can upend the best laid plans of IT vendors. The very architecture of a product may become quickly obsolete under the advent of new technology. This is what happened in the enterprise application integration (EAI) market. There were two camps of vendors in this market. Suite players offered EAI suites, which would tightly integrate many functions into their suite product such as data transformation and business process management (BPM). They found that the market shifted to other EAI vendors who had an architecturally different solution based on an enterprise services bus (ESB) (Fiorano ESB 2009, Progress Software ESB 2009). ESB vendors brought to the market a more flexible architecture where services bus was a platform for layering services such as data transformation and business process management. These value-added services could be added on top of the services bus at a pace that the customer decided upon.

- Market – Market uncertainty arises from many different sources including changes in consumer preferences, merges and acquisitions that change the structure of the market, the entry of new competitors, and the health of the broader economy. A new entrant to the market may come in with a product with new technology that disrupts the existing market structure. A case in point is the entry of Vonage into the telecommunications market with voice over IP (VOIP) phones that was disruptive to established players like Verizon and ATT.
- Standards and Regulations – Changing standards and regulation are also a source of uncertainty for vendors. In fact, the VOIP space and how it is to be regulated has generated controversy and attendant uncertainty. Vendors have been clamoring for minimal regulation whereas the US Federal Communications Commission (FCC) has sometimes taken positions on policies that are at odds with what the VOIP service carriers would like to see (Gross 2003). This had created a state of confusion and uncertainty in this space.
- Firm-Specific – Firm-specific uncertainty and risks include possible changes in management, a lawsuit being filed against the firm, being acquired by some other firm, having to fight a proxy war launched by influential shareholders, and a lack of key resources and funding to support the firm's market strategy.

Capability

The other key dimension in the CURVO Grid for competitive strategizing is capability. Over the last three decades, the theories that have had the most impact on the field of strategy are the core competency theory (Prahalad and Hamel 1990), the resource-based view (RBV) (Barney 1991, 2001) and its extensions such as the dynamic capabilities approach (DCA) (Teece et al. 1997). A firm's long-term competitive advantage in a certain market rests on the unique resources (Barney 1991, 2001) that it possesses that cannot be quickly replicated by its competitors. Teece at al.'s (1997) notion of dynamic capabilities builds on the RBV by drawing a distinction between static resources and what is needed by the firm to succeed in a rapidly-evolving Schumpeterian world of creative destruction (Schumpeter 1934). In fast-changing environments, it is not the static resources themselves but the ability of management to quickly shape these resources and configure a combination of resources to meet the needs of changing markets that results in competitive advantage (Teece et al. 1997). The notions of unique resources and dynamic managerial capabilities also find resonance in the core competency theory of competitive advantage put forth by Prahalad and Hamel (1990) where they exhort companies to focus on what they are best at and divest themselves of all non-critical functions. This view of core competency or "stick to the knitting" was the basis of much of the corporate restructuring that took place in the late 80's and early 90's.

The inward-looking view of competitive advantage which asserts the primacy of internal resources, capabilities, and competencies is at odds with Porter's (Porter 1980, 1985) view of competitive advantage. Porter's view is outwardly focused at the market. Porter's position was that a firm should look at the market landscape; find a white space where there aren't too many competitors but there is an unfulfilled need; enter the market with an offering targeted at the white space; and then entrench itself by erecting barriers to entry for newcomers, building customer and supply intimacy, and keeping competitors off-balance through product differentiation where new and innovative features are constantly added to products. As it turned out, Porter's theory was only one side of the coin and the full theory of competitive advantage also rests on the inward-looking theories of the RBV, DCA, and core competency.

Regardless of the finer distinctions between the terms resource, core competency, and dynamic capability, these terms and their associated theories are all pointing to something valuable that is internal to the firm upon which strategy for building competitive advantage can be based. In our model, we use the term "capability", as encompassing the notions of resource, core competency, and dynamic capability as defined by the RBV, DCA, and core competency schools of thought, and this capability becomes the second dimension in our CURVO Grid shown in Figure 1. The CURVO Gird is a normative heuristic for competitive strategizing that identifies the types of option chains that fit best in certain quadrants of the two-dimensional uncertainty-capability space.

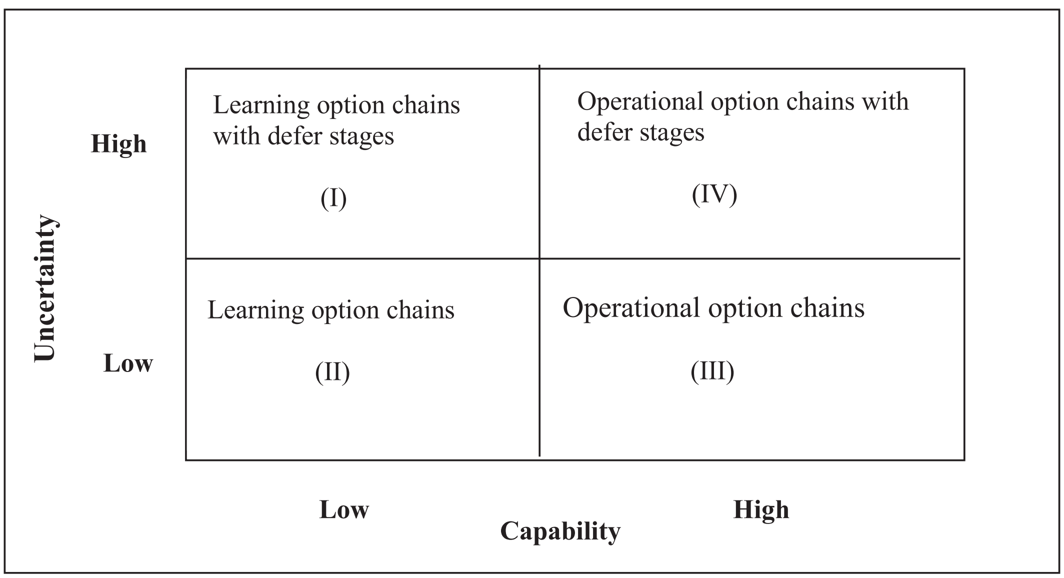


Figure 1: CURVO Repertory Grid

Does Management Always Exercise the Right Options Chain?

Recognizing the equivalence of strategy and option chains, the question in the section title could be put differently as: Does management always pursue the right strategy? Clearly not, otherwise we wouldn't have winners such as Amazon.com and losers such as Borders and Tower Records who were driven out of business by Amazon.com (Evangelista 2015). Misguided strategies abound, such as a firm doing little to explore growth opportunities. In other words, the firm is executing a defer option when it should be aggressively exercising a learning option chain followed by a strategic growth option chain. It is possible that the firm may have become too bogged down in grappling with operational issues in its existing business that it is neglecting growth opportunities. However, this has dire consequences for the survival of the firm in the long-term. The value of the CURVO Grid shown in Figure 1 is in providing management with guidance on the preferred option chains to execute under different conditions of uncertainty and firm capability. Importantly, the grid helps managers recognize when they could be going down the wrong path.

Operational option chains are a string of related actions, or exercises of options, targeted at improving the firm's existing product and service offerings. Since operational option chains are about the firm's existing business and offerings and not about a new market that the company does not know too much about just yet, operational option chains are exercised in Quadrants (III) and (IV) where the capability of the firm is high with respect to the target market. Capability includes unique resources such as technology patents and brand recognition supporting existing offerings; functional competencies such as in precision engineering and high-impact marketing; and dynamic managerial capabilities in perceiving changing market needs correctly and fashioning a response combining internal and external resources; and deploying these resources in a time-phased manner to best meet the changing conditions. Tending to the firm's existing business must be done whether the environmental uncertainty is high or not. Even in uncertain times, when the economy is in doldrums or when consumer preferences are fast changing, looking after the firm's existing business from which it derives its primary revenue must remain as one of the firm's key priorities.

A difference between the operational option chains exercised during times of high uncertainty and those in low uncertainty is that the operational option chain exercised in a highly uncertain environment will be laced with more defer stages (Quadrant III) as compared to the operational chain exercised in case of low environmental uncertainty. In times of high technology, market, or macroeconomic uncertainty it might better to temporarily suspend further action, or exercise the defer option within the option chain, till some information signal arrives that lifts the uncertainty. The signal could be about a technology standard that emerged as the winner in a battle of standards or it could be news that portends a better economy. The operational option chain could be resumed after this defer stage. Defer stages would be less necessary in times of low uncertainty, hence they are not shown in Quadrant (IV).

A firm is a collection of organizations at different levels and of different sizes ranging from ad hoc teams to units to functional organizations to full-scale business divisions. At any point in time, while some organizations, or business divisions, of the firm are tending to the firm's existing business, there must be other organizations that are exploring future growth opportunities. Often these growth opportunities enable the firm to get into a new market through technological innovation. The firm must then be engaged in learning about the new technology through internal R&D efforts and technology partnerships. The learning options chain is about gaining capability in new technologies for new markets that the firm could get into, and it is not about R&D investments to improve the firm's existing product and service offerings. Hence, it is more about discontinuous innovation rather than incremental innovation. The exercise of learning option chains is shown in Quadrants (I) and (II) which are associated with low existing capability. By definition, since the firming is learning about these new technologies and markets, it has a low current capability in this space. Again, just as tending to the existing business has to occur in an up-economy or a down-economy and in fast changing or slowly changing markets, exploring strategic growth opportunities has to be done in some place within the firm regardless of whether environmental uncertainty is high or low.

As in the case of the difference between Quadrants (III) and (IV), learning option chains in times of high uncertainty would be interspersed with more defer stages than in times of low uncertainty. It might be wise to temporarily suspend a learning option chain, or to exercise a defer option, and to wait for some information signal to arrive that lifts the uncertainty about, for example, the winner in a battle of technology standards.

A strategic growth opportunity that involves learning need not necessarily be based on new products and services enabled by discontinuous technological innovation. Entry into a major new market for the firm's existing offerings, such as a domestic firm diversifying internationally could also be viewed as a strategic growth opportunity. Again, a firm's existing capability in a foreign market is low and the firm must therefore learn about this new market via a learning options chain. The chain could begin with a joint venture with a foreign partner followed by a corporate acquisition of the partner.

Exercising Strategic Growth Option Chains

The purpose of learning is for eventually exercising strategic growth options or to execute a growth strategy. In other words, at the end of a learning options chain comes the opportunity to exercise a strategic growth option chain. Given that the CURVO Grid is a normative heuristic, the prescription it provides is that the strategic growth option chain is best executed in Quadrant (III) as shown in Figure 2.

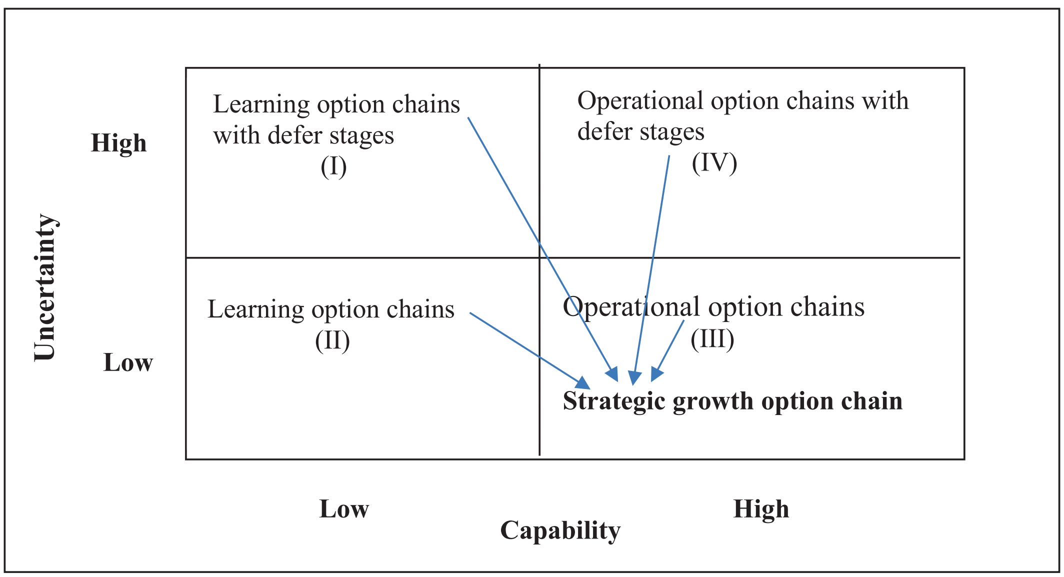


Figure 2: Strategic Growth Option Chain

To begin with, a key difference between exercising a strategic growth option chain versus an operational option chain is that growth options are about making big bets, such as leaping into a new market with innovative products undergirded by a corporate acquisition for new technology, whereas operational options are about making small bets such as enhancing service and support on an existing product. Making a big bet in a time of uncertainty involves a high risk of loss. The high risk of loss can be mitigated by having the growth strategy executed in a more stable environment, which means in Quadrant (III) and not in Quadrant (IV).

Furthermore, the growth strategy must utilize not only the knowledge gained from the learning activities performed in Quadrants (I) and (II) but, very importantly, must leverage the firm's capabilities in its existing business. It is this synergistic fusion of nascent capability in new technologies and markets acquired via the learning option chains with that of the firm's existing competencies that secures the firm's long-term competitive advantage and growth. If the firm's foray into new markets and opportunities was based solely on the learning resulting from the firm's exploration activities, then the firm's position in these new markets would be more open to competitive threat. Exploring new technologies from digital photography to artificial intelligence to business analytics is not the sole preserve of a single firm and many firms are simultaneously pursuing such new technologies. But the difference between a start-up firm introducing products based on new technologies and a firm with an established presence in a related business which is also introducing such products is the reuse of related technologies, possibly patented, and the application of its functional prowess in engineering, manufacturing, and sales and marketing from its existing to the new business. Of course, the watchword here is "related". What may appear to be related in the eyes of the senior management of a company, as the Nokia acquisition was to Steve Ballmer's Microsoft (Ribeiro 2013), may actually turn out to be not so related after all. In fact, a misperception of relatedness may turn out to be a strategic blunder as was demonstrated by the later divestment of Nokia by Microsoft (Warren 2016).

CURVO Grid and Ambidexterity Theory

While all firms have to look beyond their existing business if they are to survive in the future, which means they must simultaneously pursue exploration of future opportunities with exploitation of the current cash cow business, not all firms perform this juggling act successfully. The theory of ambidextrous organizations asserts that there is a particular type of organizational and management structure that allows for the aggressive pursuit and development of the next big opportunity along with the nurturing of the existing cash cow business (O'Reilly III and Tushman 2004). O'Reilly and Tushman (2004) claim that the pursuit of new opportunities must be housed in a separate organization with its own culture, values, structure, and business processes than the organization that tends to the firm's existing business. The exploratory and exploitative arms are by no means divorced from each other and there must be business synergy between the two. The ambidextrous organization has essentially mastered the trick of having both business synergy and organizational separation between the exploratory and exploitative arms of the firm.

O'Reilly and Tushman (2004) state that these two arms of the ambidextrous firm may indeed share resources, expertise, and customers while allowing for different cultures to flourish in these two arms. An example they provide of an ambidextrous firm is USA

Today (O'Reilly and Tushman 2004). USA Today has a traditional newspaper business, an online news service called USAToday.com, and a television operation called USAToday Direct, which are in different organizations with different structures and cultures. USAToday.com has a more entrepreneurial, fast-based, and collaborative culture as compared to that of the newspaper business. However, the pool of print reporters in the traditional newspaper business is the key resource that is being shared across the three organizations. The reporters generate the news stories that are then delivered in the three different media of print, Internet, and television. Hence, USA Today is a firm that is in the business of delivering news and it is essentially sharing "manufacturing" across its three divisions of USA Today, USAToday.com, and USAToday Direct, which are allowed to maintain their own cultures.

Operationalizing the CURVO Grid

To operationalize the CURVO Grid, the firm must first determine its own state in the two-dimensional capability-uncertainty space. The exploitative arm of the firm, by definition, is considered as having high capability because it focuses on the existing cash cow business. The exploitative arm is assumed to have worked with the firm's existing IT offerings for some years and successfully turned the existing business into a cash cow. The exploratory arm of the firm, on the other hand is, by definition, deemed to have low capability since it is just exploring new technologies and markets that the firm has not worked with before. So while the exploratory and exploitative arms of the firm can be located on the CURVO Grid with respect to the capability dimension, an instrument for measuring uncertainty needs to be developed.

<u>Uncertainty Measurement Model</u>

Our uncertainty instrument has items for assessing the economic, technology, market, and firm-specific uncertainties that the firm is facing. There are two versions of this instrument. One version is used by the organizations that are tending to the firm's existing business, or the exploitative arm of the firm, and the other by the organizations that are looking beyond the current business at future growth opportunities. The items in the uncertainty measuring instrument for the exploitative arm of the firm are as follows:

- Economy
    - The economic outlook is uncertain.
    - The economy is in recession or is expected to soon go into recession.
    - It is uncertain when we will see better economic times.
    - The economic uncertainties are causing firms to cut back on investments in IT.
- Market
    - The market outlook for our existing IT offerings is uncertain.
    - Consumer preferences may be shifting away from our IT existing offerings.
    - There are some signs that the market for our existing IT offerings could be shrinking.
    - Customer satisfaction with our existing IT offerings appear to be falling.
    - Our competitors are or will be soon introducing new and innovative IT offerings that could hurt our current market position.
    - There is an increasing threat of new players entering the market for our existing IT offerings with new and innovative offerings.

  - Our competitors are or could soon be introducing new IT offerings that are more attractively priced than our existing offerings.
- Technology
  - New technologies are making or could soon make our existing IT offerings obsolete.
  - Technology paradigms are changing fast in our markets and we could be at risk of falling behind.
  - The outlook is uncertain with regard to technology standards covering our existing IT offerings.
  - Our competitors appear to be stealing a march on us by introducing next generation technologies in their IT offerings.
  - Our competitors' IT offerings have or will soon have new features and functions that are not there in our existing offerings.
- Firm-Specific
  - The financial outlook for our company is uncertain.
  - A funding crunch may result in a reduction in financial support provided to our existing IT offerings.
  - The top management of our company is changing, and it is uncertain if the new management is committed to supporting our existing IT offerings.
  - Employee satisfaction in our firm is falling.
  - We are losing or could soon be losing key talent supporting our existing IT offerings because of employee satisfaction issues.
  - We are finding it difficult to hire the high-caliber talent we need to support our existing IT business.

The instrument for measuring uncertainty to be used by the exploratory arm of the firm has the following menu items.

- Economy
  - The economic outlook is uncertain.
  - The economy is in recession or is expected to soon go into recession.
  - It is uncertain when we will see better economic times.
  - The economic uncertainties are causing firms to cut back on investments in IT.
- Market
  - The market outlook for introducing new IT offerings based on our efforts in exploring new technologies and markets is uncertain.
  - Consumer preferences may already be shifting away from our projected IT offerings based on our efforts in exploring new technologies and markets.
  - There are some signs that the market for our projected IT offerings based on our efforts in exploring new technologies and markets could be shrinking.
  - Customer satisfaction with IT offerings similar to our projected offerings based on our efforts in exploring new technologies and markets appears to be falling.
  - Our competitors could introduce more innovative offerings compared to our projected IT offerings based on our efforts in exploring new technologies and markets

- There is an increasing threat of new players entering the market for our projected IT offerings, based on our efforts in exploring new technologies and markets, with even more innovative offerings.
- Our competitors are or could soon be introducing new IT offerings that are more attractively priced than our projected IT offerings based on our efforts in exploring new technologies and markets.

- Technology
    - New technologies are making or could soon make obsolete our projected IT offerings based on our efforts in exploring new technologies and markets.
    - Technology paradigms are changing fast in the markets for our projected IT offerings, based on our efforts in exploring new technologies markets, and there is a risk of our projected offerings' falling behind.
    - The outlook is uncertain with regard to technology standards covering our projected IT offerings based on our efforts in exploring new technologies and markets.
    - Our competitors may already be stealing a march on us by introducing next generation technologies in their IT offerings that are superior to what will be used in our projected IT offerings based on our efforts in exploring new technologies and markets.
    - Our competitors' IT next generation offerings will have features and functions that will not be there in our projected IT offerings based on our efforts in exploring new technologies and markets.
- Firm-Specific
    - The financial outlook for our company is uncertain.
    - A funding crunch may result in a reduction in financial support provided to our efforts to explore new technologies and markets.
    - The top management of our company is changing, and it is uncertain if the new management is committed to supporting our efforts to explore new technologies and markets.
    - Employee satisfaction in our firm is falling.
    - We are losing or could soon be losing key talent necessary for our efforts to explore new technologies and markets because of employee satisfaction issues.
    - We are finding it difficult to hire the high-caliber talent we need to support our efforts to explore new technologies and markets.

Each item in this instrument is measured on a 5-point Likert scale. Two management teams are formed, one from the exploitative arm of the company and the other from the exploratory arm of the company. Each management team must have both senior technical and business managers in it. For a managerial respondent, an average score is first found for each category of economic, market, technology, and firm-specific uncertainty. Then the average of these category averages taken. This average of category averages is the average level of uncertainty as perceived by that manager. The average levels of uncertainty perceived by the managers in the exploitative arm of the company are averaged to get a firm-level perception of uncertainty facing the firm's existing business. Similarly, this procedure is repeated for the management team assigned to assess the exploratory business to get a firm-level perception of uncertainty facing the exploratory arm of the firm. It is quite possible for the firm-level perception of uncertainty facing the firm's exploratory arm to be different from that facing its exploitative arm. It may well be that the firm's existing business may be in a stable market and is thus facing less uncertainty than its exploratory business. But this does not necessarily have to be true. Stability at one point in time does not mean stability forever, and it is quite possible that the firm's exploitative business could be going through a period of change.

Applying the CURVO Grid

The first step in applying the CURVO Grid is to ascertain how the firm is positioned in the grid. At a minimum the firm must be dual positioned in the grid with some part of the firm positioned in either Quadrant I or Quadrant II and another part positioned in either Quadrant III or IV. In other words, regardless of whether the firm is ambidextrous or not, it must be simultaneously tending to its existing business, and also exploring future growth opportunities as shown in Figure 3. If the firm does not have a position in Quadrant I or II along with a position in Quadrant III or IV, then that is the first indication of a poor strategy. The firm could be too bogged down in grappling with issues in its existing business that it is neglecting exploring future growth opportunities. This does not augur well for the long-term survival of the firm.

Once the firm has determined its positioning in the CURVO Grid, it must then check to see if the options chain that it is executing is the correct one for that cell. The dynamic managerial capability that Teece et al. (1997) propound, where managers can deploy internal and external resources in the right way and at the right time, is recast in this article as exercising the right options chain. An important contribution of options-based thinking in approaching investment decision-making is to consider the defer option (Dixit and Pindyck 1995). In other words, instead of the now-or-never approach of the NPV style of capital management, sometimes it is simply better to wait for more information to arrive to alleviate some of the uncertainty than to take an action now. Exercising the defer option as an alternative to taking some action now applies to the case of a single investment decision being contemplated at time $t = 0$. With regard to an option chain, where many related options are to be exercised in a certain sequential pattern, the equivalent of the defer option is a defer stage. Hence, the exploratory arm of the firm must include defer stages in the learning options chain when uncertainty as indicated by the uncertainty measurement instrument is high. Similarly, the exploitative arm of the firm must include defer stages in the operational option chain when the uncertainty measure is high.

The other key managerial prescription offered by the CURVO Grid is that, while exploring new technologies and markets is a must, learning cannot go on indefinitely. At some point, the learning option chain must be terminated, and then a growth option chain initiated. There are two salient aspects of this conversion. First it is best to convert the learning option chain into a growth option in a time of relative stability. The learning option chain was about making small bets whereas the growth option chain is about making large bets. Hence, the initiation of the growth strategy is best done in a time of relative calm as opposed to that of turmoil and change. Furthermore, the execution of the strategic growth option chain is not done in Quadrant I or II where the capability of the firm with respect to the target markets and technology is low. It is done when the capability is high. One reason for this is that the growth option chain is initiated when the learning option chain has run its course and is terminated. Hence, the firm should have acquired some capability in the target technology and markets through the course of the learning option chain. But, perhaps importantly, there must be synergy and integration between the new opportunity that the firm is getting into and the firm's existing business. The leveraging and reuse of capability from the firm's existing business is an essential part of the growth strategy. The fusion of capability acquired through learning with relevant portions of the firm's capability in its existing business endows the firm with the high capability to execute its growth strategy.

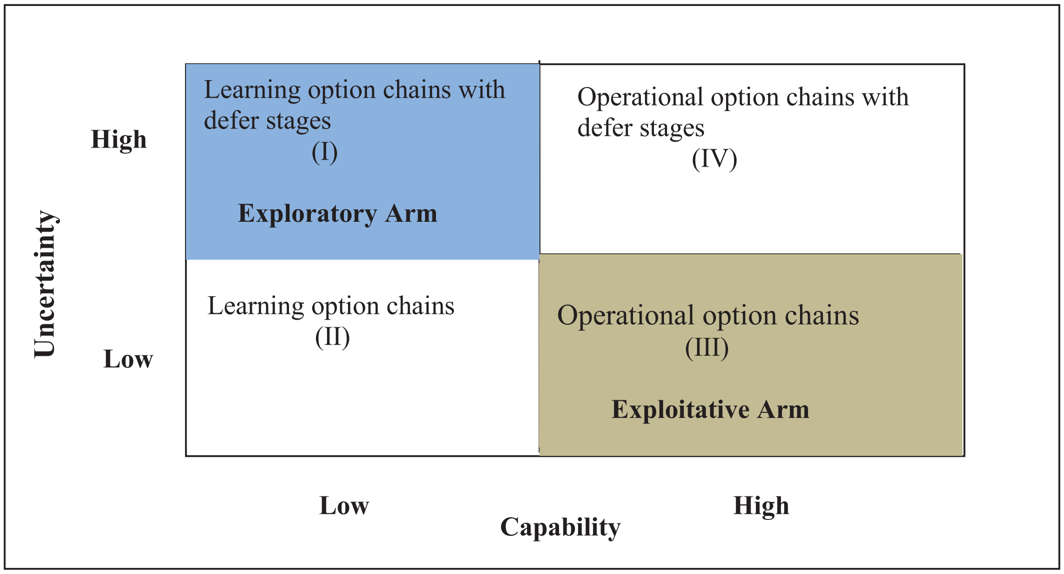


Figure 3 – Multi-Cell Positioning of Firm

CONCLUSION

This paper develops an innovative soft heuristic model, the CURVO Grid, for shaping an IT firm's competitive strategy. The CURVO Grid is informed by all the major theories and perspectives of competitive advantage including Porter's competitive forces model (Porter 1980, 1985), Prahalad and Hamel's core competency theory (Prahalad and Hamel 1990), Barney's resource-based view (Barney 1991, 2001), Teece et al.'s dynamic capabilities approach (Teece et al. 1997), and O'Reilly and Tushman's ambidexterity theory (O'Reilly and Tushman 2004). The article defines strategy in terms of exercising options chains. These include learning option chains exercised by the exploratory arm of the firm; operational option chains exercised by the exploitative arm of the firm; and strategic growth option chains which are triggered upon the culmination of learning option chains. Clearly, not all firms get competitive strategy right. The CURVO Grid provides guidelines in terms of which option chains to execute under what conditions of firm capability and environmental uncertainty.

The CURVO Grid is also a qualitative framework that brings the logic of real options into competitive strategizing without having to work through complex, quantitative options pricing models. As has been noted in RO literature, quantitative OPMs are often inexplicable and hence opaque to high-level managers and strategists (Baker et al. 2015, Teach 2003). Just as the SWOT matrix brought the structure-conduct-performance theories of industrial organization (Mason 1949, Bain 1959, Schmalensee

1989, Schmalensee and Willig 1989) within easy reach of the practitioner, the CURVO Grid does something similar by making RO accessible to practitioners.

REFERENCES

AACSB International (2008) *Impact of Research Task Force Report*, http://www.aacsb.edu/resource_centers/Research/Final/Impact_of_Research_Report-FINAL.PDF (current December 14, 2008).

Abrahamson, E. (1996) "Management fashion," *Academy of Management Review* 21(1), pp. 254-285.
Abrahamson, E., Rosenkopf, L. (1997) "Social network effects on the extent of innovation diffusion: A computer simulation," Organization Science 8(3), pp. 289-309.

Amram, M., Kulatilaka, N. (1998) *Real Options – Managing Strategic Investment in an Uncertain World*. Cambridge, MA: Harvard Business School Press.

Bain, J.S. (1959) *Industrial Organization*. New York, NY: Wiley.

Baker, H.K., Dutta, S., Saadi, S. (2015) "Management views on real options in capital budgeting," Journal of Applied Finance 21(1), pp. 18-29.

Bardhan, I., Bagchi, S., Sougstad, R. (2004) "Prioritizing a portfolio of information technology investment projects," *Journal of Management Information Systems* 21(2), pp. 33-70.

Barney, J.B. (1991) "Firm resources and sustained competitive advantage," *Journal of Management* 17, pp. 99-120.

Barney, J.B. (2001) "Is the Resource-Based "View" a useful perspective for strategic management research? Yes," *The Academy of Management Review* 26(1), pp. 41-56.

Benaroch, M. (2002) "Managing information technology investment risk: A real options perspective," *Journal of Management Information Systems* 19(2), pp. 43-84.

Benaroch, M., Kauffman, R.J. (2000) "Justifying electronic banking network expansion using real options analysis," *MIS Quarterly* 24(2), pp. 197-225.

Benaroch, M., Shah, S., Jeffery, M. (2006a) "On the valuation of multi-stage information technology investments embedding nested real options," *Journal of Management Information Systems* 23(1), pp. 239-261.

Benaroch, M., Lichenstein, Y., Robinson, K. (2006b) "Real options in information technology risk management," *MIS Quarterly* 30(4), pp. 827-864.

Benaroch, M., Jeffery, M., Kauffman, R.J., Shah, S. (2007) "Option-based risk management: A field study of sequential information technology investment decisions," *Journal of Management Information Systems* 24(2), 103-140.

Benbasat, I., Zmud, R.W. (1999) "Empirical research in information systems: The practice of relevance," *MIS Quarterly* 23(1), pp. 3-16.

Black, F., Scholes, M. (1973) "The pricing of options and corporate liabilities," *Journal of Political Economy* 81(3), 637-654.

Bowman, E.H., Hurry, D. (1993) "Strategy through the option lens: An integrated view of resource investments and the incremental-choice process," *The Academy of Management Review*, 18(4), pp. 760-782.

Brealey, R.A., Allen, F., Myers, S.C. (2016) *Principles of Corporate Finance* 12th edition. New York City, NY: McGraw-Hill Education.

Cassimon, D., Engelen P.J., Thomassen, L., Van Wouwe, M. (2004) "The valuation of a NDA using a 6-fold compound option," *Research Policy* 33, pp. 41-51.

Cohen, W.M., Levinthal, D.A. (1990) "Absorptive capacity: A new perspective on learning and innovation," *Administrative Science Quarterly* 35(1), pp. 128-152.

Collins, T. (2018) "Incredible new device that generates electricity from thin air by harvesting energy from changes in temperature could power gadgets for years," DailyMail.com https://www.dailymail.co.uk/sciencetech/article-5440613/Incredible-device-generates-electricity-air.html (current May 13, 2019)

Copeland, T.E., Weston, J.F., Shastri, K. (2005) *Financial Theory and Corporate Policy* 4th edition. Upper Saddle Hill River, NJ: Pearson Education.

DeSanctis, G., Poole, M.S. (1994) "Capturing the complexity in advanced technology use: Adaptive structuration theory," *Organization Science* 5(2), 121-147.

Dixit, A.K., and Pindyck, R.S. (1995) "The options approach to capital investment," *Harvard Business Review* 73(3), pp. 105-115.

Dos Santos, B. (1991) "Justifying investment in new information technologies," *Journal of Management Information Systems* 7(4), pp. 71-89.

Eden, C. (2004) "Analyzing cognitive maps to help structure issues or problems," *European Journal of Operational Research* 159, 673-686.

Eden, C., Jones, S. (1984) "Using repertory grids for problem construction," *Journal of the Operational Research Society* 35(9), 779-790.

Erl, T. (2005) *Service-Oriented Architecture (SOA): Concepts, Technology, and Design*. Upper Saddle Hill River, NJ: Prentice Hall/Pearson Education.

Evangelista, B. (2015) "How Amazon factor killed retailers like Borders, Circuit City," SFGate https://www.sfgate.com/business/article/How-Amazon-factor-killed-retailers-like-6378619.php (current, Jan 20, 2016)

Feinstein, S.P., Lander, D.M. (2002) "A better understanding of why NPV undervalues managerial flexibility," *The Engineering Economist* 47(4), pp. 418-35.

Fichman, R.G. (2004a) “Going beyond the dominant paradigm for information technology innovation research: Emerging concepts and methods,” *Journal of the Association for Information Systems* 5(8), pp. 314-355.

Fichman, R.G. (2004b) “Real options and IT platform adoption: Implications for theory and practice,” *Information Systems Research* 15(2), pp. 132-154.

Fiorano ESB (2009) http://www.fiorano.com/products/fesb/products_fioranoesb.php (current April 15, 2009)

Gamba, A., Fusari, N. (2009) “Valuing modularity as a real option,” *Management Science* 55(11), 1877-1896.

Ghosh, S., and Troutt, M.D. (2012) “Complex compound option models – Can practitioners truly operationalize them?” *European Journal of Operational Research* 222(3), pp. 542-552.

Ghosh, S., Troutt, M.D., Brandyberry, A.A. (2013) “An RBV and real options-based soft OR heuristic for solution strategizing.” *International Journal of Business Information Systems*, 12(4), pp. 404-422.

Gopinath, C., Hoffman, R.C. (1995) “The relevance of strategy research: Practitioner and academic viewpoints,” *Journal of Management Studies* 32(5), pp. 575-594.

Graham, J.R., Harvey, C.R. (2001) “The theory and practice of corporate finance: evidence from the field,” *Journal of Financial Economics* 60(2-3), pp. 187-243.

Gross, G. (2003) “VoIP providers: Lay off government regulation,” *Computerworld* https://www.computerworld.com/article/2573376/voip-providers--lay-off-government-regulation.html (current May 13, 2019)

Jensen, K., and P. Warren (2001) “The use of options theory to value research in the service sector,” *R&D Management* 31(2), pp. 173-180.

Karsak, E.E., Özogul, C.O. (2002) “An options approach to valuing expansion flexibility in flexible manufacturing system investments,” The Engineering Economist, 47(2), pp. 169-193.

Keil, M., Cule, P.E., Lyytinen, K., and Schmidt, R.C. (1998) “A framework for identifying software project risks,” *Communications of the ACM* 41(11), pp. 76-83.

Kiwi.com (2019) https://www.kiwi.com/us/ (current May 13, 2019)

Kogut, B., Zander, U. (1992) “Knowledge of the firm, combinative capabilities, and the replication of technology,” Organization Science 3(3), pp. 383-397.

Krafzig, D., Banke, K., Slama, D. (2005) *Enterprise SOA – Service Oriented Architecture Best Practices*. Upper Saddle Hill River, NJ: Pearson.

Kua, H.W. (2016) “A new integrated framework for stakeholder involvement in sustainability policymaking – a multidisciplinary approach," *Sustainable Development* 24(5), pp. 281-297.

Kulatilaka, N. (1988) "Valuing the flexibility of flexible manufacturing systems," *IEEE Transactions on Engineering Management* 35(4), pp. 250-257.

Kumar, R.L. (1995) "An options view of investments in expansion-flexible manufacturing Systems," *International Journal of Production Economics* 38, 281-291.

Luehrman, T.A. (1998) "Strategy as a portfolio of real options," *Harvard Business Review* September-October, pp. 89-99.

Mason, E. (1949) "The current state of the monopoly problem in the U.S.," *Harvard Law Review* 62, pp. 1265-1285.

Miller, L.T., Park, C.S. (2002) "Decision-making under uncertainty – real options to the rescue?" *The Engineering Economist* 45(1), pp. 105-150.

O'Reilly III, C.A., Tushman, M.L. (2004) "The ambidextrous organization," Harvard Business Review 82(4), pp. 74-81.

Park, C.S., Herath, H.S.B. (2000) "Exploiting uncertainty – investment opportunities as real options: A new way of thinking in engineering economics," *The Engineering Economist* 45(1), pp. 1-35.

Porter, M.E. (1980) *Competitive Strategy*. New York, NY: The Free Press.

Porter, M.E. (1985) *Competitive Advantage, Creating and Sustaining Superior Performance*. New York, NY: The Free Press.

Prahalad, C.K., Hamel, G. (1990) "The core competence of the corporation," *Harvard Business Review* (May-June), pp. 79-91.

Progress Software Sonic ESB (2009) http://www.sonicsoftware.com/products/sonic_esb/index.ssp (current February 14, 2009)

Ribeiro, J. (2013) "Update: Microsoft to acquire Nokia's mobile phone business," ComputerWorld https://www.computerworld.com/article/2484423/smartphones-update-microsoft-to-acquire-nokia-s-mobile-phone-business.html (current May 24, 2019)

Schmalensee, R. (1989) "Inter-industry studies of structure and performance," in, R. Schmalensee, and R. D. Willig, (eds.) *Handbook of Industrial Organization*. New York, NY: North Holland.

Schmalensee, R., Willig, R.D. (1989) *Handbook of Industrial Organization*. New York, NY: North Holland.

Schumpeter, J.A. (1934) *Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Taudes, A. (1998) “Software growth options,” *Journal of Management Information Systems* 15(1), pp.165-185.

Teach, E. (2003) “Will real options take root?” *CFO* 19, pp. 73-75.

Teece, D.J., Pisano, G., Shuen, A. (1997) “Dynamic capabilities and strategic management,” *Strategic Management Journal* 18(7), pp. 509-533.

TelcoWorld (2019) “Seven new smartphone features for 2019,” https://www.telcoworld.com.au/news-blog/seven-new-smartphone-features-for-2019/ (current May 13, 2019)

Triantis, A.J., Hodder., J.E. (1990) “Valuing flexibility as a complex option,” *Journal of Finance* 45(2), pp. 549-565.

Tyre, M.J., Orlikowski, W.J. (1994) “Windows of opportunity: Temporal patterns of technological adaptation in organizations,” *Organization Science* 5(1), pp. 98-118.

Warren, T. (2016) “Microsoft wasted at least $8 billion on its failed Nokia experiment,” THE VERGE https://www.theverge.com/2016/5/25/11766540/microsoft-nokia-acquisition-costs (current May 24, 2019)

Wernerfelt, B. (1984) “A resource-based view of the firm,” *Strategic Management Journal* 5(2), pp. 171-180.

Wernerfelt, B. (1989) “From critical resources to corporate strategy,” *Journal of General Management* 14, pp. 4-12.

Wernerfelt, B. (1995) “The resource-based view of the firm: Ten years after,” *Strategic Management Journal* 16, pp. 171-174.

DECISION SCIENCES INSTITUTE

Mortality Caused by Malignant Neoplasms: Its Impact on Genders and Ethnicities

Jesus A. Cardenas
University of Texas at El Paso
Email: jacardenas3@utep.edu

Leopoldo Gemoets
University of Texas at El Paso
Email: lgemoets@utep.edu

ABSTRACT

The mortality rate in the United States due to Cancers is increasing in general but some malignant neoplasm deaths are showing a trend down. This study correlates mortality trends to gender and ethnicity as a way of identifying trends which could be used by policy makers and clinicians to allocate resources appropriately for the treatment of cancers.

KEYWORDS: regression, mortality, ethnicities, sensitivity analysis and health

INTRODUCTION

The United States population is steadily growing, but there is an element that has a major impact on this growth: migration. With the increase of nuclear families for blacks and whites, the overall population is showing signs of decreasing year per year, but there is an ethnicity that is growing continuously and having a smaller death ratio, so we will analyze this part of the population, and we will compare it to the other components to identify possibilities of reducing mortality due to cancers in the US population.

Table 1: USA Population by Race and Gender in 2017

Male		Group	Female	
29.67%	95,842,609	White	98,885,235	30.61%
9.31%	30,058,779	Hispanic	30,094,010	9.32%
6.17%	19,922,595	Black	22,599,391	7.00%
3.78%	12,212,462	Others	13,381,513	4.14%

Interestingly enough, the Hispanic population occupies now the second place on regards to demographics as they represent 18.63% of the total population in the country. The white population is still a strong majority as they constitute 60.28% of the total population. Another interesting fact is that there are more women than men in all ethnicities and races, but the ratio is much larger for black women as they are 6.3% more black women in all the black population, while the ratio for white women is 1.56% and for Hispanics is only 0.06%.

GENDER MORTALITY CAUSED BY CANCER

In this study we are going to focus on mortality caused by all the different types of malignant neoplasms (cancers) in the United States, which along with cardiovascular diseases are the top two causes of death in the world. With the certainty of a R^2 of 98.67% we see the trend of total mortality caused by cancer growing with a slope of 4,358.2 deaths per year.

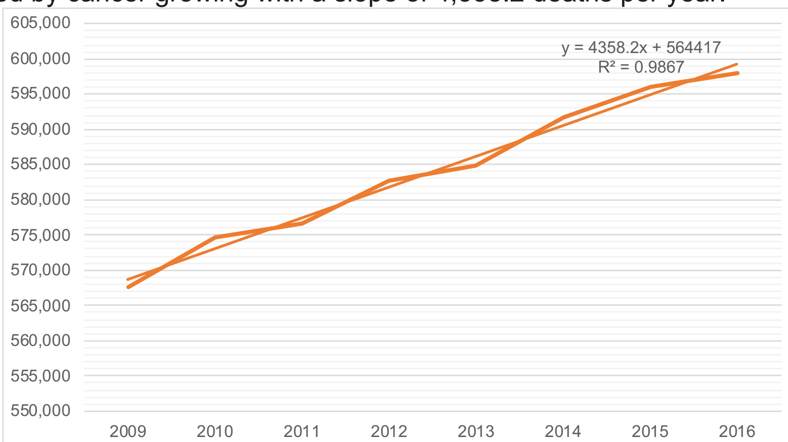


Figure 1. Total Mortality Caused by Malignant Neoplasms in the US

As we separate the mortality chart by the contribution of male's cancers, we see that the total growth of deaths by cancer is highly contributed by males. With a slope of 2,591 and a R^2 of 98.49%, males' deaths represent 59.31% of the yearly increase.

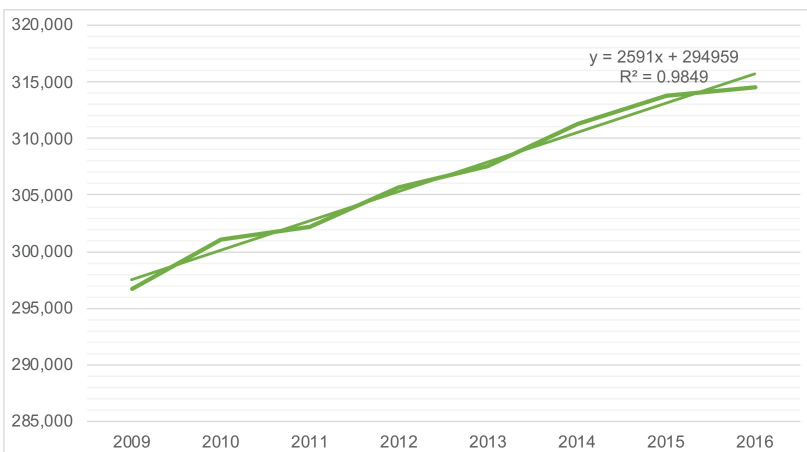


Figure 2. Males Mortality Caused by Malignant Neoplasms in the US

The female contribution to the yearly increase is only 1,767.2 deaths by cancer per year. Considering that women make more than half of the population, we are seeing a constant growth trend, with a much lower slope than men and with an intersect 25,500 already lower. In other words, the number of males' deaths due to cancer will continue increasing the gap with women in the future.

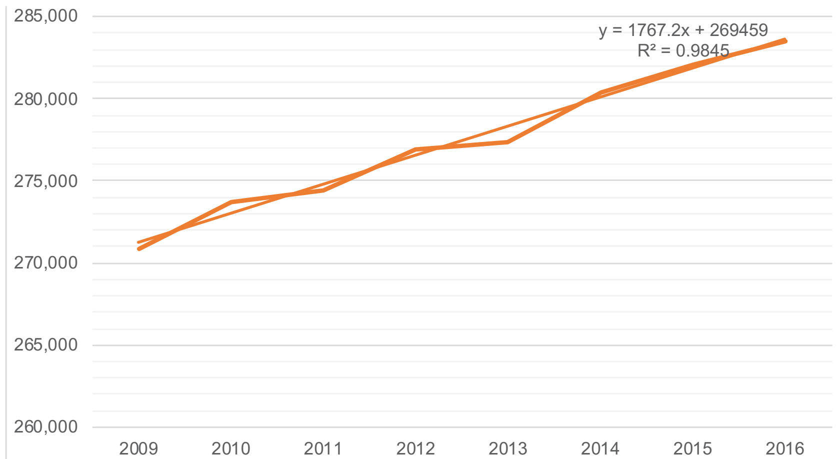


Figure 3: Female Mortality Caused by Malignant Neoplasms in the US

METHODOLOGY

Looking at these charts we can determine that cancer is causing more males' deaths in the Uniterd States, so that gender indeed has an impact on the mortality number and this trend seems to continue, so now we are going to evaluate ethnicity and race to determine if there is also a difference among them. In this project we are going to use linear regression to measure the correlation between two variables that are going to be changing from gender, ethnicity and type of malignant neoplasms; and we will also do sensitivity analyses to identify ailing points.

ETHNICITY/RACE MORTALITY CAUSED BY CANCER

In this study we are going to focus only on three major groups: White, Black and Hispanic where the White and Black groups do not include any Hispanics. The largest group and the first one to analyze is the White Non-Hispanic group. With the certainty of a R^2 of 91.79% we see the trend of White population mortality caused by cancer growing with a slope of 1,395.7 deaths per year.

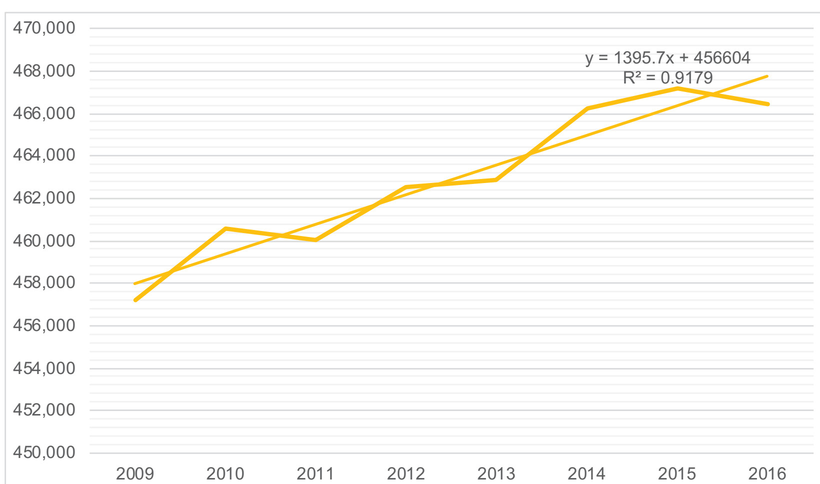


Figure 4. White Population Mortality Caused by Malignant Neoplasms in the US

With a y-intersect of 456,604 deaths, this group is definitively the major contributor of cancer's deaths, but the slope is growing only 1,395.7, that is 33.27% of the total deaths per year. Because the white population represents 60.28% of the total population, the cancer's deaths are only one third, so we can conclude that cancer deaths in the white group is being controlled and the concern about growth of mortality comes from the other groups.

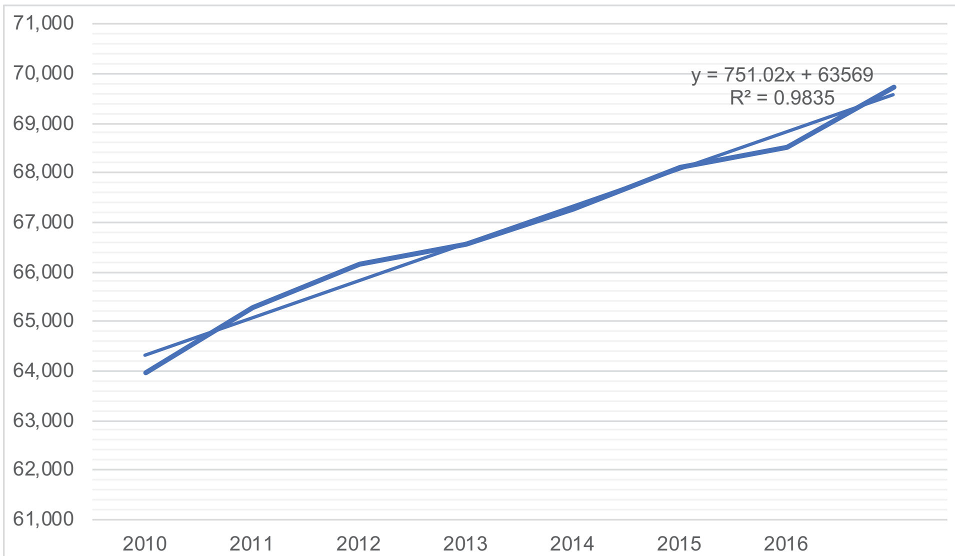


Figure 5. Black Population Mortality Caused by Malignant Neoplasms in the US

The contribution of mortality of the back population is only 751 deaths per year, that is 17.23% while the proportion of the population is 13.17%. The numbers are close, but the ratio of death is higher, so we have to be careful about the different types of cancers as we analyze them.

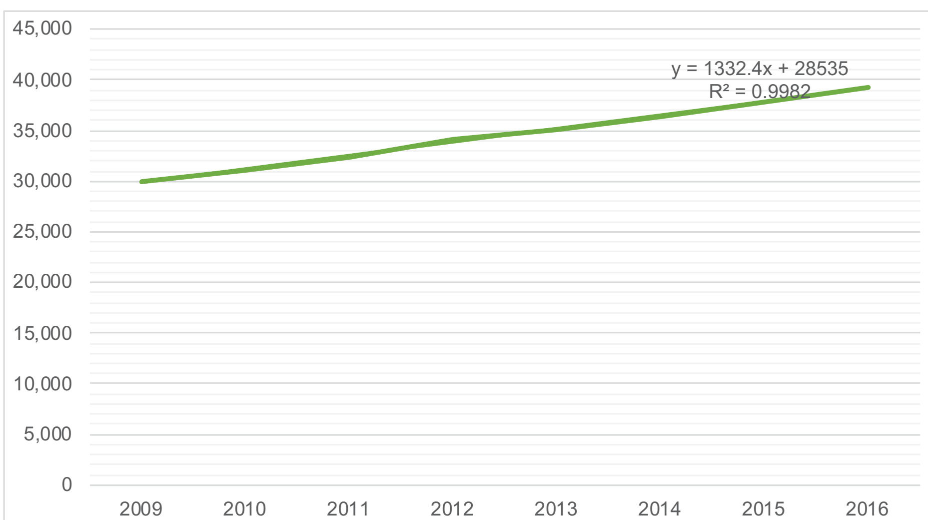


Figure 6: Hispanic Population Mortality Caused by Malignant Neoplasms in the US

We are discovering that 30.57% of the malignant neoplasms deaths every year are Hispanics while the proportion of the population is only 18.63%. In order to determine how critical this growth is, we will probe deeper into the 8 major causes of cancer's mortality and considering the contribution of gender and ethnicity.

Ethnicity	Gender	Xo	Slope	R^2
White	Male	238,870.00	1,183.00	0.9446
White	Female	217,734.00	212.71	0.6055
Black	Male	32,860.00	285.68	0.9835
Black	Female	30,709.00	465.35	0.9654
Hispanic	Male	15,117.00	669.00	0.9986
Hispanic	Female	13,418.00	663.37	0.9934

Table 2. Cancer Mortality lines for Gender and Ethnicity

Table # 2 shows very interesting facts about the cancer mortality. The Xo value for white people is around 7 times higher than the numbers for black people and around 16 times higher than Hispanics, but the yearly number of incidences (slope) for Hispanics and black women is alarming . The only two groups that are showing minimum growth are white women and black men. We will proceed to investigate the different types of cancer provoking death and break doen into these categories to find the possible cause that could drive medical efforts to find a way to reverse the trends.

TYPES OF MALIGNANT NEOPLASMS RESULTING IN DEATHS IN THE USA

The study is based on the available information at the Centers for Disease Control and Prevention (CDC). We are looking at the period between 2009 and until 2016 because these are the tables that have the most complete information.

The slopes show that there has been an important improvement on the treatment of trachea cancers for both genders, as we see decrements on mortality slopes for both. Table # 3 shows that the main contributor for the growth of deaths in males is mostly due to liver, pancreas, prostrate and bladder. On the female side, the major contributors are pancreas and liver. Comparing the contribution of the top ten causes we can quantify 2,329 deaths per year, that is 53% of all cancer related deaths.

Typically, for female's cancers, ovary and breast are the main contributors, but ovary cancer is showing a consistent trend down with a Xo=14,539 and a slope of -45.44. Breast cancer results can be seen in the table 3, and they clearly show that this type of cancer is growing minimally with 105.54 cases per year, but with an intersection of 40,630. Because we are seeing that the main issue on regards to mortality is related to liver, pancreas for both genders and prostate and bladder cancer for males only.

Cancer	Gender	Xo	Slope	R^2
Trachea	Male	89,550.00	(896.48)	0.8704
Trachea	Female	71,105.00	(204.13)	0.339
Colon	Male	26,762.00	126.88	0.7217
Colon	Female	25,332.00	(30.55)	0.1235
Breast	Male	405.75	8.25	0.5633
Breast	Female	40,630.00	105.54	0.7543
Pancreas	Male	17,343.00	564.60	0.9853
Pancreas	Female	17,268.00	423.05	0.9804
Prostrate	Male	27,155.00	315.40	0.435
Prostrate	Female	-	-	0
Liver	Male	12,463.00	704.33	0.9851
Liver	Female	5,977.90	344.81	0.9984
Leukemia	Male	12,961.00	63.00	0.3993
Leukemia	Female	9,687.40	30.66	0.4683
Lymphoma	Male	11,119.00	24.67	0.4162
Lymphoma	Female	9,240.00	(36.88)	0.5812
Bladder	Male	9,801.40	277.99	0.9713
Bladder	Female	4,121.10	73.01	0.9059
Meninges	Male	7,481.00	258.71	0.9391
Meninges	Female	5,950.90	176.36	0.937

Table 3. Cancer Mortality lines for Gender and Ethnicity

Cancer	Ethnicity	Gender	Xo	Slope	R^2
Pancreas	White	Male	14,193.00	335.33	0.9062
Pancreas	Black	Male	1,909.10	57.524	0.8944
Pancreas	Hispanic	Male	1,008.30	46.845	0.8442
Liver	White	Male	8,487.30	391.05	0.8965
Liver	Black	Male	1,848.40	83.507	0.8769
Liver	Hispanic	Male	1,484.60	81.607	0.876
Prostrate	White	Male	20,951.00	105.39	0.1156
Prostrate	Black	Male	4,566.10	33.512	0.1469
Prostrate	Hispanic	Male	1,353.50	69.714	0.9086
Bladder	White	Male	8,752.90	228.71	0.9553
Bladder	Black	Male	562.57	21.179	0.9182
Bladder	Hispanic	Male	325.46	18.536	0.8878

Table 4: Males' Cancer Mortality by Type and Ethnicity

Table # 4 shows that the white male population is the one with the largest Xo because of the high number of incidences. The largest slopes of the selected male cancers are also from the white population, because of the largest proportion of population, but we have to be careful about the Hispanic population death growth which is smaller than the white population, but percentagewise they are higher.

Cancer	Ethnicity	Gender	Xo	Slope	R^2
Pancreas	White	Female	13,883.00	184.15	0.9262
Pancreas	Black	Female	2,157.70	55.536	0.7496
Pancreas	Hispanic	Female	987.07	49.179	0.8466
Liver	White	Female	4,235.30	172.21	0.9172
Liver	Black	Female	767.54	39.298	0.9362
Liver	Hispanic	Female	688.46	51.119	0.8598
Breast	White	Female	81,621.00	-110	0.8484
Breast	Black	Female	5,849.10	51.52	0.732
Breast	Hispanic	Female	2,140.50	100.21	0.9309
Bladder	White	Female	3,483.20	44.571	0.8585
Bladder	Black	Female	420.61	13.463	0.7911
Bladder	Hispanic	Female	151.36	7.8095	0.8857
Ovary	White	Female	12,098.00	-111.4	0.9442
Ovary	Black	Female	1,248.60	10.357	0.3289
Ovary	Hispanic	Female	804.46	26.952	0.8747

Table 5. Female Cancer Mortality by Type and Ethnicity

SENSITIVITY ANALYSIS OF THE CANCERS' MORTALITY

The sensitivity analysis compares the cell versus the summation of the column and then versus the summation of the rows, this way we can determine if the contribution of the events is changing in time. In the particular case of the white males' mortality (table # 6), we can notice that the columns of the liver, meninges and bladder changed from green to red, that is, their contribution in time has been growing. Trachea, bronchus and lung's cancer is showing a positive trend as it changed from red to green.

Year	Trachea, bronchus and lung (C33-34)	Prostate (C61)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Leukemia (C91-95)	Liver and intrahepatic bile ducts (C22)	Bladder (C67)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Breast (C50)	All Others	Grand Total
2009	13.0%	12.4%	12.6%	11.4%	12.4%	10.6%	11.5%	12.6%	11.5%	12.0%	11.9%	12.2%
2010	12.9%	12.6%	12.6%	12.0%	12.3%	11.1%	11.9%	12.4%	11.6%	12.2%	12.1%	12.4%
2011	12.7%	12.3%	12.4%	12.0%	12.5%	11.8%	12.1%	12.7%	11.9%	12.7%	12.3%	12.4%
2012	12.7%	11.9%	12.4%	12.5%	12.7%	12.6%	12.3%	12.6%	12.5%	12.0%	12.4%	12.5%
2013	12.5%	12.1%	12.4%	12.5%	12.8%	13.1%	12.7%	12.4%	12.3%	12.6%	12.6%	12.5%
2014	12.4%	12.4%	12.4%	13.1%	12.7%	13.4%	12.7%	12.4%	13.0%	12.8%	12.8%	12.6%
2015	12.1%	13.1%	12.5%	13.4%	12.4%	14.1%	13.4%	12.4%	13.6%	12.5%	12.6%	12.7%
2016	11.7%	13.1%	12.5%	13.1%	12.4%	13.4%	13.4%	12.4%	13.6%	13.2%	13.2%	12.7%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Year	Trachea, bronchus and lung (C33-34)	Prostate (C61)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Leukemia (C91-95)	Liver and intrahepatic bile ducts (C22)	Bladder (C67)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Breast (C50)	All Others	Grand Total
2009	30.5%	8.9%	8.8%	6.0%	4.5%	3.6%	3.7%	3.9%	2.8%	0.1%	27.1%	100%
2010	30.0%	9.0%	8.7%	6.2%	4.4%	3.8%	3.8%	3.8%	2.8%	0.1%	27.4%	100%
2011	29.6%	8.7%	8.5%	6.2%	4.5%	4.0%	3.9%	3.9%	2.9%	0.1%	27.6%	100%
2012	29.2%	8.4%	8.4%	6.5%	4.5%	4.2%	4.0%	3.8%	3.0%	0.1%	27.8%	100%
2013	28.7%	8.5%	8.5%	6.4%	4.6%	4.4%	4.1%	3.8%	2.9%	0.1%	28.0%	100%
2014	28.1%	8.7%	8.3%	6.7%	4.5%	4.5%	4.0%	3.7%	3.1%	0.1%	28.3%	100%
2015	27.5%	9.1%	8.4%	6.8%	4.4%	4.7%	4.2%	3.7%	3.2%	0.1%	27.8%	100%
2016	26.6%	9.2%	8.4%	6.7%	4.4%	4.4%	4.2%	3.7%	3.2%	0.1%	29.0%	100%
Total	28.8%	8.8%	8.5%	6.4%	4.5%	4.2%	4.0%	3.8%	3.0%	0.1%	27.9%	100%

Table 6. Sensitivity Analysis for White Males' Mortality by Cancer

Doing the same analysis in the table #7 for the white females, we are finding red lights on the liver, meninges, bladder and pancreas as they change from green to red color code. Breast and colon cancers are showing a slim reduction on the past eight years, so whatever is being done, seems to be working.

Year	Trachea, bronchus and lung (C33-34)	Breast (C50)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Liver and intrahepatic bile ducts (C22)	Bladder (C67)	All Others	Grand Total	Trachea, bronchus and lung (C33-34)	Breast (C50)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Liver and intrahepatic bile ducts (C22)	Bladder (C67)	All Others	Grand Total
2009	12.7%	12.7%	12.8%	11.9%	12.4%	13.0%	11.9%	10.8%	11.9%	12.3%	12.5%	27.7%	14.5%	9.1%	6.4%	3.6%	3.6%	2.5%	2.0%	1.6%	29.1%	100%
2010	12.7%	12.7%	12.7%	12.1%	12.5%	12.9%	11.6%	11.4%	12.2%	12.4%	12.5%	27.5%	14.4%	9.0%	6.5%	3.6%	3.5%	2.4%	2.1%	1.6%	29.4%	100%
2011	12.6%	12.5%	12.6%	12.2%	12.7%	12.5%	12.0%	11.8%	12.4%	12.4%	12.5%	27.3%	14.3%	8.9%	6.6%	3.7%	3.4%	2.5%	2.2%	1.7%	29.3%	100%
2012	12.6%	12.5%	12.5%	12.5%	12.4%	12.6%	12.3%	12.4%	12.4%	12.4%	12.5%	27.3%	14.3%	8.8%	6.7%	3.6%	3.4%	2.5%	2.3%	1.7%	29.3%	100%
2013	12.5%	12.4%	12.4%	12.5%	12.5%	12.3%	12.8%	12.9%	12.6%	12.4%	12.5%	27.2%	14.2%	8.8%	6.7%	3.6%	3.4%	2.6%	2.4%	1.7%	29.4%	100%
2014	12.5%	12.4%	12.4%	12.9%	12.7%	12.4%	12.7%	13.4%	12.4%	12.5%	12.5%	27.0%	14.1%	8.7%	6.9%	3.7%	3.4%	2.6%	2.4%	1.7%	29.5%	100%
2015	12.4%	12.4%	12.4%	13.0%	12.3%	12.2%	13.4%	13.9%	13.1%	12.6%	12.5%	26.8%	14.1%	8.7%	7.0%	3.6%	3.3%	2.7%	2.5%	1.8%	29.5%	100%
2016	12.0%	12.4%	12.2%	12.9%	12.3%	12.2%	13.4%	13.4%	13.1%	13.0%	12.5%	25.9%	14.0%	8.6%	6.9%	3.6%	3.3%	2.7%	2.4%	1.8%	30.7%	100%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	27.1%	14.2%	8.9%	6.7%	3.6%	3.4%	2.6%	2.3%	1.7%	29.5%	100%

Table 7. Sensitivity Analysis for White Females' Mortality due to Cancers

Year	Trachea, bronchus and lung (C33-34)	Prostate (C61)	Colon, rectm and anus (C18-21)	Liver and intrahepatic bile ducts (C22)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Bladder (C67)	Meninges, brain & others (C70-72)	Breast (C50)	All Others	Grand Total	Trachea, bronchus and lung (C33-34)	Prostate (C61)	Colon, rectm and anus (C18-21)	Liver and intrahepatic bile ducts (C22)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Bladder (C67)	Meninges, brain & others (C70-72)	Breast (C50)	All Others	Grand Total
2009	12.6%	12.5%	11.8%	10.7%	11.2%	12.2%	12.7%	11.2%	10.5%	9.5%	12.1%	12.1%	29.0%	14.3%	10.2%	5.7%	5.9%	3.0%	2.4%	1.8%	1.3%	0.2%	26.1%	100%
2010	12.8%	12.7%	12.5%	11.0%	12.0%	11.6%	12.9%	11.6%	10.9%	13.3%	12.1%	12.3%	28.9%	14.3%	10.7%	5.8%	6.2%	2.8%	2.4%	1.8%	1.3%	0.2%	25.5%	100%
2011	12.6%	12.2%	12.3%	11.8%	11.7%	12.5%	12.7%	11.6%	11.0%	11.7%	12.4%	12.3%	28.4%	13.7%	10.5%	6.2%	6.1%	3.1%	2.4%	1.8%	1.4%	0.2%	26.3%	100%
2012	12.7%	12.1%	12.5%	12.5%	12.1%	12.4%	11.9%	11.9%	13.2%	10.2%	12.5%	12.4%	28.5%	13.4%	10.5%	6.5%	6.1%	3.0%	2.2%	1.8%	1.6%	0.2%	26.1%	100%
2013	12.8%	11.9%	12.5%	13.4%	12.8%	13.0%	12.2%	12.9%	12.8%	14.0%	12.4%	12.6%	28.3%	13.0%	10.4%	6.9%	6.5%	3.1%	2.2%	2.0%	1.5%	0.2%	25.8%	100%
2014	12.4%	12.0%	12.4%	13.4%	13.3%	12.6%	12.6%	13.6%	13.5%	14.3%	12.8%	12.6%	27.4%	13.1%	10.3%	6.9%	6.7%	3.0%	2.3%	2.1%	1.6%	0.2%	26.4%	100%
2015	12.2%	13.3%	12.8%	14.0%	13.6%	12.9%	12.5%	13.6%	14.0%	14.7%	12.3%	12.7%	26.8%	14.4%	10.6%	7.1%	6.8%	3.0%	2.3%	2.1%	1.7%	0.2%	25.1%	100%
2016	11.9%	13.3%	13.1%	13.4%	13.3%	12.9%	12.5%	13.6%	14.0%	12.4%	13.4%	12.9%	25.6%	14.3%	10.7%	6.8%	6.6%	3.0%	2.2%	2.0%	1.7%	0.2%	27.0%	100%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	27.8%	13.8%	10.5%	6.5%	6.3%	3.0%	2.3%	1.9%	1.5%	0.2%	26.0%	100%

Table 8. Sensitivity Analysis for Black Males' Mortality due to Cancers

The next group to analyze are the black males' mortality which is also showing that the meninges, liver and bladder's cancers are the ones showing a continuous increase on proportion in the past years. A success story is the small and consistent reduction of deaths related to trachea, bronchus and lung's cancers.

Because black females' mortality was also raised as an issue, we did the sensitivity analysis and found the liver, bladder and meninges' cancers the major concerns. Breast cancer is showing a very positive reduction for white females, but it is not doing the same for black women. We ought to see what is happening in the treatment for both groups and determine if something can be done to reduce the number of deaths.

Year	Trachea, bronchus and lung (C33-34)	Breast (C50)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Liver and intrahepatic bile ducts (C22)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Bladder (C67)	Meninges, brain & others (C70-72)	All Others	Grand Total	Trachea, bronchus and lung (C33-34)	Breast (C50)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Liver and intrahepatic bile ducts (C22)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Bladder (C67)	Meninges, brain & others (C70-72)	All Others	Grand Total
2009	12.0%	11.9%	12.6%	11.4%	10.6%	11.8%	11.5%	11.0%	11.6%	11.5%	11.8%	21.3%	18.7%	11.0%	7.1%	2.6%	2.8%	2.0%	1.4%	1.3%	31.7%	100%
2010	12.4%	12.3%	12.5%	11.5%	11.0%	11.8%	12.5%	11.8%	11.9%	11.8%	12.1%	21.6%	18.9%	10.6%	7.0%	2.6%	2.7%	2.1%	1.4%	1.4%	31.7%	100%
2011	12.5%	12.6%	12.9%	11.9%	12.0%	12.5%	12.1%	12.8%	11.5%	12.2%	12.4%	21.2%	18.9%	10.6%	7.1%	2.8%	2.8%	2.0%	1.5%	1.3%	31.9%	100%
2012	12.6%	12.5%	12.0%	12.7%	12.0%	12.4%	12.6%	11.6%	12.4%	12.3%	12.4%	21.4%	18.7%	9.9%	7.5%	2.8%	2.7%	2.1%	1.4%	1.4%	32.0%	100%
2013	12.9%	12.4%	12.1%	12.9%	12.8%	12.1%	12.3%	12.4%	11.1%	12.6%	12.6%	21.5%	18.3%	9.9%	7.5%	2.9%	2.6%	2.0%	1.5%	1.2%	32.6%	100%
2014	12.8%	12.6%	12.2%	12.8%	14.0%	12.4%	13.3%	13.0%	14.9%	12.9%	12.8%	20.9%	18.2%	9.8%	7.4%	3.1%	2.7%	2.1%	1.5%	1.6%	32.7%	100%
2015	12.4%	12.8%	12.8%	13.8%	13.7%	13.5%	12.9%	13.7%	13.3%	12.8%	12.9%	20.2%	18.5%	10.2%	7.9%	3.1%	2.9%	2.1%	1.6%	1.4%	32.2%	100%
2016	12.4%	12.9%	12.9%	12.8%	14.0%	13.5%	12.9%	13.7%	13.3%	13.8%	13.2%	19.8%	18.1%	10.1%	7.2%	3.1%	2.8%	2.0%	1.5%	1.4%	34.0%	100%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	21.0%	18.5%	10.3%	7.3%	2.9%	2.7%	2.1%	1.5%	1.4%	32.4%	100%

Table 8. Sensitivity Analysis of Black Females' Mortality due to Cancer

The last group to analyze are the Hispanics. Doing the sensitivity analysis to the cancers' mortality of Hispanic males, we are finding that most of the cancers are showing a trend up in mortality as the green color is noticeable in the first years and the last years are mostly red. The prostrate, meninges and bladder cancers contribution to the total mortality in the past 8 years are the one having larger contributions, so we strongly recommend doing further analysis on these three malignant neoplasms illnesses on the Hispanic males.

Year	Trachea, bronchus and lung (C33-34)	Colon, rectm and anus (C18-21)	Liver and intrahepatic bile ducts (C22)	Prostate (C61)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Bladder (C67)	Breast (C50)	All Others	Grand Total	Trachea, bronchus and lung (C33-34)	Colon, rectm and anus (C18-21)	Liver and intrahepatic bile ducts (C22)	Prostate (C61)	Pancreas (C25)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Bladder (C67)	Breast (C50)	All Others	Grand Total
2009	11.7%	11.0%	10.2%	10.8%	10.8%	10.8%	11.6%	10.7%	10.5%	9.8%	10.6%	10.9%	18.8%	10.9%	9.6%	9.1%	6.7%	4.8%	4.7%	3.1%	2.2%	0.1%	30.0%	100%
2010	12.0%	11.5%	10.7%	11.5%	10.7%	11.2%	11.1%	11.2%	10.8%	10.4%	11.4%	11.3%	18.5%	10.8%	9.6%	9.3%	6.4%	4.8%	4.3%	3.1%	2.1%	0.1%	30.9%	100%
2011	12.4%	11.6%	11.9%	11.8%	12.2%	11.8%	11.4%	11.4%	11.4%	11.7%	11.7%	11.9%	18.4%	10.5%	10.2%	9.2%	6.9%	4.8%	4.2%	3.0%	2.2%	0.1%	30.4%	100%
2012	12.4%	12.0%	12.6%	11.9%	12.6%	12.7%	12.1%	12.1%	12.5%	8.6%	12.3%	12.3%	17.6%	10.5%	10.5%	8.9%	6.9%	5.0%	4.3%	3.1%	2.3%	0.1%	30.8%	100%
2013	12.7%	13.1%	13.5%	12.3%	12.4%	13.1%	12.8%	11.9%	13.8%	14.7%	12.3%	12.7%	17.6%	11.1%	10.9%	8.9%	6.6%	5.0%	4.4%	3.0%	2.5%	0.1%	30.0%	100%
2014	12.8%	13.0%	13.6%	12.8%	13.5%	13.3%	13.4%	13.7%	12.8%	14.1%	13.2%	13.1%	17.0%	10.6%	10.6%	9.0%	6.9%	4.9%	4.5%	3.3%	2.2%	0.1%	30.9%	100%
2015	13.2%	13.8%	14.0%	14.5%	14.4%	13.5%	13.8%	14.5%	14.1%	17.2%	13.3%	13.7%	16.9%	10.8%	10.5%	9.7%	7.1%	4.8%	4.4%	3.4%	2.3%	0.1%	30.0%	100%
2016	12.8%	14.1%	13.6%	14.5%	13.5%	13.5%	13.8%	14.5%	14.1%	13.5%	15.3%	14.1%	15.9%	10.7%	9.8%	9.4%	6.4%	4.6%	4.3%	3.2%	2.3%	0.1%	33.3%	100%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	17.5%	10.7%	10.2%	9.2%	6.7%	4.8%	4.4%	3.2%	2.3%	0.1%	30.8%	100%

Table 9. Sensitivity Analysis of Hispanic Males' Mortality due to Cancer

Table # 10 is presenting a very worrying picture because most of the cancers are increasing for this ethnicity in the United States. The Hispanic population is growing rapidly and that might be the reason behind the mortality caused by cancers, but the meninges, bladder and liver malignant neoplasms are areas recommended for further investigation.

Year	Breast (C50)	Trachea, bronchus and lung (C33-34)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Liver and intrahepatic bile ducts (C22)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Bladder (C67)	All Others	Grand Total	Breast (C50)	Trachea, bronchus and lung (C33-34)	Colon, rectm and anus (C18-21)	Pancreas (C25)	Liver and intrahepatic bile ducts (C22)	Leukemia (C91-95)	Non-Hodgkins lymphoma (C81-85)	Meninges, brain & others (C70-72)	Bladder (C67)	All Others	Grand Total
2009	10.9%	10.8%	11.4%	10.5%	10.2%	11.4%	11.4%	10.6%	10.7%	10.6%	10.8%	16.0%	12.8%	9.8%	7.2%	5.3%	4.6%	4.1%	2.9%	1.1%	36.2%	100%
2010	11.1%	11.4%	11.5%	11.3%	10.2%	11.6%	12.0%	9.7%	11.2%	11.2%	11.2%	15.6%	13.0%	9.5%	7.5%	5.1%	4.5%	4.2%	2.5%	1.1%	36.9%	100%
2011	11.4%	11.5%	12.0%	11.5%	11.0%	11.6%	12.0%	10.4%	12.4%	11.7%	11.6%	15.6%	12.7%	9.6%	7.3%	5.3%	4.3%	4.0%	2.6%	1.2%	37.2%	100%
2012	12.6%	12.5%	12.1%	12.2%	12.4%	12.3%	12.7%	12.8%	11.8%	12.4%	12.4%	16.0%	12.8%	9.0%	7.3%	5.6%	4.3%	4.0%	3.0%	1.1%	36.8%	100%
2013	12.9%	12.7%	12.7%	13.1%	13.6%	13.1%	12.5%	13.0%	12.0%	12.6%	12.8%	15.9%	12.7%	9.2%	7.5%	5.9%	4.4%	3.8%	3.0%	1.1%	36.4%	100%
2014	13.6%	13.5%	12.9%	13.3%	13.7%	13.0%	12.7%	12.9%	13.4%	13.2%	13.3%	16.2%	13.0%	9.0%	7.4%	5.8%	4.3%	3.7%	2.8%	1.1%	36.6%	100%
2015	13.2%	14.0%	13.5%	14.7%	15.1%	13.5%	13.3%	15.3%	14.3%	13.3%	13.7%	15.3%	13.0%	9.2%	7.9%	6.2%	4.3%	3.8%	3.3%	1.2%	36.0%	100%
2016	14.2%	13.6%	14.1%	13.3%	13.7%	13.5%	13.3%	15.3%	14.3%	15.0%	14.3%	15.7%	12.2%	9.1%	6.9%	5.4%	4.1%	3.6%	3.1%	1.1%	38.7%	100%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	15.8%	12.8%	9.3%	7.4%	5.6%	4.3%	3.9%	2.9%	1.1%	36.9%	100%

Table 10. Sensitivity Analysis of Hispanic Females' Mortality due to Cancer

FINAL MULTIPLE REGRESION ANALYSIS

In order to determine the impact of the different types of cancers that we have analyzed, we conducted a multiple regression analysis to determine which different types of cancers can help us predict the potential number of people due to malignant neoplasms. The elements of gender, ethnicity and year were also included for further analysis.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000[a]	1.000	1.000	828.74721

a. Predictors: (Constant), Prostate, Breast, Liver, Lymphoma, Bladder, Pancreas, Colon

b. Dependent Variable: Cancer deaths

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	457780131077.456	7	65397161582.494	95217.053	.000[b]
	Residual	27472877.794	40	686821.945		
	Total	457807603955.250	47			

a. Dependent Variable: Cancer deaths

b. Predictors: (Constant), Prostate, Breast, Liver, Lymphoma, Bladder, Pancreas, Colon

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-838.382	414.823		-2.021	.050
	Colon	5.230	.620	.443	8.431	.000
	Pancreas	7.395	.804	.483	9.194	.000
	Liver	-2.005	.514	-.068	-3.900	.000
	Lymphoma	4.944	.681	.184	7.255	.000
	Bladder	-1.098	.636	-.039	-1.727	.092
	Breast	-.425	.351	-.048	-1.209	.234
	Prostate	.273	.408	.022	.669	.507

a. Dependent Variable: Cancer_deaths

In general, the pancreas and colon cancers are the ones that are mostly driving the number of deaths for all genders and ethnical groups followed by the lymphoma.

CONCLUSION

In this paper, we analyzed the impact of malignant neoplasms in the mortality by genders and ethnicities and we recommend medical research on regards to cancers in the pancreas, liver and bladder. Although these are not the top causes of mortality, they are showing signs of growth that could result on more deaths in the near future. The number of deaths per year of white people is growing but at a slowed rate, while the Hispanics' line is going rapidly upwards. Interestingly the two groups that have a lower growth rate are white women and black men, which could be benchmarked to the other ethnicities in order to reduce the number of deaths.

REFERENCES

Aier, Imlimaong; Semwal, Rahul; Sharma, Anju; Kumar, Pritish Varadwaj. "A systematic assessment of statistics, risk factors, and underlying features involved in pancreatic cancer," Cancer Epidemiology. Vol. 58 (2019); 104-110.

German, Robert R.; Fink, Aliza K.; Heron, Melonie; Stewart, Sherri L.; Johnson, Chris J.; Finch, Jack L.; "The accuracy of cancer mortality statistics based on death certificates in the United States," Cancer Epidemiology. Vol. 35 (2011); 126-131.

Jemal, Ahmedin; Siegel, Rebecca; Ward, Elizabeth; Taylor Murray; Xu, Jiaquan; Thun, Michael J.; "Cancer Statistics 2009" CA: A Cancer Journal for Clinicians. Vol. 59. Pages: 225-249.

Jemal, Ahmedin; Siegel, Rebecca; Xu, Jiaquan; Ward, Elizabeth. "Cancer Statistics, 2010" CA: A Cancer Journal for Clinicians. Vol. 60. Pages: 277-230

Jemal, Ahmedin; Bray, Freddie; Center, Melissa M.; Ferlay, Jacques; Ward, Elizabeth; Forman, David. “Cancer Statistics, 2011” CA Cancer Journal for Clinicians; Vol. 61. Pages: 69–90

Siegel, Rebecca; Naishadham, Deepa; Jemal, Ahmedin. “Cancer Statistics, 2012” CA Cancer Journal for Clinicians; Vol. 62. Pages: 10–29

Siegel, Rebecca; Naishadham, Deepa; Jemal, Ahmedin. “Cancer Statistics, 2013” CA Cancer Journal for Clinicians; Vol. 63. Pages: 11–30

Siegel, Rebecca; Ma, Jiemin; Zou, Zhaohui; Jemal, Ahmedin. ”Cancer Statistics, 2014” CA: A Cancer Journal for Clinicians. Vol. 64. Pages: 9-29

Siegel, Rebecca L.; Miller, Kimberly D.; Jemal, Ahmedin. “Cancer statistics, 2015” CA: A Cancer Journal for Clinicians. Vol. 65. Pages: 5-29

Siegel, Rebecca L.; Miller, Kimberly D.; Jemal, Ahmedin. “Cancer statistics, 2016” CA: A Cancer Journal for Clinicians. Vol 66. Pages: 7-30

Texas Department of State. Health Services. http://www.dshs.texas.gov/chs/vstat/annrpts.shtm

National Vital Statistics Reports, “Deaths: Final Data for 2014” Vol. 65, No.4, June 30, 2016 https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_04.pdf

National Vital Statistics Reports, “Deaths: Final Data for 2015” Vol. 66, No.6, November 27, 2017 https://www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr66_06.pdf

National Vital Statistics Reports, “Deaths: Final Data for 2016”, Vol. 67, No. 5, July 26, 2018 https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_05_tables.pdf

DECISION SCIENCES INSTITUTE

Optimal Transportation and Pricing Policies for Perishable Products Supply Chains with RFID-Generated Quality Information

Sara Babaee
Wilfrid Laurier University
Email: baba4590@mylaurier.ca

Mojtaba Araghi
Wilfrid Laurier University
Email: maraghi@wlu.ca

ABSTRACT

We study the problem of maximizing profits for a distributor who procures perishable products, transports them along the supply chain, and sails them at retail stores. The quality of perishable products deteriorates during the transportation process. Products with different levels of quality are substitutable depending on their prices and the customer's quality sensitivity. The distributor may utilize radio frequency identification (RFID) temperature tags, to track the quality of the products along the supply chain. Based on a quadratic programming problem, we determine the profit-maximizing transportation and pricing policies and evaluate the value of RFID-generated quality information.

KEYWORDS: Technology adoption, Cold chain, RFID, Transportation, Pricing

INTRODUCTION

Perishable goods, from fresh produce to pharmaceutical drugs, must be kept and distributed in a temperature-controlled environment along the entire supply chain. Controlling other factors such as humidity, light, and air pressure may also affect the quality of products delivered to end customers. Failing to transport products in the desired environment increases their quality deterioration rates and yields food (and other perishable products) waste. Food and Agriculture Organization of the United Nations estimates that 40% of the food produced in the world is wasted annually (Gustavsson *et al.*, 2011). Proper refrigeration along the food distribution network in the United States can decrease the food waste by 12% (NRDC, 2012) which significantly reduces its $218 billion annual economic impact. (Mercier *et al.*, 2017)
The series of activities and equipment for producing, storing, and transporting perishable products in a refrigerated environment is referred to as the cold chain. The cold chain maintains the desired low-temperature range for perishable products. However, there is a significant temperature heterogeneity inside each transportation equipment (Raab *et al.*, 2008; Margeirsson *et al.*, 2012) which can considerably affect the quality and the remaining shelf-life of fresh products. Therefore, it is critical to closely monitor the temperature profile of each perishable item during the transportation process and update the distribution policy accordingly. Radio frequency identification (RFID) tags have been widely used in temperature-monitoring systems in cold chains in recent years (Mercier et al. 2017 and references therein). However, the value of the information provided by this technology is not well quantified in the literature. Hence, producers and distributors are reluctant to invest in enhancing their cold chain management systems by implementing RFID temperature tags. RFID generated information can be implemented in inventory system to prioritize selling those pallets with lower qualities.

Also, it can enable retailers to better understand the quality of the products that they receive from suppliers and improve their inventory management and shopper satisfaction. The other area that could take advantage of these new level of information is transportation. In this paper, we are trying to quantify the benefits of implementing RFID and sensor technology in transportation of perishable products and analyze how the profit and waste change after the adoption of RFID and sensor technology from a distributer's point of view.

In this work, we investigate a supply chain in which a distributor procures perishable products, transports them along the supply chain, and sells them at retail stores. The distributor's goal is to maximize its profit by jointly optimizing the transportation (route selection and freshness keeping efforts) and pricing policies. The contributions of this paper are threefold: (1) We develop a mathematical programming model to optimize transportation and pricing policies for a cold chain in which the demand is vertically differentiated at retailers and customers are both price and quality sensitive; (2) We compare the profit of the cold chain with and without RFID implementation to quantify the value of the information provided by this technology; (3) We show that in the absence of strict environmental regulations, a distributor may exploit the RFID technology to increase its profit at the expense of the food waste.

The rest of this paper is organized as follows. First, we briefly review the related literature. Then, we develop our mathematical programming model and using a numerical example, we discuss the benefits and costs of using RFID technology. Finally, we present our concluding remarks in the last section.

LITERATURE REVIEW

There is a large body of literature descriptively studying the benefits of implementing RFID tags in long-haul transportation of perishable products including fresh crops (Delen, Sharda and Hardgrave, 2011), seafood (Abad *et al.*, 2009; Trebar, Lotrič and Fonda, 2015), meat (Narsing, 2005; Raab, Petersen and Kreyenschmidt, 2011), and medicine (Monteleone, Sampaio and Maia, 2017). In particular, Raab et al. (2011) observe that while the continuous control of environmental conditions is critical in the cold chain management, its implementation cost is a major obstacle. To address this issue, . (Badia-Melis *et al.*, 2018) propose the integration of RFID-based technologies with temperature estimation methods to reduce the number of required sensors in cold chain monitoring systems

In the operations management literature, more attention has been paid over the last decade to analytically improving the efficiency of perishable products supply chains. We refer to (He *et al.*, 2018) for a review of the related analytical models.

Tracking the perishable products plays a vital role in the inventory and revenue management. (Gaukler, Ketzenberg and Salin, 2017) show that retailers may benefit from eliminating the risk of selling contaminated products by using RFID tags to determine dynamic expiration dates for perishable products. (Piramuthu and Zhou, 2013) study an inventory management and shelf-space allocation model for fresh produce, where detailed item-level quality information is generated using auto-ID technologies such as RFID with necessary sensors. They assume the demand rate depends on both the display inventory and the freshness of products. (Liu, Tang and Huang, 2008) study an inventory-pricing problem for the supermarkets using the RFID technology to gather accurate information on the quality of perishable items. Their model shows promising benefits for both deterministic and stochastic demand functions. In contrast to these papers which are focused on the inventory and the revenue management, we investigate transportation policies in cold chains.

(Grunow and Piramuthu, 2013) use the RFID-generated item-level information to determine the accurate remaining shelf-life for highly perishable products. They derive conditions under which the RFID implementation in the food supply chain is beneficial for distributors, retailers, or consumers. (Shi, Zhang and Qu, 2010) optimize the distribution strategy when frequent updates of the product quality are available through RFID tags. By sequentially adjusting transportation decisions, the distributor minimizes the cost of meeting the demand while preserving the product quality. Unlike our paper, these studies do not consider the pricing as a lever for the distributor to control the demand for different quality levels of a product.

It is worth to briefly discuss the operations management literature on the transportation and pricing of perishable products without using the updated quality information. (Cai *et al.*, 2010) analyze a supply chain in which a distributor buys fresh products from a manufacturer and distributes them in a target market. They consider different levels of freshness keeping efforts during the transportation which yields different levels of quality at the target market. Assuming the demand depends on both the freshness and the price of products, they propose a model which determines the order quantity, the level of freshness-keeping efforts, and the wholesale (set by the manufacturer) and the retail (set by the distributor) prices. (Cai and Zhou, 2014) consider the problem of supplying a market with a perishable product, using an inexpensive public transportation system which is subject to frequent disruptions. They study both make-to-stock and make-to-order systems and minimize the expected loss by determining the optimal production and delivery schedule. (Ahumada and Villalobos, 2011) present an integrated production-distribution model considering the resource availability, pricing, transportation, and inventory decisions. (Soysal *et al.*, 2018) study an inventory routing problem by considering the perishability of products, demand uncertainty, and logistics cost and fuel consumption. They illustrate the benefit of the horizontal collaboration among suppliers and its sensitivity to the supplier size and the maximum shelf life of products.

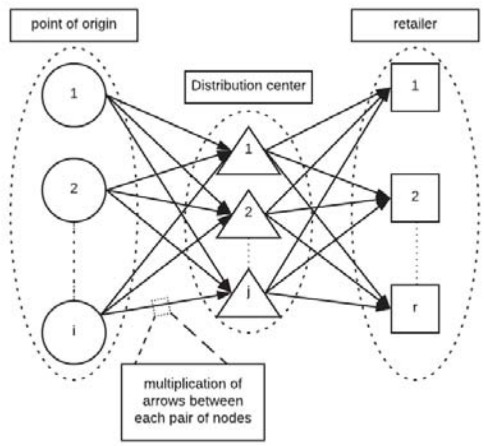
The dynamic pricing of perishable products is frequently discussed in the revenue management literature. (Akçay, Natarajan and Xu, 2010) consider a firm selling horizontally or vertically differentiated substitutable and perishable products. Solving a dynamic pricing problem, they determine the relationship between the optimal prices and inventory for different levels of quality. (Chew *et al.*, 2014) show that how a firm can optimally shift the demand between different quality levels of substitutable and perishable products using dynamic pricing and ordering decisions. (Li, Zhang and Tang, 2014) consider perishable products in a stochastic inventory system and determine the optimal pricing and inventory policies. In a series of papers, (Herbon, 2015, 2016, 2017, 2018) studies dynamic quality-based pricing of perishable products considering price discrimination for homogeneous and heterogeneous customers who are sensitive to the price and freshness of products.

MODEL FORMULATION

We investigate a supply chain in which a distributor procures fresh but perishable products from a set of suppliers, ships them through a set of distribution centers (DCs), and sells them in a set of retail stores after choosing their retail prices. The distributor's goal is to maximize its profit by jointly optimizing transportation and pricing policies.

To determine the optimal transportation policy through the supply chain, we need to consider series of transportation-related decisions including the packaging, route selection, transportation mode and distance, and freshness keeping efforts such as controlling the temperature and humidity. We call each combination of these decisions a transportation option and include all corresponding costs of each transportation option in its transportation cost per unit of the product. Figure 1 presents a schematic view of our cold supply chain network. Note that each arrow may represent a series of transportation options between the corresponding nodes.

Figure 1: Schematic View of the cold supply chain



For the sake of the model simplicity, we consider a single product type which can be categorized into three broad classes based on its quality: High-quality, Acceptable-quality, and Unacceptable-quality. We also assume that all suppliers offer the product with the same initial quality, normalized to be considered as the high-quality. While these assumptions simplify our notation, they will not affect the insights of the model, and the results can be easily extended to include models where suppliers offer different levels of quality of different products.
After the procurement, the distributor chooses among the transportation options to ship the perishable items from the suppliers to DCs. All items going through the same transportation option receive the same level of freshness keeping efforts. However, as discussed in the introduction, they encounter an inevitable variation in the temperature, humidity, and other environmental conditions. Therefore, the deterioration of each specific item during the transportation can be seen as a stochastic process with a rate that depends on the chosen transportation option.
While we assume all items are at the high-quality level at suppliers, by the time they reach DCs, some of them may deteriorate to acceptable-quality level. We separate items at DCs based on their quality levels using the information gathered by RFID tags. Then we choose the transportation options from DCs to retailers according to the quality of the items. One intuitive example of such transportation policy is sending high-quality items with a slower but less expensive option, while expediting the transportation of acceptable-quality products through an expensive transportation mode in order to deliver them to retailers before they deteriorate to the unacceptable-quality level.
The distributor then sets the price for each quality level at each retailer's location. We let the distributor choose different prices at different retail stores, considering the cost of transportation options as well as the retailers' demand and quality preferences. Finally, at each retailer, customers choose among the items at different quality levels and different prices, which we discuss in detail in the next section.

Quality Based Demand Allocation

While we ship the high-quality items through the cold chain, their quality may remain as high (we call it quality level 1) or decrease to an acceptable (quality level 2) or even an unacceptable level. At each retailer location r, products are vertically differentiated based on their quality level. Let α^q denote the desirability of a product with quality level $q \in \{1,2\}$. A customer's utility from purchasing one item at quality level q from the retailer r, is given by

$$\mu_r^q = \theta_r \alpha^q - p_r^q \qquad (1)$$

where θ_r is the quality sensitivity of the retailer r 's customers, and p_r^q is the price of an item with the quality level q at the retailer r. The quality sensitivity factor θ_r is a customer specific attribute. We let customers patronizing different stores have different quality sensitivities, e.g., due to the demographic characteristics of the retailer's location. For the retailer r 's customers, we let θ_r to be uniformly distributed between 0 and τ_r.

We assume the distributor is looking to segment the market such that both quality levels can be sold at the retailers. Hence, the retail prices at each retailer r should follow Lemma 1, which is in line with pricing schemes for vertically differentiated products in (Akçay, Natarajan and Xu, 2010).

Lemma 1 Assume the customer quality sensitivity θ_r is uniformly distributed between 0 and $\tau_r$at the retailer r. To segment the market into both quality levels, the retailer prices for the quality level 1 and 2 should satisfy

$$0 \leq \frac{p_r^2}{\alpha^2} \leq \tau_r \qquad (2)$$

$$0 \leq \frac{p_r^1 - p_r^2}{\alpha^1 - \alpha^2} \leq \tau_r \qquad (3)$$

Proof The price to desirability ratios in (2) are larger than zero by definition and are smaller than τ_r to keep the utility for both quality levels positive. Moreover, if p_r^2/α^2 becomes larger than p_r^1/α^1, all customers will prefer items with quality 1 regardless of their corresponding quality sensitivity θ_r. Similarly, the ratio in (3) is positive as a higher quality and desirability is corresponding to a higher price. Finally, if $p_r^1 - p_r^2$ grows τ_r times larger than $\alpha^1 - \alpha^2$, then the utility of purchasing a quality level 2 item becomes greater than the one for a quality level 1. Thus, all customers prefer items at quality level 2 regardless of their θ_r.

Assuming the total demand at the retailer r is D_r, under the price setting suggested by Lemma 1, it is straightforward to show that the expected demand for the high-quality items is:

$$D_r^1 = \left(1 - \frac{p_r^1 - p_r^2}{\tau_r(\alpha^1 - \alpha^2)}\right) D_r \qquad (4)$$

and for the acceptable quality items is:

$$D_r^2 = \left(\frac{p_r^1 - p_r^2}{\tau_r(\alpha^1 - \alpha^2)} - \frac{p_r^2}{\tau_r \alpha^2}\right) D_r \qquad (5)$$

Note that if the firm decides not to sell a specific quality level at a specific retailer (e.g., due to the excessive cost of delivering high-quality products to a remote retailer), it will have the flexibility to modify the prices to shift the demand of one quality level to another. More specifically, for a given p_r^2, by increasing p_r^1 to a level that makes $(p_r^1 - p_r^2)/(\alpha^1 - \alpha^2) = \tau_r$, the firm can eliminate the demand of high-quality products and increase D_r^2 to $(1 - p_r^2/\tau_r \alpha^2) D_r$. Similarly, if the firm decides to sell only the high-quality products, it is enough to increase p_r^2 to a level that $p_r^2/\alpha^2 = p_r^1/\alpha^1$. At this price, the demand for the quality 2 products shifts to the quality 1 products. Hence D_r^1 increases to $(1 - p_r^1/\tau_r\alpha^2)D_r$.

Model In The Presence Of RFID

In this section, we propose a quadratic optimization problem based on the cold chain network in Figure 1. We present the set of suppliers by S, DCs by D and retailers by R. We assume that there are N_{ij} transportation options between node $i \in \{S, D\}$ and node $j \in \{D, R\}$. We define x_{ijn}^1 and x_{ijn}^2 as the number of items with quality level 1 and 2 to be shipped from node i to node j by transportation option $n \in N_{ij}$, respectively. The corresponding transportation cost is denoted by C_{ijn} per item independent of the item quality, and the procurement cost of one high-quality item at the supplier s is denoted by C_s'. Finally, we let $z_r^1(z_r^2)$ present the number of items with quality 1(2) sold at the retailer r with the selling price $p_r^1(p_r^2)$.
The rate of quality deterioration depends on both the chosen transportation option and the quality of the item. Let λ_{ijn}^1 denote the percentage of high-quality items (quality level 1) that deteriorates to acceptable-quality items during the transportation from nodes i to node j through the transportation option n (we assume the rest of items remain as high-quality). Similarly, let λ_{ijn}^2 denote the percentage of acceptable-quality items (quality level 2) whose quality depreciates to a level that is no longer acceptable by customers and should be discarded when the transportation option n is chosen between nodes i and j. Table 1 summarizes our notation.

Table 1: Summary of Notations	
Indices	
S	Set of suppliers
D	Set of DCs
R	Set of retailers
N_{ij}	Set of transportation options between node i and j
Decision Variables	
z_r^q	Number of quality level $q \in \{1,2\}$ item sold at retailer r
p_r^q	Price of quality level $q \in \{1,2\}$ item at retailer r
x_{ijn}^q	Number of items with quality level $q \in \{1,2\}$ to be shipped between node i and node j by transportation option $n \in N_{ij}$
System Parameters	
α^q	Desirability of a product with quality level $q \in \{1,2\}$
θ_r	Quality sensitivity of a retailer r's customer, uniformly distributed in $[0, \tau_r]$
D_r	Total expected demand at retailer r
C_s'	Procurement cost of one high quality item at supplier s
C_{ijn}	Cost of shipping one item from node i to node j by transportation option n
λ_{ijn}^q	Percentage of quality level $q \in \{1,2\}$ items deteriorates to the next (lower) quality level while getting shipped from node i to node j by transportation option n

Using our notation, we present our profit-maximizing quadratic programming problem in (6-16).

$$\max \pi = \sum_{r\in R}(p_r^1 z_r^1 + p_r^2 z_r^2) - \sum_{s\in S}\sum_{d\in D}\sum_{n\in N_{sd}}\left(C_s^{'} + C_{sdn}\right)x_{sdn}^1 - \sum_{d\in D}\sum_{r\in R}\sum_{n\in N_{dr}} C_{rdn}(x_{drn}^1 + x_{drn}^2) \tag{6}$$

Subject to

$$z_r^1 \leq \sum_{d\in D}\sum_{n\in N_{dr}} x_{drn}^1\left(1-\lambda_{drn}^1\right) \qquad \forall\, r\in R \tag{7}$$

$$z_r^1 \leq \left(1 - \frac{p_r^1 - p_r^2}{\tau_r(\alpha^1-\alpha^2)}\right)D_r \qquad \forall\, r\in R \tag{8}$$

$$z_r^2 \leq \sum_{d\in D}\sum_{n\in N_{dr}} x_{drn}^1\lambda_{drn}^1 + x_{drn}^2\left(1-\lambda_{drn}^2\right) \qquad \forall\, r\in R \tag{9}$$

$$z_r^2 \leq \left(\frac{p_r^1 - p_r^2}{\tau_r(\alpha^1-\alpha^2)} - \frac{p_r^2}{\tau_r\alpha^2}\right)D_r \qquad \forall\, r\in R \tag{10}$$

$$\sum_{s\in S}\sum_{n\in N_{sd}} x_{sdn}^1\left(1-\lambda_{sdn}^1\right) \geq \sum_{r\in R}\sum_{n\in N_{dr}} x_{drn}^1 \qquad \forall\, d\in D \tag{11}$$

$$\sum_{s\in S}\sum_{n\in N_{sd}} x_{sdn}^1\lambda_{sdn}^1 \geq \sum_{r\in R}\sum_{n\in N_{dr}} x_{drn}^2 \qquad \forall\, d\in D \tag{12}$$

$$\frac{p_r^2}{\alpha^2} \leq \frac{p_r^1}{\alpha^1} \qquad \forall\, r\in R \tag{13}$$

$$\frac{p_r^1}{\alpha^1} \leq \tau_r \qquad \forall\, r\in R \tag{14}$$

$$\frac{p_r^1 - p_r^2}{(\alpha^1-\alpha^2)} \leq \tau_r \qquad \forall\, r\in R \tag{15}$$

$$x_{sdn}^1, x_{drn}^1, x_{drn}^2, p_r^1, p_r^2, z_r^1, z_r^2 \geq 0 \qquad \forall\, s,d,r,n \tag{16}$$

The distributor's objective (6) is to maximize the profit of selling different quality levels of products to potential customers by taking into account the procurement cost of high-quality products, the cost of delivering high-quality products from suppliers to DCs, and the cost of delivering high- and acceptable-quality products from DCs to retailers.
Constraints (7) and (8) together guarantee that the sales of quality 1 items at any retailer r are smaller than the minimum of available quality 1 items, and the demand for quality 1, respectively. Similarly, constraints (9) and (10) together guarantee that, at any retailer r, the number of sold quality 2 items is smaller than the minimum of its inventory and demand. Constraints (11) and (12) are DC balancing constraints, limiting the total number of quality 1 and 2 items sent out from each DC to the number of quality 1 and 2 items received at that DC, respectively. Finally, constraints (13)-(15) restrict the candidate prices to the set discussed in Lemma 1 to keep both products appealing to customers, and constraint (16) is the set of non-negativity constraints.
It follows from (11) and (12) that the firm may choose to dispose of some items in DCs, instead of shipping them out. This waste happens despite paying the procurement and transportation cost to DCs. Using RFID tags, the distributor can separate high- and acceptable-quality products and may decide not to ship some of the acceptable-quality items which eventually would not be sold at retailers (either due to the demand or the final quality).

Note that we have not included the RFID adoption cost in our model. Instead, we will compare the optimal profit of this model with the optimal profit of a cold supply chain not enhanced by RFID. Determining the extra profit generated by adopting this technology enables practitioners to choose among different available types of RFID and sensor technologies.

Model in The Absence Of RFID

Without the implementation of the RFID technology, the distributor cannot separate different quality levels at DCs. Therefore, at each supplier, the distributor is essentially choosing the complete route each item goes through to reach the retailers. We define a new decision variable x^1_{sdrnm} which represent the number of high-quality items shipped from the supplier s to DC d thorough the option $n \in N_{sd}$, and then from DC d to the retailer r through the option $m \in N_{dr}$. As all items are fresh at suppliers, there is no item with quality level 2 at the node $s \in S$.
We model the transportation-pricing problem of the distributor as follows:

$$max\,\pi = \sum_{r\in R}(p_r^1 z_r^1 + p_r^2 z_r^2) - \sum_{s\in S}\sum_{d\in D}\sum_{r\in R}\sum_{n\in N_{sd}}\sum_{m\in N_{dr}}\left(C_s' + C_{sdn} + C_{rdm}\right)x^1_{sdrnm} \qquad (17)$$

Subject to

$$z_r^1 \leq \sum_{s\in S}\sum_{d\in D}\sum_{n\in N_{sd}}\sum_{m\in N_{dr}} x^1_{sdrnm}\left(1-\lambda^1_{sdn}\right)\left(1-\lambda^1_{drn}\right) \qquad \forall\, r\in R \qquad (18)$$

$$z_r^1 \leq \left(1 - \frac{p_r^1 - p_r^2}{\tau_r(\alpha^1-\alpha^2)}\right)D_r \qquad \forall\, r\in R \qquad (19)$$

$$z_r^2 \leq \sum_{s\in S}\sum_{d\in D}\sum_{n\in N_{sd}}\sum_{m\in N_{dr}} x^1_{sdrnm}\left(\left(1-\lambda^1_{sdn}\right)\lambda^1_{drn} + \lambda^1_{sdn}\left(1-\lambda^1_{drn}\right)\right) \qquad \forall\, r\in R \qquad (20)$$

$$z_r^2 \leq \left(\frac{p_r^1 - p_r^2}{\tau_r(\alpha^1-\alpha^2)} - \frac{p_r^2}{\tau_r\alpha^2}\right)D_r \qquad \forall\, r\in R \qquad (21)$$

$$\frac{p_r^2}{\alpha^2} \leq \frac{p_r^1}{\alpha^1} \qquad \forall\, r\in R \qquad (22)$$

$$\frac{p_r^1}{\alpha^1} \leq \tau_r \qquad \forall\, r\in R \qquad (23)$$

$$\frac{p_r^1 - p_r^2}{(\alpha^1-\alpha^2)} \leq \tau_r \qquad \forall\, r\in R \qquad (24)$$

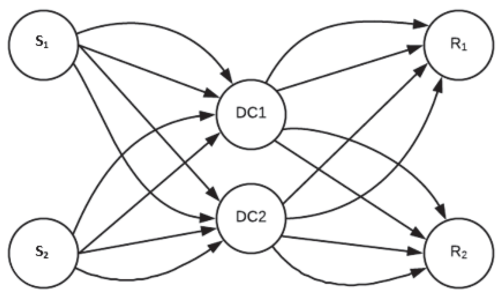
$$x^1_{sdrnm}, p_r^1, p_r^2, z_r^1, z_r^2 \geq 0 \qquad \forall\, s,d,r,n,m \qquad (25)$$

The modeling intuition behind each equation remains similar to the one in (6)-(16) discussed in the proposed model in presence of RFID. However, considering our new transportation decision variables, the objective function and the right-hand side of constraints (18) and (20) are updated. Moreover, the balancing constraints for DCs are implicitly included in the definition of x^1_{sdrnm} variables, and thus omitted.

NUMERICAL STUDY

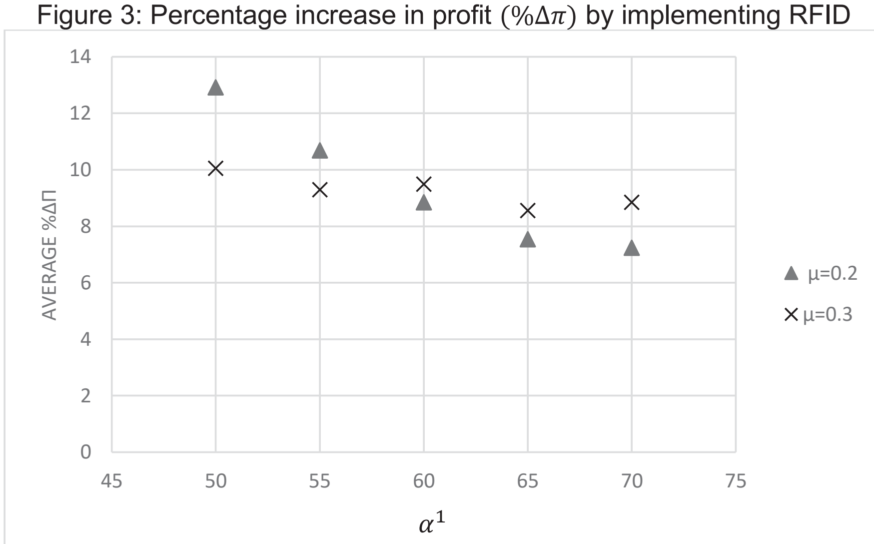
To evaluate the performance of our proposed transportation-pricing model, we use GUROBI software in R platform to find the maximum profit in the presence and absence of RFID tags for a series of numerical examples. We consider a stylized network shown in Figure 2, consists of two suppliers, two DCs and two retailers. We assume the distributor has two transportation options between each pair of nodes. The procurement cost at both suppliers is $20 per item. Intuitively, there is an inverse relation between the transportation cost and deterioration rate of each transportation option: Higher freshness keeping efforts and faster transportation both increase the transportation cost and decrease the deterioration rate. In particular, we assume that the cost of shipping an item through a transportation option equals the inverse of the square root of its deterioration rate for high-quality items. In order to model the faster deterioration of items with lower quality, we assume that the deterioration rate for quality level 2 items is twice as large as the deterioration rate for quality level 1 items for each transportation option. The expected demand at retailer 1 and 2 are assumed to be 1500 and 1000 items, respectively.

Figure 2: supply chain network used in numerical example



We assume that the deterioration rate of high-quality products is normally distributed with mean μ and standard deviation σ. We consider two sets of scenarios with $\mu = 0.2$ and $\mu = 0.3$, with $\sigma = 0.05$ for both sets. In each set, we fix the desirability of items with quality level 2 at $\alpha^2 = 40$ and consider five cases for the desirability of items with quality level 1, $\alpha^1 = \{50,55,60,65,70\}$. For each of our 10 scenarios, we simulate 100 randomly generated replications.
We calculate the percentage change in the profit, denoted by $\%\Delta\pi$, by comparing the maximum profit with and without enhancing the supply chain by RFID-based sensors. This percentage represents the value gained by adopting this technology. For all replications, $\%\Delta\pi$ is positive or close enough to zero, such that the reduction in profit can be explained by approximation and rounding errors.

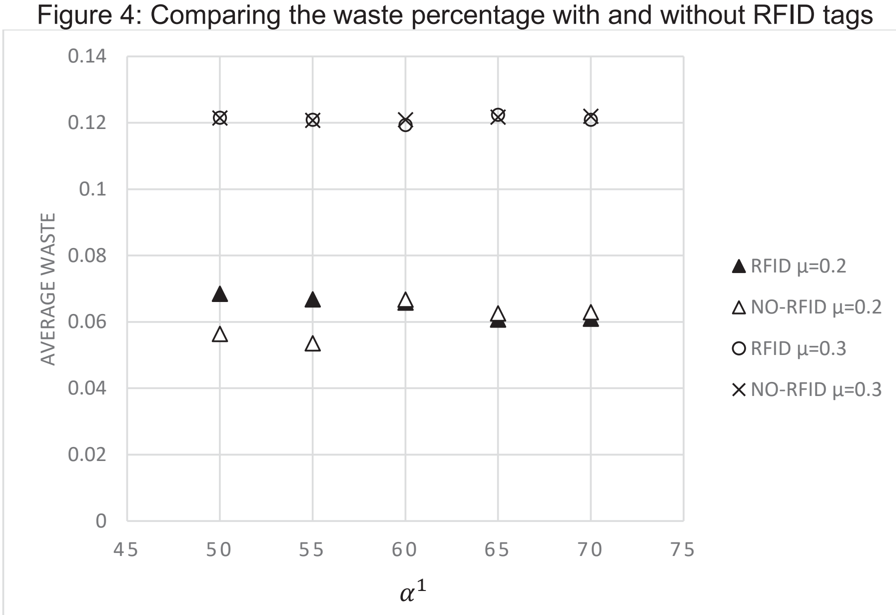
Figure 3: Percentage increase in profit ($\%\Delta\pi$) by implementing RFID



The primary benefit of the updated quality information in our model is to optimally handle items with already reduced quality at DCs. When we use RFID, a transportation option with lower deterioration rate and higher cost may be used for shipping quality level 2 items, if they are valuable enough to cover the extra cost of the expedited transportation. Therefore, we expect a higher value of the RFID implementation when acceptable-quality items are almost as desirable as high-quality items. When the desirability of level 1 increases in comparison to level 2, we observe in Figure 3 that the value of the RFID technology decreases at $\mu = 0.2$. We are not observing the same profit gain when $\mu = 0.3$, as the higher average deterioration rates strike items at quality level 2 harder.

We also compare the percentage of the high-quality items which are wasted, i.e., bought from suppliers but not sold at retailers (discarded at DCs or retailers, or deteriorated to unacceptable quality during the transportation). Figure 4 shows the average percentage of items wasted under different scenarios (with and without RFID tags).

Figure 4: Comparing the waste percentage with and without RFID tags



It is not surprising that a higher average deterioration rate yields a higher percentage of items wasted. However, as it can be seen in Figure 4, the adoption of RFID does not improve the waste level. In contrast, by comparing Figures 3 and 4, we observe that the RFID implementation increases the profit in the cost of tolerating an increase in the food waste. This is the result of our objective function which is focused on increasing the profit. In other words, we cannot merely rely on sensor technologies to decrease the waste. We need some additional constraints or penalties to make sure that the RFID technology does not encourage the distributor to ship more items knowing that it can separate and discard lower quality items later at DCs.

CONCLUDING REMARKS

We study the profit-maximizing transportation and pricing problem for a distributor with price and quality sensitive customers. The distributor decides on the procurement level from the suppliers, transportation policy from suppliers to distribution centers to retailers, and selling prices for different quality levels at different retailers. Comparing the maximum profit in the presence and absence of wireless sensor technologies that can track the products' quality along the supply chain, we derive the value of implementing such a technology.

Our results show that in the absence of strict environmental regulations, enhancing the cold chain by RFID temperature tags provides the distributor with incentives to increase its procurement level, as the temperature abused items can be identified and separately handled later at DCs.

Our study categorizes the perishable products into three groups based on their quality levels: high, acceptable, and unacceptable. It is straightforward to extend our mathematical model to consider multiple quality levels. In this case, the value of implementing RFID tags increases as it will provide more opportunities to customize the pricing and distribution policies. Continuous quality score can also be considered for the products using exponential or Gamma distribution (Yousefi *et al.*, 2019) for the quality.

In this work, we focused on products whose quality are observable for the end customers. If the quality of the product is not observable at retailers (e.g., due to packing the product in a sealed container), then the value of implementing RFID tags will increase. Under this scenario, we cannot differentiate between different quality levels even at the retailers. Thus, we should set a single retail price at each location based on the expected quality of the product and its desirability.

Finally, we can further extend our model by considering the capacity of our transportation options. Adding restrictions on full or less than truckload shipping, or on the total capacity of each option, will broaden the applicability of our model for practical purposes.

REFERENCES

Abad, E. *et al.* (2009) ‘RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain’, *Journal of Food Engineering*. Elsevier, 93(4), pp. 394–399. doi: 10.1016/J.JFOODENG.2009.02.004.

Ahumada, O. and Villalobos, J. R. (2011) ‘A tactical model for planning the production and distribution of fresh produce’, *Annals of Operations Research*. Springer US, 190(1), pp. 339–358. doi: 10.1007/s10479-009-0614-4.

Akçay, Y., Natarajan, H. P. and Xu, S. H. (2010) ‘Joint Dynamic Pricing of Multiple Perishable Products Under Consumer Choice’, *Management Science*. INFORMS , 56(8), pp. 1345–1361. doi: 10.1287/mnsc.1100.1178.

Badia-Melis, R. *et al.* (2018) ‘New trends in cold chain monitoring applications - A review’, *Food Control*. Elsevier, 86, pp. 170–182. doi: 10.1016/J.FOODCONT.2017.11.022.

Cai, X. *et al.* (2010) ‘Optimization and coordination of fresh product supply chains with freshness-keeping effort’, *Production and Operations Management*. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1937-5956.2009.01096.x/full.

Cai, X. and Zhou, X. (2014) ‘Optimal Policies for Perishable Products When Transportation to Export Market Is Disrupted’, *Production and Operations Management*. Wiley/Blackwell (10.1111), 23(5), pp. 907–923. doi: 10.1111/poms.12080.

Chew, E. P. *et al.* (2014) ‘Optimal dynamic pricing and ordering decisions for perishable products’, *International Journal of Production Economics*, 157, pp. 39–48. doi: 10.1016/j.ijpe.2013.12.022.

Delen, D., Sharda, R. and Hardgrave, B. (2011) ‘The promise of RFID-based sensors in the perishables supply chain’, *IEEE Wireless Communications*, 18(2), pp. 82–88. doi: 10.1109/MWC.2011.5751300.

Gaukler, G., Ketzenberg, M. and Salin, V. (2017) ‘Establishing dynamic expiration dates for

perishables: An application of rfid and sensor technology', *International Journal of Production Economics*, 193, pp. 617–632. doi: 10.1016/j.ijpe.2017.07.019.
Grunow, M. and Piramuthu, S. (2013) 'RFID in highly perishable food supply chains – Remaining shelf life to supplant expiry date?', *International Journal of Production Economics*, 146(2), pp. 717–727. doi: 10.1016/j.ijpe.2013.08.028.
Gustavsson, J. *et al.* (2011) *Global food losses and food waste extent, causes and prevention*. Available at: http://www.fao.org/3/a-i2697e.pdf (Accessed: 13 May 2019).
He, Y. *et al.* (2018) 'Quality and Operations Management in Food Supply Chains: A Literature Review', *Journal of Food Quality*. Hindawi, 2018, pp. 1–14. doi: 10.1155/2018/7279491.
Herbon, A. (2015) 'Optimal two-level piecewise-constant price discrimination for a storable perishable product', *International Journal of Production Research*. doi: 10.1080/00207543.2015.1018451org/10.1080/00207543.2015.1018451.
Herbon, A. (2016) 'Optimal piecewise-constant price under heterogeneous sensitivity to product freshness', *International Journal of Production Research*. Taylor & Francis, 54(2), pp. 365–385. doi: 10.1080/00207543.2014.997402.
Herbon, A. (2017) 'Should retailers hold a perishable product having different ages? The case of a homogeneous market and multiplicative demand model', *International Journal of Production Economics*. Elsevier, 193, pp. 479–490. doi: 10.1016/J.IJPE.2017.08.008.
Herbon, A. (2018) 'Optimal Pricing and Replenishment of an Expiring Inventoried Product under Heterogeneous Consumer Sensitivities', *Decision Sciences*. Wiley/Blackwell (10.1111), 49(3), pp. 522–552. doi: 10.1111/deci.12276.
Li, S., Zhang, J. and Tang, W. (2014) 'Joint dynamic pricing and inventory control policy for a stochastic inventory system with perishable products', *International Journal of Production Research*, 53(10), pp. 2937–2950. doi: 10.1080/00207543.2014.961206.
Liu, X., Tang, O. and Huang, P. (2008) 'Dynamic pricing and ordering decision for the perishable food of the supermarket using RFID technology', *Asia Pacific Journal of Marketing and Logistics*, 20(1), pp. 7–22. doi: 10.1108/13555850810844841.
Margeirsson, B. *et al.* (2012) 'Temperature fluctuations and quality deterioration of chilled cod (Gadus morhua) fillets packaged in different boxes stored on pallets under dynamic temperature conditions', *International Journal of Refrigeration*. Elsevier, 35(1), pp. 187–201. doi: 10.1016/J.IJREFRIG.2011.09.006.
Mercier, S. *et al.* (2017) 'Time-Temperature Management Along the Food Cold Chain: A Review of Recent Developments', *Comprehensive Reviews in Food Science and Food Safety*. John Wiley & Sons, Ltd (10.1111), 16(4), pp. 647–667. doi: 10.1111/1541-4337.12269.
Monteleone, S., Sampaio, M. and Maia, R. F. (2017) 'A novel deployment of smart Cold Chain system using 2G-RFID-Sys temperature monitoring in medicine Cold Chain based on Internet of Things', in *2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*. IEEE, pp. 205–210. doi: 10.1109/SOLI.2017.8120995.
Narsing, A. (2005) *RFID and supply chain management: an assessment of its economic, technical, and productive viability in global operations - ProQuest*, *Journal of applied business research*. doi: 38163658.
NRDC (2012) *Wasted: How America Is Losing Up to 40 Percent of Its Food from Farm to Fork to Landfill | NRDC*. Available at: https://www.nrdc.org/resources/wasted-how-america-losing-40-percent-its-food-farm-fork-landfill (Accessed: 13 May 2019).
Piramuthu, S. and Zhou, W. (2013) 'RFID and perishable inventory management with shelf-space and freshness dependent demandPiramuthu, S. and Zhou, W. (2013) "RFID and perishable inventory management with shelf-space and freshness dependent demand", International Journal of Production Econo', *International Journal of Production Economics*, 144(2), pp. 635–640. doi: 10.1016/j.ijpe.2013.04.035.
Raab, V. *et al.* (2008) 'Generic model for the prediction of remaining shelf life in support of cold

chain management in pork and poultry supply chains', *Journal on Chain and Network Science*. Wageningen Academic Publishers , 8(1), pp. 59–73. doi: 10.3920/JCNS2008.x089.
Raab, V., Petersen, B. and Kreyenschmidt, J. (2011) 'Temperature monitoring in meat supply chains', *British Food Journal*. Emerald Group Publishing Limited, 113(10), pp. 1267–1289. doi: 10.1108/00070701111177683.
Shi, J., Zhang, J. and Qu, X. (2010) 'Optimizing distribution strategy for perishable foods using RFiD and sensor technologies', *Journal of Business & Industrial Marketing*. Emerald Group Publishing Limited, 25(8), pp. 596–606. doi: 10.1108/08858621011088338.
Soysal, M. *et al.* (2018) 'Modeling a green inventory routing problem for perishable products with horizontal collaboration', *Computers & Operations Research*, 89, pp. 168–182. doi: 10.1016/j.cor.2016.02.003.
Trebar, M., Lotrič, M. and Fonda, I. (2015) 'Use of RFID temperature monitoring to test and improve fish packing methods in styrofoam boxes', *Journal of Food Engineering*, 159, pp. 66–75. doi: 10.1016/j.jfoodeng.2015.03.007.
Yousefi, N. *et al.* (2019) 'Optimization of on-condition thresholds for a system of degrading components with competing dependent failure processes', *Reliability Engineering & System Safety*. Elsevier, p. 106547. doi: 10.1016/J.RESS.2019.106547.

DECISION SCIENCES INSTITUTE

Enhancing Supply Chain Resiliency: An Organizational Information Processing Perspective

Ahmed Hamdi
Rabat Business School
Email: ahmed.hamdi@uir.ac.ma

Arash Azadegan
Rutgers Business School
Email: aazadegan@business.rutgers.edu

Bouchaib Bahli
Ted Rogers School of Management - Ryerson University
Email: bahli@ryerson.ca

ABSTRACT

To ensure business continuity and protect its value, the firm needs to enhance its supply chain resiliency. Based on the organizational information processing theory, we consider the enhancement of supply chain resiliency as an information-based process. In our study, we posit that implementing supply chain risk management practices creates redundancy that enables a reduction of the information processing needs. Whereas, the investment in supply chain integration activities will help the firm increase its information processing capacity. Analysis of data collected from 283 US and French manufacturers, supports our arguments and offers theoretical and managerial insights.

KEYWORDS: Supply chain resilience, Supply chain integration, Supply chain risk management, Organizational Information processing theory.

INTRODUCTION

Understanding and building supply chain resiliency is increasingly becoming a necessity as supply chains are becoming more complex and supply chain disruptions often impact the organizations severely. This has been evident in recent natural disasters. The recent hurricane Maria at Puerto Rico on 20 September 2017, has been recorded as the most powerful storm to hit the Caribbean island in 85 years (Lewin, Roegiers, Tuazon, & Almond, 2017). Puerto Rico is a critical node in the US pharmaceutical industry supply chain as it hosts 13 of top pharmaceutical companies and medical devices manufactures in the US (PharmaBoardroom, 2016). Therefore, disruptive events that strike the region could have a devastating impact on the supply of critical drugs, medical devices but more importantly, the life of patients. Hurricane Maria caused severe damage to the infrastructure of Puerto Rico. The Federal Drug Administration (FDA) issued alarming reports about the potential shortage of 40 critical medication and more than 50 manufacturing companies afflicted by damaged facilities (Kelly, 2017). Thus, understanding the factors contributing to building resiliency will enable organizations to be better prepared to face disruptive events.

The research on supply chain resiliency is at a nascent stage and requires further understanding of its enhancers (Blackhurst, Dunn, & Craighead, 2011). Previous studies (Martin

Christopher & Peck, 2004; Pettit, Croxton, & Fiksel, 2013; Pettit, Fiksel, & Croxton, 2010) have discussed various capabilities that will contribute to build and improve the resiliency of the supply chain based on the main four principles advanced by Christopher & Peck (2004) namely : re-engineering, collaboration, agility and culture of risk management. In their research paper (Scholten, Sharkey, Scott, & Fynes, 2014) designed an integrated framework based on four distinct phases (mitigation, preparedness, immediate response and recovery) wherein each phase is related to a set of capabilities to build supply chain resilience. Although, supply chain resiliency enhancers have been discussed from an information-based perspective, the interplay between the varying degree of uncertainty faced by the company and the capabilities it needs to develop is still unclear. In our paper, we intend to study how companies can deal with uncertainty from an organizational information processing theory (OIPT) to enhance resiliency by either increasing its access to information or by reducing its need for information. We set to contribute to the literature on supply chain resiliency, by studying the mechanisms through which the firm can balance its information needs and information capacity to mitigate disruptive events in a reactive and proactive manner. We empirically test the effects of Supply Chain Integration (SCI) and Supply Chain Risk Management practices (SCRM) on Supply Chain Resiliency (SCRE). Thus, we intend to answer the following research question: To what extent the interplay between SCI and SCRM enhance supply chain resiliency?

In line with extant literature (Sheffi & Rice, 2005; Wieland & Wallenburg, 2012) we conceptualize supply chain resiliency as consisting of two dimensions: Agility “Reactive Approach” and Robustness “Proactive Approach.” Robustness focuses on the ability of the organization to maintain operations in the face of disruptive events (Kitano, 2004), while Agility pertains to the ability of the organization to quickly react to changes by returning to a normal situation (Chowdhury & Quaddus, 2017; Wieland & Wallenburg, 2012). Considering the critical role played by agility and robustness in protecting operational and business performance of the organization (Hendricks & Singhal, 2005b, 2005a), we adopt the view of Brandon-Jones et al. (2014) by considering robustness and agility as performance outcomes.

Past studies have demonstrated the disastrous effect supply chain risks can have on firm performance (Hendricks & Singhal, 2005a). Since supply chain disruptions are a form of risk (Chopra & Sodhi, 2004), firms need to implement supply chain risk management (SCRM) practices to withstand its adverse effects. SCRM practices are the integrated process of identifying, analyzing accepting or mitigating the risks along the supply chain (Blome & Schoenherr, 2011). Previous studies argued that SCRM practices can effectively mitigate supply chain risks through the implementation of tactics and tools which seek to build redundancy in the supply chain, such as slack capacity, dual sourcing (Lavastre, Gunasekaran, & Spalanzani, 2012; Wiengarten, Humphreys, Gimenez, & McIvor, 2016).

Previous studies and meta-analysis demonstrated the benefits of supply chain integration in its external and internal form on the organization’s performance (Ataseven & Nair, 2017; Flynn, Huo, & Zhao, 2010; Schoenherr & Swink, 2012). Research on SCI focused mainly on studying the impact of integrating internally and externally on operational performance (Danese & Bortolotti, 2014; Flynn et al., 2010; Wong, Boon-Itt, & Wong, 2011), business performance (Chang, Ellinger, Kim, & Franke, 2016; Leuschner, Rogers, & Charvet, 2013; Swink, Narasimhan, & Wang, 2007). Although it is fundamental to investigate the impact of SCI on classical forms of firm performance, an organization can derive other benefits from SCI. Thus, by investigating the role of SCI in building supply chain resiliency, we argue that integrating with supply chain partners allow the firm to improve its resiliency and protect its value from disruptive events.

In this paper, we draw on the organizational information processing theory (OIPT) (Galbraith, 1973, 1977) to understand the relationship between supply chain integration, supply chain risk management practices, and supply chain resiliency. According to the OIPT, when facing high levels of uncertainty, organizations need greater levels of information (Galbraith, 1973). Thus, we posit that SCI and SCRM are instrumental in reducing the level of uncertainty to improve the supply chain resiliency. By implementing these supply chain practices, firms can balance their information needs and capacity to enhance their robustness and agility. We build on literature from SCI and supply chain risk management to develop our hypothesis.

We contribute to the literature of supply chain resiliency, by studying supplier and customer integration impact as enhancers of robustness and agility. Based on the OIPT, we view the development of resiliency as an information-based process, where information processing needs and information processing capacity act as levers to reduce the high uncertainty characterizing the identification, processing, and mitigation of disruptive risks. In the following sections, we will review the literature on our main constructs and develop the hypothesis for our research model. Then, we will discuss the methodology and results found after testing our hypotheses. Finally, we will discuss the results and suggest theoretical and managerial implications for our paper.

THEORETICAL FOUNDATIONS

Organizational processing theory (OIPT): a link to resiliency enhancers

According to OIPT, when facing high uncertainty stemming from their business environment, organizations need to deploy the adequate processing needs to improve their organizational performance (Tushman & Nadler, 1978). Therefore, organizations need to collect, transform and interpret the information to reduce uncertainty (Daft & Lengel, 1986). For Galbraith (1974), organizations can manage low uncertainty contexts through a coordination mechanism by utilizing rules, hierarchy, targets, and goals to overcome exceptional situations. However, when the level of uncertainty increases such a “mechanistic approach” is no longer feasible. Galbraith (1974) suggests that organizations can adjust by either decreasing its information processing needs by creating slack or buffers or by increasing its information processing capacity through lateral relations and investing in vertical information systems. For instance, the implementation of some SSCRM practices aims to mitigate disruption risks by positioning safety stock in critical nodes of the supply chain, while firms can create lateral relations through external integration with customers and suppliers.

Disruptive events exert great uncertainty on organizations, thus impeding their ability to develop the adequate measures to withstand disruptions. In addition to this, the occurrence of a disruptive event necessitates collaboration and information exchange with supply chain partners in order to withstand its devastating effects. Organizations seeking supply chain integration will increase their information processing capacity. While implementing SSCRM practices is going to reduce the need for information processing needs of the organization. The joint action of both SCI and SCRM balances the information needs and capacity a firm requires to successfully enhance its resiliency. Thus, OIPT is a suitable theoretical lens to investigate the relationship between SCI, SCRM, and Supply Chain resiliency.

Supply Chain Integration:

Supply chain integration refers to the extent to which a strategic collaboration and the intra- and inter-organization processes are managed between an organization and its partners, to process the flows of products & services, information, funds and decisions effectively and efficiently, to maximize the customer value ultimately (Flynn et al., 2010; Frohlich & Westbrook, 2001). Prior studies differentiated between internal integration and external integration (Flynn et al., 2010; Leuschner et al., 2013; Wong et al., 2011). Internal integration focuses on the in-house processes and aims to synchronize and connect the intra-organizational processes by overcoming the existing functional silos (Kahn & Mentzer, 1996). External integration is oriented towards external partners of the firm namely suppliers and customers. It refers to the extent to which a firm collaborates with external partners through inter-organizational practices to synchronize and streamline processes and improve the supply chain performance (Stank & Keller, 2001). Since we are investigating how the need for information – by increasing the capacity to capture information or reduce the need for it - is crucial to manage disruptive events, we will focus only on external integration as a practice that creates external linkages with supply chain partners in order to collaborate and share information.

Building resiliency in the supply chain requires the collaboration of the organization with its upstream and downstream partners. Based on the OIPT perspective, overcoming exceptional scenarios characterized by high uncertainty such as supply chain disruptions necessitates the investment in information processing needs. To this end, firms can increase their information processing capacity via lateral relations and vertical information systems (Galbraith, 1973). We consider SCI as information exchange relationship involving extensive information gathering, sharing and interpreting. Although high levels of information sharing might lead to detrimental effects on performance (Flynn, Koufteros, & Lu, 2016), it is offset by the instrumental role SCI plays in the development of shared meanings, shared supply chain experiences (Hult, Ketchen, & Slater, 2004) and common understanding (Huber, 1991). As a result, supply chain partners can operate effectively as an integrated network and make sense more easily of greater levels of information. Therefore, we use the OIPT as a lens to argue that organizations need to implement supply chain integration to build resiliency in the supply chain.

Supply Chain Risk Management:

With the extended boundaries of the focal firm and the close inter-organizational relationships, various types of risks may impede supply chain from meeting its customer demand. In our paper, we orient our study specifically to supply chain risks. They can be defined as the probability and impact of different events capable of distorting and causing failures across the organizational levels (operational, tactical and strategic) along with any part of a supply chain (Ho, Zheng, Yildiz, & Talluri, 2015). Supply chain disruption is identified by (Chopra & Sodhi, 2004) as one of the primary sources of Supply chain risks. The occurrence of supply chain risks is associated with uncertainty because of the unexpected nature of risks that may materialize and hinder supply chain activities (Waters, 2007). Based on OIPT framework, we argue that organizations facing a high level of uncertainty will be able to build resiliency by increasing their information processing capacity through lateral relationships such as supply chain integration and by decreasing their information processing needs through the creation of redundancy via the implementation of SCRM practices.

Supply chain risk management is a collaborative approach that involves the external partners (suppliers and customers) in identifying, analyzing and mitigating the risks and uncertainties along the supply chain (Blome & Schoenherr, 2011; Ho et al., 2015; Manuj & Mentzer, 2008). At first, an organization should identify the risks and their drivers, which can hinder the operations of the supply chain. Then, it will analyze the information at its disposal and implement the adequate mitigation processes (Craighead, Blackhurst, Rungtusanatham & Handfield, 2007; Hendricks & Singhal, 2003; Tomlin, 2006). The processes of risk identification aim to scan the environment at an intra-organizational and inter-organizational level for potential risks (Kern, Moser, Hartmann, & Moder, 2012). Being able to precisely and quickly identify the supply chain risk, gives the organization the opportunity to assess its severity and engage the suitable mitigation strategies to limit or avoid its impact on the organizational performance (Blackhurst et al., 2011; Tomlin, 2006; Zsidisin, Ellram, Carter, & Cavinato, 2004). Disruptive events exert high uncertainty on supply chain decisions that requires greater information processing needs. Thus, gathering and interpreting accurate information along the supply chain is critical for the organization to withstand the effects of disruptions. In the hypotheses section, we argue that instead of increasing its information processing capacity, firms can use SCRM practices to create redundancy in the supply chain. This latter will act as an absorber of uncertainty and will allow the firm to demonstrate robustness and agility in case of a disruption.

Supply Chain Resiliency: a proactive and reactive approach

Given the importance of disruption effects on organizational performance, the literature on supply chain resiliency (SCRE) is growing, and multiple definitions of the SCRE are proposed (Ambulkar, Blackhurst, & Grawe, 2015; Bhamra, Dani, & Burnard, 2011; Ponomarov & Holcomb, 2009). For instance, Brandon-Jones et al. (2014) makes a distinction between resilience and robustness and consider them as two distinct concepts. The first is related to the ability of the firm to recover rapidly from a disruption, while the second concept pertains to the ability of the firm to sustain its operations in the face of disruption. Other researchers conceptualize resiliency as a mix of both proactive (robust) and reactive (agile) strategies (Chowdhury & Quaddus, 2017; Hohenstein, Feisel, Hartmann, & Giunipero, 2015; Melnyk, Closs, Griffis, Zobel, & Macdonald, 2014; Wieland & Wallenburg, 2012). In this paper, we adopt this view and consider an organization to be resilient if it is capable of maintaining its operations functioning while adjusting its supply chain rapidly to the new situation to face a disruptive event.

Robustness involves an ex-ante preparation for predictable and unpredictable disruptions. Researchers refer to supply chain robustness as “proactive strategy” (Durach, Wieland, & Machuca, 2015; Klibi, Martel, & Guitouni, 2010; Wieland & Wallenburg, 2012). An organization achieves robustness if it can maintain the continuity of the operations against disruptions (Brandon-Jones et al., 2014). To prepare for disruptive events, firms implement preventive measures across the supply chain in order to be able to operate despite the occurrence of disruption by using safety stock or backup options for instance (Azadegan, Patel, Zangoueinezhad, & Linderman, 2013; Zsidisin & Wagner, 2010). However, developing these “second options” should be done ex-ante via early detection of possible sources of disruptions to enable the organization to withstand disruptive events.

Conversely, to the capability of enduring disruption (robustness), agility is concerned with the reacting capability to disruptive events on an ex-post basis (Husdal, 2010: 14). According to (Wieland & Wallenburg, 2012), supply chain agility refers to the ability of an organization to react rapidly to distortions in the supply chain by readjusting its original configuration. An organization will be identified as agile, if it can respond promptly and recover from disruption with minimal impact (Pettit et al., 2013; Sheffi & Rice, 2005). As supply chain agility is observed after the occurrence of a disruption, the speed at which an organization reacts becomes a critical capability to face disruptions effectively.

HYPOTHESES DEVELOPMENT:

In our research model (*Figure 1*), we look at the direct effects of supply chain integration and supply chain risk management practices on supply chain resiliency in its proactive (robustness) and reactive (agility) form.

Supply Chain Integration Effect on Supply Chain Resiliency:

Facing disruptive events exerts a high level of uncertainty on managers to take the adequate decisions. OIPT suggests to firms in such cases firms to deal with increased uncertainty either by increasing their information processing capacity through lateral relations (SCI), or by creating buffers and slack (Galbraith, 1973). Respectively, engaging in supply chain integration activities with the upstream and downstream partners will allow the firm to leverage its inter-organizational linkages to reduce the uncertainty(Koufteros, 2014; Wong, Wong, & Boon-itt, 2015). The alignment of supply chain processes characterizing SCI will enable the firm to access new valuable information, interpret it and make informed decisions to prepare for and withstand supply chain disruptions.

Based on OIPT, closer integration with customers and suppliers enables the firm to gather, transform and exploit the information collected thanks to its external linkages (Galbraith, 1974; Srinivasan & Swink, 2015). Previous studies have demonstrated that tighter integration with supply chain partners allows the firm to access more accurate and high-quality information (Barratt & Oke, 2007).The development of information sharing among the firm and its suppliers and customers is instrumental in enhancing the supply chain visibility (M. Christopher & Lee, 2004). Armed with an enhanced visibility and established routines of information sharing, the firm can detect possible sources of disruptions in its supply chain and build the adequate measures to face it in a proactive manner. For example, thanks to the sharing of production plans among partners, the firm can detect possible bottlenecks and work with partners to alleviate them to withstand disruptive events, hence demonstrating robustness in face of a disruptive event. Therefore, we suggest the following hypotheses:

H_{1a}: Supplier integration is positively related to robustness

H_{1b}: Customer integration is positively related to robustness

Moving from an arm's length relationship to embedded relationships such as SCI, entails the development of trust and commitment (D. Q. Chen, Preston, & Xia, 2013; Zhang & Huo, 2013). These relational antecedents have been shown to play an essential role in reinforcing of the willingness of the firms to share risks and rewards amongst each other (I. J. Chen & Paulraj, 2004; Mayer, Davis, & Schoorman, 1995; Sambasivan & Yen, 2010). Furthermore, integration with suppliers and customers enable the supply chain partners to develop shared goals, understanding, codified routines, policies and procedures (Rothaermel & Deeds, 2006). Subsequently, by integrating with its suppliers and customers, the firm will be able to share, interpret and act on the available information more quickly since the supply chain partners already have developed routines and procedures to operate their processes. Thus, they will be able to react more quickly to disruptive events since they can process and treat information more quickly than others. Indeed, Liang & Huang (2006) have found that higher levels of external integration allow the firm to improve its agility and responsiveness to market uncertainty. Based on OIPT, closer integration with customer and supplier allow the firm to increase its information processing capacity to gather and exploit data more effectively to improve its agility. As a result, we propose the following hypotheses:

H_{2a}: Supplier integration is positively related to agility

H_{2b}: Customer integration is positively related to agility

Supply chain risk management effect on supply chain resiliency:

From an OIPT perspective, the second lever managers can activate to deal with the uncertainty implied by disruptive events is the creation of redundancy/slack (Galbraith, 1973, 1974). Consistent with this line of reasoning, we suggest that firms can deal with the uncertainty created by supply chain disruptions by implementing SCRM practices. In fact, the different SCRM processes are designed in a way to successfully identify potential sources of risk and apply adequate mitigation measures. To mitigate potential supply chain risks, firms build redundancy in the form of safety stocks, the dual or multiple sourcing, maintaining capacity slack (Lavastre et al., 2012). Implementing these practices enables the firm to lessen the effect the uncertainty by creating “shock absorbers” (Bode, Wagner, Petersen, & Ellram, 2011). However, to effectively mitigate the disruptive risks, the firm needs to identify these risks successfully. Indeed, Kern (2012) empirically demonstrated the positive relationship between risk identification and the success of the subsequent SCRM processes and the quality of the mitigation.

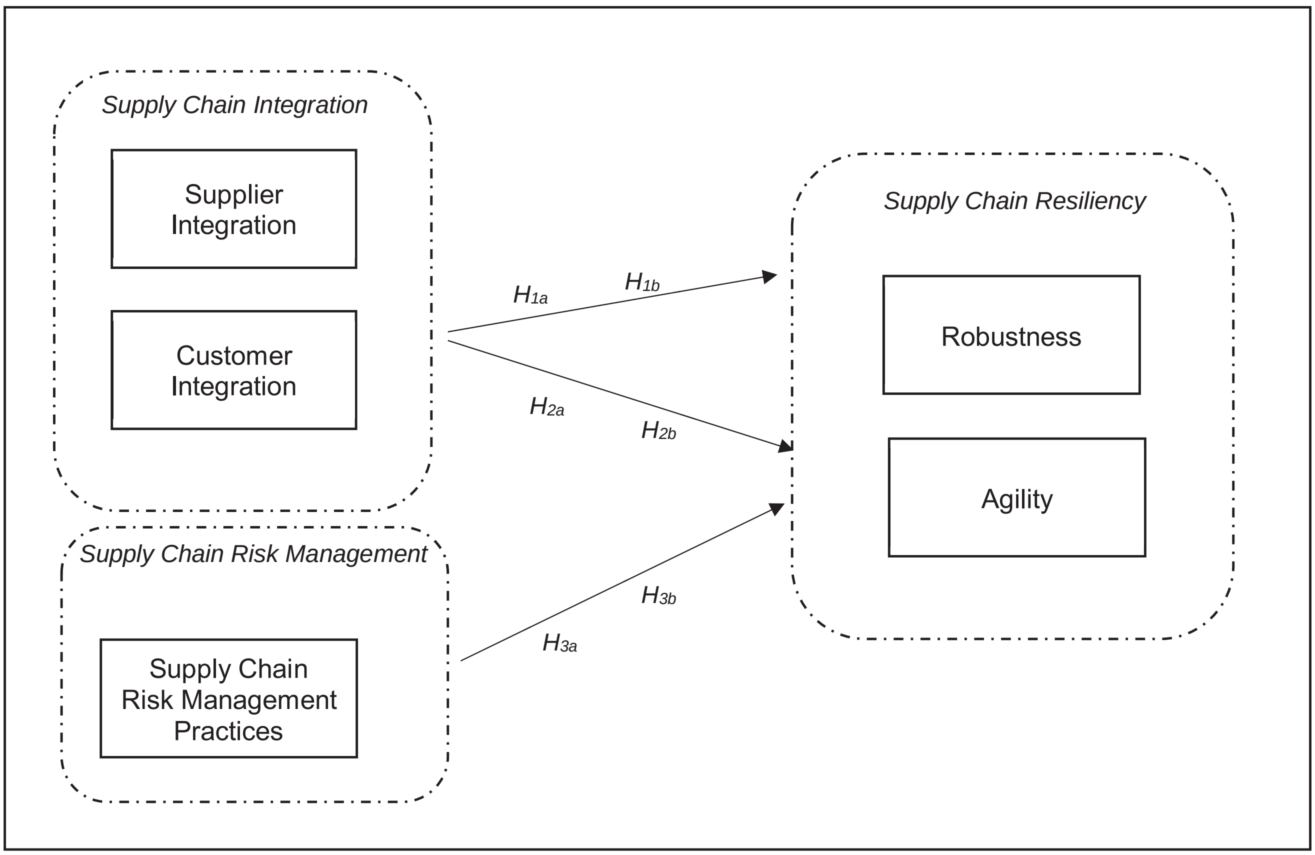
Risk identification is a critical part of the SCRM process (Blome & Schoenherr, 2011; Manuj & Mentzer, 2008). Developing risk identification capabilities will allow the firm to detect possible sources of disruptions and address them in a proactive manner. Risk identification is defined as the ability of the firm to systematically and exhaustively evaluate and recognize potential supply chain related risks (Tummala & Schoenherr, 2011). If this latter is executed correctly, it will allow the firm to effectively create and allocate the slack in its different forms to manage the uncertainty accompanying SC disruptions. As a result, the firm can prepare ex-ante for the disruptive events. For instance, after scanning its environment, the firm detected a vulnerable supplier that might be easily affected by the hazardous event. Based on the assessment of the collected information from, the firm can proceed to ask this supplier to build a safety stock to be able to withstand disruptive events. Thus, implementing SCRM is instrumental in developing robustness in a proactive manner to face SC disruptions. Indeed, practices such as early warning systems and supplier development programs (Azadegan, Kach, Golara, & Mousavi, 2017; Wiengarten et al., 2016) are directly related to developing robustness in response to disruptive events.

Further, the time necessary to react to a disruptive event is a major factor in reducing its impact on the supply chain (Craighead et al., 2007). Risk mitigation techniques (i.e., rethinking and restructuring supply and distribution strategy) (Wiengarten et al., 2016) endow the firm with the ability to promptly redesign and adapt its supply chain in the case of a disruption. Previous studies have stressed the importance of considering the creation of redundancy/slack in the supply chain as a way of mitigating supply chain disruptions (Sarathy, 2006; Sheffi & Rice, 2005). For example, after identifying a distribution-center as weather-sensitive, the firm can build a dual distribution-center that can be utilized in response to disruption. Therefore, by implementing this mitigation technique, the firm will be able to promptly reroute its distribution strategy and recover quickly. From an OIPT perspective, the slack resources used to mitigate the risks absorb the uncertainty created by the disruptive event and allow the firm to recover rapidly from a disruption. Previous studies have demonstrated the relationship between risk mitigation and recovery speed (Blackhurst et al., 2011; Craighead et al., 2007). Based on these arguments, we suggest the following hypotheses:

H_{3a}: Supply chain risk management practices are positively related to robustness

H_{3b}: Supply chain risk management practices are positively related to agility

Figure 1: Research Model



RESEARCH METHODOLOGY:

Data Collection and Sampling:

To empirically test our research model about enhancers of supply chain resiliency, we collected data through an online based survey of 2042 supply chain managers and executives from France and the United States. The questionnaire was developed for a single respondent using the organization level as a unit of analysis. To enhance the generalizability of our results we collected the data from two countries and across various industries. All the measurement instrument used in our questionnaire were adopted from past validated studies. For the French sample, the questionnaire was translated from English to the French language by a professional translator, then resulting version in French was back-translated into English by another bi-lingual researcher. The two versions were compared and checked for discrepancies (Flynn et al., 2010). We submitted the questionnaire to academicians and then to supply chain managers/executives for review. After receiving their comments and reviews, we integrated them to improve the clarity of the questionnaire. Then, we used this latest version to pre-test our questionnaire with 15 supply chain professionals belonging to different industrial sectors. We included the feedback from the supply chain managers to draw the final version questionnaire. Following a key informant approach in our study, we decided to collect data from top managers and executives working in supply chain management activities such as operations, purchasing and supply chain management. These individuals are knowledgeable about their organizations' activities and can readily access sensitive information (Liu, Wei, Ke, Wei, & Hua, 2016; Montabon, Daugherty, & Chen, 2018). Following the key informant approach, we carefully screened the responses and eliminated those whose titles are not reflecting a supply chain related activity or didn't have the required experience inside the company. For the US data, 40% of respondents are supply chain directors and VPs, and 17.8% occupy a Manager or VP position in Supply chain related activities (Materials handling and purchasing). Also, 65.8% are highly experienced professionals in supply chain activities. For the French sample, 58% of the respondents work as Supply Chain Directors and VP's, and 12.7% work in supply chain related activity either as Managers' or VP's. Furthermore, in the US sample, 65.8% of the respondents have at least nine years of experience in the supply chain field. This figure is about 48.5% on the French sample.

To collect data from France, we collaborated with "Association Française pour la Logistique" (ASLOG), the largest association of supply chain management professionals in the country. The organization has more than 1,200 supply chain professionals working across various industrial sectors. The president of the organization signed and sent an email to invite all its members to participate in the study and explaining the objectives of the study and its importance to advance supply chain. Plus, a summary report of the results was promised to the participants to the study. Then, an email that contains the web-link to the questionnaire was sent to the 1200 members of the organization of supply chain management professionals. After two waves of emails and reminders through the social networks of the association, we received 133 usable responses, corresponding to an 11.25% response rate. To collect data in the US, we have commissioned Qualtrics, a third-party online survey administration company to collect data from US supply chain managers and executives. Qualtrics Panels has been employed in previous studies and is identified as a reliable instrument to collect data (Courtright, Gardner, Smith, & McCormick, 2016; Hazen, Bradley, Bell, In, & Byrd, 2017). Qualtrics Panel services sent an email containing the online survey link to a pool of 842 participants. 152 completed questionnaires were recorded corresponding to a response rate of 18.05%.

Measures:

We surveyed the existing literature to identify validated and reliable measures for our study. All measurements scales used in our questionnaire were adopted from existing literature (Malhotra & Grover, 1998).
Supply Chain Integration: We measured supply chain integration in its supplier and customer dimension using scales adopted from (Flynn et al., 2010; Frohlich & Westbrook, 2001). The items used, measured the level of coordination with key suppliers and customers of the flow of information and planning decisions.
Supply chain Robustness: we measured supply chain robustness using a scale adopted from (Brandon-Jones et al., 2014). This scale measures the ability of the firm to keep performing its operations normally in case a disruption occurs. It assesses the extent to which the firm will be able to meet customer demand and achieve the expected level of performance despite the disruption.
Supply chain agility: we measured supply chain agility using a scale adopted from (Wieland & Wallenburg, 2012). It examines the speed at which the firm can modify its operations in case of unexpected changes. It estimates the ability of the firm to rapidly adapt activities such as manufacturing lead-times or delivery reliability to face changes.
Supply chain risk practices: To measure supply chain risk management practices we adopted a scale from (Wiengarten et al., 2016). It evaluates the effort level devoted to implementing programs to identify and mitigate risks, such as vendor monitoring and rating programs and early warning systems and contingency programs (Blome & Schoenherr, 2011; Manuj & Mentzer, 2008).
Control Variables: To enhance the robustness of our results, we included relevant variables as control variables to our research model. We controlled for firm size by measuring the total number of employees, since small-sized companies are less exposed to disruptions (Wagner & Neshat, 2012). Second, the level of SCRM and SCI practices implementation may differ across industrial sectors in terms of exposure to disruptions (Min & Galle, 2001). Thus, we decided to level out the effect of the industry on the results. Third, we included the country as a dummy variable to account for possible differences among the two countries (US and France). Finally, the relationship duration between the firm and its upstream and downstream partners may affect the results since long established firms enjoy more information sharing routines and procedures to improve their performance to withstand disruptive events (Kotabe, Martin, & Domoto, 2003).

Measurement Model:

The analysis of the confirmatory factor analysis 'CFA' results (see Table 1) indicates that the data effectively fit the model: $(\mathcal{X}^2 = \mathbf{401.58}\ (\boldsymbol{df} = \mathbf{178}, \mathcal{X}^2/\boldsymbol{df} = \mathbf{2.256}\ )$; $\boldsymbol{CFI} =\mathbf{.94}$; $\boldsymbol{TLI} = \mathbf{.93}$; $\boldsymbol{RSMEA} =\mathbf{.06})$, thus ensuring strong model fit. Additionally, we assessed the convergent validity of the constructs by analyzing the factors loading, average variance extracted (AVE) and composite reliability (CR). The analysis results show that all factor loadings are greater than 0.5 and significant at ($p < .001$) (Hair, Anderson, Black, & Babin, 2010), thus we can establish high convergence of our measurement instruments. The CR values of all constructs used in our model was found to be greater than 0.7 (Hair et al., 2010), thus ensuring internal consistency and convergent validity. Also, the AVE for each measure was greater than 0.5, hence convergent validity is supported. Furthermore, the AVE results exceeded the squared correlations of the remaining measures (Hair et al., 2010), indicating the support of the discriminant validity. Since all our scales exceed the recommended thresholds for each of the tests, it is safe to say that all our measurements have good reliability, convergent and discriminant validity.

Table 1- Confirmatory Factor Analysis

Construct[a]	Loading
Robustness (α = .89; CR = .90; AVE = .66)	
To what extent do these statements apply to your supply chain should disruptions of operations occur (1= Not at all; 7= To a very great extent):	
ROB1 Operations would be able to continue	,780
ROB2 We would still be able to meet customer demand	,836
ROB3 Performance would not deviate significantly from targets	,829
ROB4 The supply chain would still be able to carry out its regular functions	,820
Agility (α = .85; CR = .86; AVE = .62)	
Please indicate the speed of reaction with which your company can engage in the following activities should changes occur (1= Slow; 7= Fast):	
AGI1 Adapt manufacturing lead-times	,692
AGI2 Adapt level of customer service	,804
AGI3 Adapt delivery reliability	,762
AGI4 Adapt responsiveness to changing market needs	,854
Supply chain risk management practices (α = .89; CR = .89; AVE = .66)	
What level of effort did your company invest in action programs within the previous 3 years (1= None, 7 = High):	
SSCRM1 Rethinking and restructuring supply strategy and the organization and management of supplier portfolio	,811
SSCRM2 Implementing supplier development and vendor rating programs	,820
SSCRM3 Rethinking and restructuring distribution strategy **in order to** change the level of intermediation	,839
SSCRM4 Implementing practices including early warning system, effective contingency programs for possible supply chain disruptions	,785
Supplier Integration (α = .91; CR = .91; AVE = .68)	
Please indicate the extent of integration or information sharing between your organization and your major supplier in the following areas (1 = Not at all; 7 = Extensive):	
SI1 Our major supplier shares their production schedule with us	,840
SI2 Our major supplier shares their production capacity with us	,897
SI3 Our **major** supplier shares available inventory with us	,809
SI4 We share our production plans with our **major** supplier	,804
SI5 We share our inventory levels with our **major** supplier	,735
Customer Integration (α = .84; CR = .84; AVE = .57)	
Please indicate the extent of integration or information sharing between your organization and your major customer in the following areas (1 = Not at all; 7 = Extensive):	
CI1 Our major customer shares Point of Sales (POS) information with us	,756
CI2 Our major customer shares demand forecast with us	,771
CI3 We share our available inventory with our major customer	,686
CI4 We share our production plan with our major customer	,716
The first item in each scale was fixed to a loading of 1.0 in the initial run to set the scale of the construct. Observed CFA fit statistics were: X2(401) = 178.25; TLI = .933; incremental fit index = .944; comparative fit index = .943; root mean square error of approximation = .06	

ANALYSES:

Robustness and agility are two closely related concepts that form the supply chain resiliency concept (Wieland & Wallenburg, 2012, 2013). Thus, error terms in the regression equations may correlate. To overcome this issue, we deem the use of Seemingly Unrelated Regression (SUR) as a suitable method to test the hypotheses of our research model. SUR is adapted to our regression equations because it accounts for the correlations among the error terms (Habermann, Blackhurst, & Metcalf, 2015). Indeed, the SUR equation modeling is suitable to generate robust regression estimates when a set of equations includes identical independent variables (Habermann et al., 2015).

Table 2a: SUR Results for Robustness

Seemingly Unrelated Regression : DV1 Robustness

	Control Model		***Full model***	
Variables	*B*	S.E	*B*	S.E
Controls				
Firm Size	-.03	(.05)	-.08*	(.04)
Supplier relationship duration	.15*	(.07)	.11*	(.06)
Country dummy	.86***	(.18)	.19	(.16)
Customer relationship duration	.01	(.07)	.02	(.06)
Industry-type[a]				
Main effects				
Supply chain risk management			.33***	(.05)
Supplier Integration			.14**	(.05)
Customer Integration			.18***	(.05)
Model Summary				
R^2	.20		.48	
ΔR^2			.28	

Sample size for all models, N = 283.
[a]Industry type was included as a dummy variable, but for the sake of brevity, it was not included in the table.
Results are presented as coefficients and standard errors in (parentheses).
* *p* < .05,
** *p* < .01,
*** *p* < .000.

To assess multicollinearity, we used traditional OLS regression to estimate the variance inflation factors (VIF) for regression coefficients among our independent variables (Habermann et al., 2015). The VIF values vary from 1.38 to 2.38, indicating that our results are below the recommended cut-offs for multicollinearity problems (Hair, Black, Babin, & Anderson, 2006).

The results of the seemingly unrelated regression are presented in the table 2a and table 2b. The first group of hypotheses proposed a positive direct effect of supplier integration (H1a) and customer integration (H1b) on robustness and a similar positive direct effect between supplier integration (H2a) and customer integration (H2b) on agility. The results of the SUR provided full support for this set of hypotheses H1a, H1b, H2a and H2b, thus demonstrating that SCI positively and significantly enhance supply chain resiliency in its robustness and agility dimension. The second group of hypotheses suggested that SCRM is positively related to robustness (H3a) and agility (H3b). The results indicate that SCRM is positively and significantly related to robustness and agility. Hence both our hypotheses are supported.

Hypotheses 1 proposed a positive relationship between SCI and robustness. The relationship was positive and significant for the supplier integration $(\boldsymbol{\beta = .14},\ \boldsymbol{\rho < .01})$ and customer integration $(\boldsymbol{\beta = .18},\ \boldsymbol{\rho < .00})$, hence providing full support for hypotheses H1a and H1b. Also, our results supported the hypotheses 2, where we suggested that SCI has a positive and significant effect on agility both for supplier integration hypothesis H2a $(\boldsymbol{\beta = .12},\ \boldsymbol{\rho < .05})$ and customer integration Hypothesis H2b $(\boldsymbol{\beta = .18},\ \boldsymbol{\rho < .00})$. Finally, the results confirm the hypotheses H3a and H3b and suggests that SCRM is positively and significantly enhancing robustness H3a $(\boldsymbol{\beta = .33},\ \boldsymbol{\rho < .00})$ and agility $(\boldsymbol{\beta = .26},\ \boldsymbol{\rho < .00})$.

Table 2b: SUR Results for Agility

Seemingly Unrelated Regression : DV2 Agility

	Control Model		***Full model***	
Variables	**B**	**S.E**	**B**	**S.E**
Controls				
Firm Size	-.10*	(.05)	-.14***	(.04)
Supplier relationship duration	.14*	(.07)	.10	(.05)
Country (dummy variable)	.55*	(.17)	-.04	(.16)
Customer relationship duration	-.06	(.06)	-.05	(.05)
Industry-Type[a]				
Main effects				
Supply chain risk management			.26***	(.05)
Supplier Integration			.12*	(.05)
Customer Integration			.18***	(.05)
Model Summary				
R^2	.17		.41	
ΔR^2			.24	

[a]Industry type was included as a dummy variable, but for the sake of brevity, it was not included in the table.
Results are presented as coefficients and standard errors in (parentheses).
* *p* < .05,
** *p* < .01,
*** *p* < .000.

DISCUSSION:

By studying supply chain resiliency from an information-based view, we suggest that firms can enhance their resiliency by balancing their information needs and capacity. As a result, we complement the existing findings enhancers of resiliency (Pettit et al., 2010; Scholten et al., 2014) by nuancing the supply chain practices that company can build on to improve its robustness and agility. Hence, managers can vary the levels of information capacity and information needs according to the set of resources they have in order to enhance the resiliency of their supply chain. Our findings identify supply chain integration and supply chain risk management practices as two important factors that contribute to enhancing the robustness and agility of the firm.
The findings of this study offer a threefold contribution to supply chain management and OIPT literature. First, past studies on SCI focused on investigating its effects on various forms of firm performance such as operational performance (Flynn et al., 2010; Swink et al., 2007), financial performance (Chang et al., 2016; Vickery, Jayaram, Droge, & Calantone, 2003), and business performance (Cao & Zhang, 2011; Flynn et al., 2010). Yet, the study of other types of SCI benefits is still scarce. The results of this paper demonstrate the role SCI plays in enhancing supply chain resiliency in its two dimensions (robustness and agility). We considered SCI as an information exchange relationship that involves repeated interactions between the firm and its supply chain partners to collect and interpret data. In accordance with OIPT principles (Galbraith, 1977), SCI enables the firm to cope with high uncertainty faced to endure and recover from a disruptive event. In fact, SCI acts as a mechanism to increase the information processing capacity of the firm. By closely integrating with suppliers and customers, the firm can enhance its visibility of the supply chain to prepare beforehand for disruptive events. Moreover, SCI allows the firm to leverage the established routines and procedures to quickly make sense out collected information and react quickly to disruptive events.

Second, by considering SCRM practices as a mechanism to reduce the information processing needs of the firm, we clarified how these practices could contribute to supply chain resiliency. Consistent with OIPT principles, we explained that the practices of SCRM allow the firm to build the necessary redundancy in the supply chain to absorb the possible shocks of disruptive events. Thanks to the practices aiming to identify sources of risks, the firm can detect and position more effectively the buffers ex-ante to disruptions. Whilst, mitigation practices will allow the firm to react promptly to disruptions by capitalizing on the created slacks (safety stock, dual sourcing,etc.).

Third, by viewing the enhancement of supply chain resiliency as an information-based process, we deemed suitable to apply the OIPT to explain how the combination of SCI and SCRM practices can allow the firm to cope with the high uncertainty involved in building robustness and agility to deal with SC disruptions. By studying these two factors, we answer the call of researchers to investigate enhancers of supply chain resiliency (Blackhurst et al., 2011). Consistent with OIPT principles, we suggest that managers can cope with the uncertainty implied by disruptive events either by increasing their information processing capacity through SCI or by decreasing their needs for information processing through the implementation of SCRM practices that seeks to build redundancy in the supply chain. Therefore, we contribute to OIPT by offering two mechanisms that can be used to manage or lessen the uncertainty in order to improve supply chain resiliency.

CONCLUSIONS AND LIMITATIONS

Our study extends and clarifies prior research on supply chain resiliency by investigating two mechanisms through the firm can enhance its robustness and agility. Specifically, we study how supply chain integration and supply chain risk management practices enhance the resiliency of the firm to disruptive events. By applying OIPT principles, we identified two mechanisms through which the firm can cope with uncertainty. Our results demonstrate that SCI allows the firm to increase its information processing capacity by leveraging the lateral benefits relations with supplier and customers can confer to the firm. Whilst implementing SCRM practices enable the firm to create the required redundancy to reduce its information processing needs. Combining these two mechanisms allow the firm to enhance robustness and agility thanks to a reduction of the uncertainty created by disruptive events.

The findings of our study offer various practical implications. First, by differentiating between robustness and agility, we aim to help managers understand that each dimension of resiliency requires different capabilities to be enhanced. Second, by investigating SCI and SCRM practices, we offer managers a solution to leverage supply chain practices to enhance their resiliency. Firms can act on the two levers (SCI and SCRM) to increase their information processing capacity or decrease their information processing needs depending on their situations.

Along with the results of this study, you should read it considering its limitations. First, cross-sectional data was used to test the hypotheses of this research paper. As a result, we cannot draw causal relationships between our indigenous and exogenous variables. Thus, future research can attempt to replicate our study to add on its validation or use longitudinal data to analyze the effect of SCI and SCRM practices on agility and robustness over time. Second, we limited our model to two possible enhancers of supply chain resiliency, but other unexplored factors might contribute to supply chain resiliency. We invite future research to explore other factors that may improve robustness and agility.

REFERENCES:

Ambulkar, S., Blackhurst, J., & Grawe, S. (2015). Firm's resilience to supply chain disruptions: Scale development and empirical examination. *Journal of Operations Management*, *33–34*, 111–122.

Ataseven, C., & Nair, A. (2017). Assessment of supply chain integration and performance relationships: A meta-analytic investigation of the literature. *International Journal of Production Economics*, *185*, 252–265.

Azadegan, A., Kach, A., Golara, S., & Mousavi, N. (2017). Corporate environmental investments: A cross-national study on managerial decision making. *International Journal of Production Economics*.

Azadegan, A., Patel, P. C., Zangoueinezhad, A., & Linderman, K. (2013). The effect of environmental complexity and environmental dynamism on lean practices. *Journal of Operations Management*, *31*(4), 193–212.

Barratt, M., & Oke, A. (2007). Antecedents of supply chain visibility in retail supply chains: a resource-based theory perspective. *Journal of Operations Management*, *25*(6), 1217–1233.

Bhamra, R., Dani, S., & Burnard, K. (2011). Resilience: The concept, a literature review and future directions. *International Journal of Production Research*, *49*(18), 5375–5393.

Blackhurst, J., Dunn, K. S., & Craighead, C. W. (2011). An Empirically Derived Framework of Global Supply Resiliency. *Journal of Business Logistics*, *32*(4), 374–391.

Blome, C., & Schoenherr, T. (2011). Supply chain risk management in financial crises - A multiple case-study approach. *International Journal of Production Economics*, *134*(1), 43–57.

Bode, C., Wagner, S. M., Petersen, K. J., & Ellram, L. M. (2011). Understanding Responses to Supply Chain Disruptions : Insights from Information Processing and Resource Dependence Perspectives. *Academy of Management Journal*, *54*(4), 833–856.

Brandon-Jones, E., Squire, B., Autry, C. W., & Petersen, K. J. (2014). A Contingent Resource-Based Perspective of Supply Chain Resilience and Robustness. *Journal of Supply Chain Management*, *50*(3), 55–73.

Cao, M., & Zhang, Q. (2011). Supply chain collaboration: Impact on collaborative advantage and firm performance. *Journal of Operations Management*, *29*(3), 163–180.

Chang, W., Ellinger, A. E., Kim, K. K., & Franke, G. R. (2016). Supply chain integration and firm financial performance: A meta-analysis of positional advantage mediation and moderating factors. *European Management Journal*, *34*(3), 282–295.

Chen, D. Q., Preston, D. S., & Xia, W. (2013). Enhancing hospital supply chain performance: A relational view and empirical test. *Journal of Operations Management*, *31*(6), 391–408.

Chen, I. J., & Paulraj, A. (2004). Towards a theory of supply chain management: The constructs and measurements. *Journal of Operations Management*, *22*(2), 119–150.

Chopra, S., & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. *MIT Sloan Management Review*, *46*(1), 53–62.

Chowdhury, M. M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. *International Journal of Production Economics*, *188*(September 2015), 185–204.

Christopher, M., & Lee, H. (2004). Mitigating supply chain risk through improved confidence. *International Journal of Physical Distribution & Logistics Management*, *34*(5), 388–396.

Christopher, Martin, & Peck, H. (2004). Building the Resilient Supply Chain. *The International Journal of Logistics Management*, *15*(2), 1–14.

Courtright, S. H., Gardner, R., Smith, T. A., & McCormick, B. W. (2016). My Family Made Me Do It: a Cross-Domain, Self-Regulatory Perspective on Antecedents to Abusive supervision. *Academy of Management Journal*, *59*(5), 1630–1652.

Craighead, C. W., Blackhurst, J., Rungtusanatham, J. M., & Handfield, R. B. (2007). The Severity of Supply Chain Disruptions : Design Characteristics and Mitigation Capabilities. *Decision Sciences*, *38*(1), 131–156.

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, *32*(5), 554–571.

Danese, P., & Bortolotti, T. (2014). Supply chain integration patterns and operational performance: A plant-level survey-based analysis. *International Journal of Production Research*, *52*(23), 7062–7083.

Durach, C. F., Wieland, A., & Machuca, J. A. D. (2015). Antecedents and dimensions of supply chain robustness: a systematic literature review. *International Journal of Physical Distribution & Logistics Management*, *45*(1/2), 118–137.

Flynn, B. B., Huo, B., & Zhao, X. (2010). The impact of supply chain integration on performance: A contingency and configuration approach. *Journal of Operations Management*, *28*(1), 58–71.

Flynn, B. B., Koufteros, X., & Lu, G. (2016). On Theory in Supply Chain Uncertainty and its Implications for Supply Chain Integration. *Journal of Supply Chain Management*, *52*(3), 3–27.

Frohlich, M. T., & Westbrook, R. (2001). Arcs of integration: An international study of supply chain strategies. *Journal of Operations Management*, *19*(2), 185–200.

Galbraith, J. R. (1973). *Designing Complex Organizations.* Reading, MA.: Addison-Wesley.

Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces*, *4*(3), 28–36.

Galbraith, J. R. (1977). *Organization design.* Reading, MA.: Addison-Wesley.

Habermann, M., Blackhurst, J., & Metcalf, A. Y. (2015). Keep Your Friends Close? Supply Chain Design and Disruption Risk. *Decision Sciences*, *46*(3), 491–526.

Hair, J. F., Anderson, R. E., Black, W., & Babin, B. J. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Hair, J. F., Black, W., Babin, B. J., & Anderson, R. E. (2006). *Multivariate data analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Hazen, B. T., Bradley, R. V., Bell, J. E., In, J., & Byrd, T. A. (2017). Enterprise architecture: A competence-based approach to achieving agility and firm performance. *International Journal of Production Economics*, *193*(July), 566–577.

Hendricks, K. B., & Singhal, V. R. (2003). The effect of supply chain glitches on shareholder wealth. *Journal of Operations Management*, *21*(5), 501–522.

Hendricks, K. B., & Singhal, V. R. (2005a). An Empirical Analysis of the Effect of Supply Chain Disruptions on Long-Run Stock Price Performance and Equity Risk of the Firm. *Production & Operations Management*, *14*(1), 35–52.

Hendricks, K. B., & Singhal, V. R. (2005b). Association between Supply Chain Glitches and Operating Performance. *Management Science*, *51*(5), 695–711.

Ho, W., Zheng, T., Yildiz, H., & Talluri, S. (2015). Supply chain risk management: A literature review. *International Journal of Production Research*, *53*(16), 5031–5069.

Hohenstein, N.-O., Feisel, E., Hartmann, E., & Giunipero, L. (2015). Research on the phenomenon of supply chain resilience: A systematic review and paths for further investigation. *International Journal of Physical Distribution & Logistics Management*, *45*(1/2), 90–117.

Huber, G. P. (1991). Organizational learning: The contributing processes and literature. *Organizational Science*, *2*(1), 88–115.

Hult, G. T., Ketchen, D. J., & Slater, S. F. (2004). Information processing, knowledge development and strategic supply chain performance. *Academy of Management Journal*, *47*(2), 241–253.

Husdal, J. (2010). A Conceptual Framework for Risk and Vulnerability in Virtual Enterprise Networks. In S. Ponis (Ed.), *Managing Risk in Virtual Enterprise Networks: Implementing Supply Chain Principles* (pp. 1–27).

Kahn, K. B., & Mentzer, J. T. (1996). Logistics and interdepartmental integration. *International Journal of Physical Distribution & Logistics Management*, *26*(8), 6–14.

Kelly, L. (2017). FDA: Medical saline bags in short supply in U.S. from Hurricane Maria in Puerto Rico - Washington Times. *The Washington Time*. Retrieved from https://www.washingtontimes.com/news/2017/oct/13/fda-medical-saline-bags-short-supply-us-hurricane-/

Kern, D., Moser, R., Hartmann, E., & Moder, M. (2012). Supply risk management: model development and empirical analysis. *International Journal of Physical Distribution & Logistics Management*, *42*(1), 60–82.

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, *5*(11), 826–837.

Klibi, W., Martel, A., & Guitouni, A. (2010). The design of robust value-creating supply chain networks: A critical review. *European Journal of Operational Research*, *203*(2), 283–293.

Kotabe, M., Martin, X., & Domoto, H. (2003). Gaining from vertical partnerships: Knowledge transfer, relationship duration, and supplier performance improvement in the U.S. and Japanese automotive industries. *Strategic Management Journal*, *24*(4), 293–316.

Koufteros, X. (2014). The Effect Of Performance Measurement Systems On Firm Performance: A Cross-Sectional And A Longitudinal Study. *Journal of Operations Management*.

Lavastre, O., Gunasekaran, A., & Spalanzani, A. (2012). Supply chain risk management in French companies. *Decision Support Systems*, *52*(4), 828–838.

Leuschner, R., Rogers, D. S., & Charvet, F. F. (2013). A Meta-Analysis of Supply Chain Integration and Firm Performance. *Journal of Supply Chain Management*, *49*(2), 34–57.

Lewin, L., Roegiers, B., Tuazon, B., & Almond, K. (2017). In pictures: Hurricane Maria pummels Puerto Rico. Retrieved December 20, 2017, from edition.cnn.com website: http://edition.cnn.com/interactive/2017/09/world/hurricane-maria-puerto-rico-cnnphotos/

Liang, W., & Huang, C. (2006). Agent-based demand forecast in multi-echelon supply chain. *Decision Support Systems*, *42*(1), 390–407.

Liu, H., Wei, S., Ke, W., Wei, K. K., & Hua, Z. (2016). The configuration between supply chain integration and information technology competency: A resource orchestration perspective. *Journal of Operations Management*, *44*, 13–29.

Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: from constructs to theory. *Journal of Operations Management*, *16*(4), 407–425.

Manuj, I., & Mentzer, J. T. (2008). Global Supply Chain Risk Management. *Journal of Business Logistics*, *29*(1), 133–155.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

Melnyk, S. A., Closs, D. J., Griffis, S. E., Zobel, C. W., & Macdonald, J. R. (2014). Understanding supply chain resilience. *Supply Chain Management Review*, *January/Fe*, 34–41.

Min, H., & Galle, W. P. (2001). Green purchasing practices of US firms. *International Journal of Operations & Production Management*, *21*(9), 1222–1238.

Montabon, F. L., Daugherty, P. J., & Chen, H. (2018). Setting Standards for Single Respondent Survey Design. *Journal of Supply Chain Management*, *54*, 35–41.

Pettit, T. J., Croxton, K. L., & Fiksel, J. (2013). Ensuring supply chain resilience: Development and implementation of an assessment tool. *Journal of Business Logistics*, *34*(1), 46–76.

Pettit, T. J., Fiksel, J., & Croxton, K. L. (2010). Ensuring Supply Chain Resilience: Development of a Conceptual Framework. *Journal of Business Logistics*, *31*(1), 1–21.

PharmaBoardroom. (2016). *Healthcare & Life Sciences Review: Puerto Rico*. London.
Ponomarov, Y., & Holcomb, M. C. (2009). Understanding the concept of supply chain resilience. *The International Journal of Logistics Management*, *20*(1), 124–143.
Rothaermel, F. T., & Deeds, D. L. (2006). Alliance type, alliance experience, and alliance management capability in high-technology ventures. *Journal of Business Venturing*, *21*(4), 429–460.
Sambasivan, M., & Yen, C. N. (2010). Strategic alliances in a manufacturing supply chain: Influence of organizational culture from the manufacturer's perspective. *International Journal of Physical Distribution & Logistics Management*, *40*(6), 456–474.
Sarathy, R. (2006). Security and the global supply chain. *Transportation Journal*, *45*(4), 28–51.
Schoenherr, T., & Swink, M. (2012). Revisiting the arcs of integration: Cross-validations and extensions. *Journal of Operations Management*, *30*(1–2), 99–115.
Scholten, K., Sharkey, Scott, P., & Fynes, B. (2014). Mitigation processes – antecedents for building supply chain resilience . *Supply Chain Management: An International Journal*, *19*(2), 211–228.
Sheffi, Y., & Rice, J. B. (2005). A Supply Chain View of the Resilient Enterprise Resilient. *MIT Sloan Management Review*, *47*(1), 41–48.
Srinivasan, R., & Swink, M. (2015). Leveraging Supply Chain Integration through Planning Comprehensiveness: An Organizational Information Processing Theory Perspective. *Decision Sciences*, *46*(5), 823–861.
Stank, T. P., & Keller, S. B. (2001). Supplychain Collaboration And Logisticalservice Performance. *Journal of Business Logistics*, *22*(1), 29–48.
Swink, M., Narasimhan, R., & Wang, C. (2007). Managing beyond the factory walls: Effects of four types of strategic integration on manufacturing plant performance. *Journal of Operations Management*, *25*(1), 148–164.
Tomlin, B. (2006). On the Value of Mitigation and Contingency Strategies for Managing Supply Chain Disruption Risks. *Management Science*, *52*(5), 639–657.
Tummala, R., & Schoenherr, T. (2011). Assessing and managing risks using the Supply Chain Risk Management Process (SCRMP). *Supply Chain Management: An International Journal*, *16*(6), 474–483.
Tushman, M. L., & Nadler, D. A. (1978). Information processing as an integrating concept in organizatonal design. *Academy of Management Review*, *3*(3), 613–624.
Vickery, S. K., Jayaram, J., Droge, C., & Calantone, R. (2003). The effects of an integrative supply chain strategy on customer service and financial performance: An analysis of direct versus indirect relationships. *Journal of Operations Management*, *21*(5), 523–539.
Wagner, S. M., & Neshat, N. (2012). A comparison of supply chain vulnerability indices for different categories of firms. *International Journal of Production Research*, *50*(11), 2877–2891.
Waters, D. (2007). *Supply Chain Risk Management: Vulnerability and Resilience in Logistics*. London, UK: Kogan Page.
Wieland, A., & Wallenburg, C. M. (2012). Dealing with supply chain risks: Linking risk management practices and strategies to performance. *International Journal of Physical Distribution & Logistics Management*, *42*(10), 887–905.
Wieland, A., & Wallenburg, C. M. (2013). The influence of relational competencies on supply chain resilience: a relational view. *International Journal of Physical Distribution & Logistics Management*, *43*(4), 300–320.
Wiengarten, F., Humphreys, P., Gimenez, C., & McIvor, R. (2016). Risk, risk management practices, and the success of supply chain integration. *International Journal of Production Economics*, *171*, 361–370.
Wong, C. Y., Boon-Itt, S., & Wong, C. W. Y. (2011). The contingency effects of environmental

uncertainty on the relationship between supply chain integration and operational performance. *Journal of Operations Management*, *29*(6), 604–615.
Wong, C. Y., Wong, C. W., & Boon-itt, S. (2015). Integrating environmental management into supply chains. *International Journal of Physical Distribution & Logistics Management*, *45*(1/2), 43–68.
Zhang, M., & Huo, B. (2013). The impact of dependence and trust on supply chain integration. *International Journal of Physical Distribution & Logistics Management*, *43*(7), 544–563.
Zsidisin, G. A., Ellram, L. M., Carter, J. R., & Cavinato, J. L. (2004). An analysis of supply risk assessment techniques. *International Journal of Physical Distribution & Logistics Management*, *34*(5), 397–413.
Zsidisin, G. A., & Wagner, S. M. (2010). Do Perceptions Become Reality? The Moderating Role of Supply Chain Resiliency on Disruption Occurrence. *Journal of Business Logistics*, *31*(2), 1–20.

DECISION SCIENCES INSTITUTE

Building Disaster Immunity through Absorptive Capacity in Supply Chain Management

Javad Sadeghi, Elisabeth Struckell, Divesh Ojha
G. Brint Ryan College of Business, University of North Texas
Email: Javad.Sadeghi@unt.edu; Elisabeth.Struckell@unt.edu; Divesh.Ojha@unt.edu

ABSTRACT

In supply chain management, we examine how to build the capability of managing and avoiding disasters – disaster immunity. Using complex adaptive systems theory, this study supports a novel three dimensional framework of absorptive capacity that includes absorptive infrastructure, complementary knowledge, and human resource management. Our model investigates the influence of two resources – information quality and change management capability – in creating the disaster immunity capability. We used the data collected from 264 US service firms to test our hypothesized model. Our empirical results indicate both information quality and change management capability completely mediate the effect of absorptive capacity on disaster immunity.

KEYWORDS: Absorptive Capacity, Disaster Immunity, Change Management Capability, Information Quality, Complex Adaptive Systems.

INTRODUCTION

It is surprising that given increased focus on prevention at government and corporate levels, the frequency and intensity of natural and manmade disasters and their consequences are increasing in the US, over a fifty-year period. Natural disasters such as hurricanes and earthquakes negatively affect the competitive capabilities of organizations and supply chains by creating disruptions.

A viable approach is developing resilience for managing disasters – disaster immunity capability. In this paper, we build on Ojha (2008) and consider disaster immunity as an organization's ability to manage disasters more effectively than its competitors. Disaster immunity (DI) refers to the organization's ability to avoid, recover, and mitigate negative impact from disasters.

Companies and their supply chain organizations are motivated by the substantial toll and increased frequency of disasters to develop capabilities to mitigate their damaging effects. Prior empirical studies have substantiated the positive relationship between absorptive capacity and organizational resilience to disasters. Absorptive capacity, defined by Cohen and Leventhal (1990) as "an ability to recognize the value of new information, assimilate it, and apply it to commercial ends", is especially useful to organizations in uncertain situations. Natural and manmade disasters are unplanned and epitomize uncertain conditions for impacted organizations. The contribution of this paper is the use of complex adaptive systems theory to empirically support a novel three dimensional framework of absorptive capacity that includes absorptive infrastructure, complementary knowledge, and human resource management. The model presented herein investigates the influence of two mediating resilience enhancing resources - information quality and change management capability – as means to build organizational resilience to disasters – disaster immunity capability. We address gaps surfaced in a review of supply chain recovery literature by (1) increasing the focus on this important organizational phenomenon, (2) providing empirical rigor to the proposed solution and (3) introducing a model that takes into consideration avoidance, recovery, and mitigation. We also contribute by answering calls to increase the awareness of the importance of disaster immunity capability for its humanitarian results beyond efficiency and by broadening the framework of disasters to encompass both natural and manmade disasters, consistent with organizational adoption of risk management focus.

What follows is a theoretical model, literature review, support for hypotheses, our methodology, researching findings, discussion and management implications.

LITERATURE REVIEW

We present a theoretical model, building disaster immunity capability through absorptive capacity and resilience enhancing resources.

Disaster immunity (DI) refers to an organization's ability to avoid disasters, minimize disaster effects, and recover quickly from disasters when compared with competitors.

Learning is a fundamental capability to build supply chain immunity against disasters. Cohen and Levinthal (1990) introduced absorptive capacity as “new perspective on organizational learning” defined as “the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends” (p128).

No study was identified to examine the role of absorptive capacity on a firm’s disaster immunity capability, however, the Roberts et al. (2012) review showed a significant body of literature, that considered absorptive capacity as an important enabler to improve firms’ performance.

We support our model with the theory of complex adaptive systems. In the context of this study, when a disaster occurs, external stimuli is introduced to the system – unpredictable in its timing and, at least temporarily uncontrollable. A disaster constitutes a complex organizational phenomenon and the organization entity which includes external interacting partnerships and customers represents a “system”. The firm impacted by the disaster represents the internal mechanism, which consists of agents (organizational business units) that operate under pre-defined rules or schema – organizational routines. The disaster is external to the everyday operating system and is dynamic in nature - constantly changing and non-uniform. In the disaster context, the schema can refer to both the everyday routines of the business as well as special schema that have been devised to prepare for disasters. Coevolution between the system and environment is a result of the feedback loops (learning), which are enhanced by the quality of information exchanged, and other resilience capabilities (change management capability) developed by the system that are inherent in the mutual interactions of the agents. Feedback is motivated by the changes that alter the state of equilibrium in the system and the nature of the organization to restore the equilibrium as quickly as possible and with as little negative impact as possible – disaster immunity capability and orientation.

HYPOTHESES DEVELOPMENT

Absorptive Capacity and Change Management Capability

Change management capability is the ability of a manager to understand how new knowledge should be used, and how to motivate and manage personnel through the change process. Absorptive capacity plays an important role in managing a business unit’s new knowledge. Change management is an ability to continually renew an organization’s direction, structure, and reconfigure capabilities with new knowledge to serve the ever-changing needs of external and internal customers. The importance of change management stems from its ability to help an organization to be competitive in addressing ever-changing customer needs and to survive in a highly competitive market. Feedback loops are an integral part of complex adaptive systems and a mechanism for organizational units to continuously adapt and learn. Torodovo and Durisin (2007) added feedback loops to the model of absorptive capacity. Applying the complex adaptive system model, business units will leverage their past knowledge, using organizational infrastructure, through repetitive interactions. This will give rise to emerging patterns, which will define the ability of the organizations to respond to calamity. We, thus, hypothesize –

H_1: Absorptive capacity is positively related to change management capability.

Change Management Capability and Disaster Immunity

Change management capability positively facilitates supply chain success, and at the same time has been found to mitigate supply chain risks. Applying complex adaptive systems theory, agents (business units) will coevolve naturally, as they strive to achieve the shared goal of survival in turbulent times with minimum possible losses. This shared vision between the business subunits of the organization will directly affect the motivation of the people managing the change, further influencing the firm's resilience to disasters. Thus, we hypothesize –

H_2: Change management capability is positively related to disaster immunity.

Absorptive Capacity and Information Quality

Information quality refers to "accuracy, reliability, relevance, adequacy, ease of access, and timeliness of the information shared across the supply chain" (Lusiantoro, Yates, Mena, & Varga, 2018). For information to be effective, it needs to be accurate (high quality) especially in situations of organizational stress and disruption. Reliance on information is important to decision making and progress. Access to quality information is critical to decision makers in supply chain management. Absorptive capacity can play an important role in identifying, assimilating and employing information – quality information.

The ability to identify the value of new knowledge influences the quality of information. Absorptive capacity motivates firms to renovate their information technology infrastructures that allow them to accurately and quickly process information throughout their supply chain. In their empirical study, Kyoon et al. (2011) showed that absorptive capacity is a critical antecedent to significantly increasing information quality. The quality of exchanging information can be improved by absorptive capacity.

In our research context, mutual interactions of the business subunits, when channeled through robust organizational infrastructure, by people high in intellect and diligence, will result in stronger inter unit connections, which foster effective and efficient communication in challenging times. We, thus, hypothesize that –

H_3: Absorptive capacity is positively related to information quality.

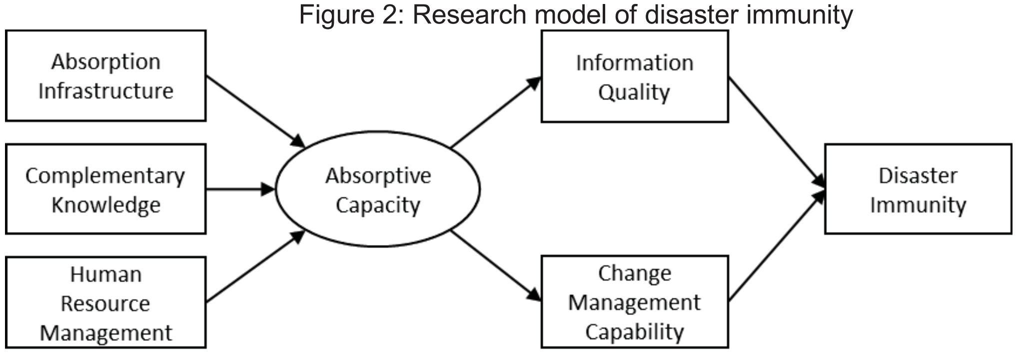
Information Quality and Disaster Immunity

Rich information exchange potentially enhances partnerships and ultimately impacts a firm's performance. Information quality leads to higher decision quality, business performance, and operational performance. Information quality plays an important role in firms to increase supply chain efficacy.

As firms strive to collaborate amongst themselves, the quality of information shared is critical (Işık, Jones, & Sidorova, 2013). Firms seek to establish robust processes to achieve effective inter-unit communication over multiple layers and enhance efficient but secure sharing of dynamic information, with minimum risk and uncertainty. These processes correspond to the coevolution focus of the complex adaptive systems, in which business subunits will collectively build resilience to forces of the external environment, as the patterns, emerging through dynamic interactions, will gravitate towards a coherent unity, after dropping the anti-bonding elements in the inter-unit connections. We, thus, hypothesize that –

H_4: Information quality is positively related to disaster immunity.

Figure 2: Research model of disaster immunity



METHODOLOGY

Analysis

The data was collected using online survey of mid to high level managers from US service industry. To analyze the research model data, we employ structural equation modeling (SEM) as a popular data analysis approach in empirical research. SEM enables us to simultaneously evaluate and estimate relationships in the entire system of variables. To evaluate our structural model, we apply a two-step process (1) evaluate the measurement model and (2) evaluate the structural model (Anderson & Gerbing, 1988). The statistical software utilized to conduct the analysis is IBM SPSS AMOS 25. Moreover, Bootstrapping method verifies the mediation roles of change management capability and information quality in the relationship between absorptive capacity and disaster immunity.

Results

We tested the nomological validity of our hypothesized model using overall fit indices. The structural model has a good overall model (χ^2[df]=719.29 [245], p≤0.000, CFI=.932, RMSEA=.086, RMR= 0.141).

The direct effects were evaluated using the significance of regression coefficients or the hypothesized relationships. The first hypothesis which states that absorptive capacity is positively related to change management capability (β = 0.169 and $p \le 0.001$), is supported. The second hypothesis which states that change management capability is positively related to disaster immunity is also supported (β = 0.283 and $p \le 0.001$). The third hypothesis which states that absorptive capacity change management capability is positively related to disaster immunity information quality is supported (β = 0.167 and $p \le 0.001$). The fourth hypothesis which states that information quality is positively related to disaster immunity is also supported (β = 0.332 and $p \le 0.001$).

Finally, we evaluated the indirect effects or implied mediation in our hypothesized model. We used Bootstrapping procedure in AMOS to test the indirect effect hypotheses. The indirect effect of absorptive capacity on disaster immunity through change management capability and information quality was supported ($\beta = 0.103$, $p = 0.024$).

DISCUSSION

The research supports the theoretical model introduced in the paper. First, the study proposed and research supports a three-dimensional framework of absorptive capacity as appropriate for the disaster immunity context - absorptive infrastructure, complementary knowledge, and human resource management. Second, using complex adaptive systems theory, we empirically demonstrate how business units are able to build disaster immunity capability (avoid, mitigate damages, and increase recovery time) through the absorptive capacity collective as well as through the development of two resilience enhancing mediating resources – (information quality and change management capability).

CONCLUSION

Natural and manmade disaster are increasing in frequency and intensity. Organizations are aware that the damage from disasters reaches well beyond business interruption and can include loss of life and reputation ruin. Given the impact to the value of the enterprise, risk management is a topic that is discussed at the board level of organizations. Proactive organizations are developing capabilities to help them cope with disasters – to avoid them, to mitigate their impact and to recovery more quickly – we call this disaster immunity capability. The research uses complex adaptive systems theory to provide empirical support for how companies can develop disaster immunity capability through the development of absorptive capacity, change management capability and the quality of information exchanged.

REFERENCES

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, *103*(3), 411.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Işık, Ö., Jones, M. C., & Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management*, *50*(1), 13–23.

Kyoon Yoo, D., Vonderembse, M. A., & Ragu-Nathan, T. S. (2011). Knowledge quality: antecedents and consequence in project teams. *Journal of Knowledge Management*, *15*(2), 329–343.

Lusiantoro, L., Yates, N., Mena, C., & Varga, L. (2018). A refined framework of information sharing in perishable product supply chains. *International Journal of Physical Distribution & Logistics Management*, *48*(3), 254–283.

Ojha, D. (2008). *Impact of strategic agility on competitive capabilities and financial performance*. Ph.D., Clemson University, South Carolina, USA.

Roberts, Galluch, Dinger, & Grover. (2012). Absorptive capacity and information systems research: review, synthesis, and directions for future research. *MIS Quarterly*, *36*(2), 625.

Todorova, G., & Durisin, B. (2007). Absorptive capacity: Valuing a reconceptualization. *Academy of Management Review*, *32*(3), 774–786.

DECISION SCIENCES INSTITUTE

Exploring the Relationship Between Organizational Culture, Total Quality Management, and Green Organizations

Lillian Y. Fok
Western Washington University
fokl@wwu.edu

Yun-Chen Morgan
Southeastern Louisiana University
ymorgan@selu.edu

Susan M. L. Zee
Southeastern Louisiana University
susan.zee@selu.edu

ABSTRACT

In this study, we examine multiple relationships among organizational variables and companies' overall green orientation and the resulting effects. How employees perceive cultural practices, quality management maturity, organizational impact of green initiatives, and their companies' overall performance are analyzed. The results are linked to organizations' green orientation embedded in their corporate strategies.

KEYWORDS: Green Movement; Quality Management; Organizational Culture

INTRODUCTION

Organizational efforts to preserve our environment are increasingly being scrutinized. In order to respond to the heightened customer awareness and demands, businesses need to place priority on incorporating green initiatives in corporate strategies in order to thrive among intense global competitions. The green movement is no longer a superficial or short-term focus to alter reputations and increase profit; it has become a vital part in most companies' operations. While "going green" may be costly in the beginning, business leaders realize the long-term benefits should not be ignored. This study explores the relationships between various organizations' green orientation and factors within organizations that can be influenced by the degree of green orientation. Specifically, we focus on the connection between the green movement and sustainability issues, and the linkages between green orientation and organizational culture, quality management maturity, organizational impact, and organizational success.

LITERATURE REVIEW

Companies choose to practice in greener ways because they realize they can contribute to the preservation of our environment, and also increase their profits. Deif (2011) suggested that there are three reasons that justify investing and implementing green manufacturing techniques: efficiency, market share, and government support and regulations. Employing greener manufacturing can save time and money, because the idea is to make the same product using fewer resources or energy. New consumer demands and heightened consumer awareness (Willson, 2011; Ackerman, 1997) combined with increased global competitions pressure manufacturing enterprises to review their strategies in order to secure market shares and sharpen their competitive edges. Pressure from governments to evolve into green manufacturing is also increasing. As regulations change, penalties and tax benefits increase, becoming greener at different manufacturing stages may become a mandate rather an option. Companies worldwide are concerned about the ability to grow and remain profitable, and are pressured by stakeholders to be more eco-efficient (Klebnikoff, 1996).

Total Quality Management (TQM), Sustainability, and Green Initiatives

Total Quality Management (TQM) has its roots in the integrative approach in customer satisfaction and a company's overall success (Chin, Pun, & Hua, 2001). Organizations are constantly facing intense pressure from competitions, and they realize the need to incorporate sustainable development and TQM in order to reach higher level of improvement and ultimate profitability (Hitchcock & Willard, 2002; Jonker, 2000; McAdam & Leonard, 2003). It is a logical continuation of research to expand TQM into a concept that also includes sustainability, long-term survival and growth with the emphasis of globalizing economies (Dervitsiotis, 2001; Wilkinson, Hill, & Gollan, 2001; Zairi, 2002).

Employee attitudes regarding green initiatives may be influenced by their perceptions toward Total Quality Management (TQM) programs implemented in the workplace. Previous research examined employee involvement (Rapp & Eklund, 2002), human resource management and leadership (Daily & Huang, 2001), commitment (Matta, Davis, Mayer, & Conlon, 1996), and personality traits (Ahmad & Schroeder, 2002) all expressed possible connections between perceptions and attitudes. How employees perceive the effectiveness of varying TQM tools employed in the workplace should have significant consequences on employees' attitudes about going green.

Organizational Culture

Employees of the same organization often share similar values and attitudes, and in turn, behave similarly. Therefore, organizational culture often is an invaluable tool to implement company strategies. When green movement inevitably becomes a part of organizational strategy, the culture that binds members of the organization together also impact the level of success on green initiatives. Organizational realize their brand image is on the line if they neglect their part in implementing green initiatives internally. Companies incorporating these initiatives within part of their corporate culture could become the leaders in the green movement and reach higher level of sustainability and profitability (Acharya, Vadher, & Acharya, 2014).

RESEARCH MODEL

In this research, we extend the examination of green movement issues to consider organizational culture, employee perceptions of organizational commitment to the green movement and the relationships which may exist between personal and organizational commitment and the QM Maturity of the organization. Additionally, where there is a higher level of perceived commitment to the green movement and where more mature QM systems are in place, we expect that, in the overall, the organization itself will be seen as "doing better" and the impact of the green movement will be perceived as favorable. Thus, QM and perceptions of the organization's green movement will be seen as having positive impacts upon organizational outcomes. The research model shows the linkages we expect and relates linkages to the corresponding research questions. In our study, we believe that the green movement within the organization should be related to or affected by the organizational culture. Additionally, as organizations become more green-oriented, the organization itself will be seen as "doing better" in general and the impact of the green movement will be more positive. Furthermore, organizations with more desirable organizational culture should be more supportive of the green movement. We also expect that organizational culture is related to the organization's performance in general. Finally, we also believe that as more organizations with more desirable organizational culture, where QM systems are in place, and the organization is "doing better," the employees will perceive the impact of the green movement to be even better.

Research Question 1: Organizational Green Orientation is related to Organizational Culture.

Research Question 2: Organizations which are described by employees as higher in Organizational Green Orientation, will also report more positive feelings about the impact of the green movement.

Research Question 3: Organizations which are described by employees as higher in Organizational Green Orientation will also report more positive feelings about the organization's performance.

Research Question 4: Organizations which are described by employees as higher in Organizational Performance, they will also report more positive feelings about the impact of the green movement.

Research Question 5: Organizations which are described by employees as higher in Organizational Culture, they will also report more positive feelings about the impact of the green movement.

Research Question 6: Organizations which are described by employees as higher in Organizational Culture will also report more positive feelings about the organization's performance.

Research Question 7: Organizations which are described by employees as higher in Organizational Green Orientation will also be described as having higher QM Maturity.

Research Question 8: Organizations which are described by employees as higher in QM Maturity, they will also report more positive feelings about the impact of the green movement.

Research Question 9: Organizations which are described by employees as higher in QM Maturity will also report more positive feelings about the organization's performance.

Research Question 10: Organizations which are described by employees as higher in Organizational Culture will also be described as having higher QM Maturity.

METHODOLOGY

Subjects of the Study

Subjects in the sample were approximately 331 full-time employees from a wide variety of industries in the South. The subjects were roughly 51.7 % male and 48% female with an average age of 41.49. Among these 331 subjects, there are approximately nine categories of industries in our survey instrument. The three largest percentages of the subjects work in retail, health care, and manufacturing industries with 22.1%, 20.8%, and 12.7%, respectively. The results are shown in Table 1.

The results are shown in Table 1.

Table 1 Types of Industry and Annual Revenue of the Organizations

Industry Types		Frequency	Percent
	Manufacturing	42	12.7
	Financial Services	15	4.5
	Retail	73	22.1
	Utilities	9	2.7
	High Technology	7	2.1
	Education	18	5.4
	Health Care	69	20.8
	Government	5	1.5
	Other	92	27.8
	Total	330	99.7
Missing	System	1	.3
Total		331	100.0
Revenue Groups		**Frequency**	**Percent**
	Over $1,000 Million	103	31.1
	$501-$1,000 Million	14	4.2
	$251-$500 Million	13	3.9
	$101-$250 Million	14	4.2
	$51-$100 Million	15	4.5
	$25-$50 Million	21	6.3
	Less than $25 Million	113	34.1
	Total	293	88.5
Missing	Unkmown	37	11.2
	System	1	.3
	Total	38	11.5
Total		331	100.0

Approximately 31.1% of the subjects are employed in a company which has over $1,000 million annual revenue, 4.2% of the subjects are employed in a company which has $501 to $1,000 million annual revenue, 3.9% of the subjects are employed in a company which has $251 to $500 million annual revenue, 4.2% of the subjects are employed in a company which has $101 to $250 million annual revenue, 4.5% of the subjects are employed in a company which has $51 to $100 million, 6.3% of the subjects are employed in a company which has $25 to 50 million annual revenue, and 34.1% of the subjects are employed in a company which has less than $25 million annual revenue. Subjects responded to a survey which asked about their perceptions and experiences about the green movement, quality management, and organizational culture in their own firms. In this study, we will concentrate on the relationships among commitment to the green movement, quality management, organizational performance, and impacts of organizational culture.

Instrument

Organizational Green Orientation

Based on the previous research (Li, Hartman, and Zee, 2009), we measured the Organizational Green Movement by using the survey which provides twenty-one organizational green initiatives question items. The "Green Orientation" variable is a count of the total number of green initiatives implemented in an organization.

Quality Management (QM) Maturity

In this study, QM Maturity refers, in a qualitative sense, to the *degree* of QM implementation in an organization. We suggest, and previous research has shown (Ahire & Golhar, 1996; Flynn, Schroeder, & Sakakibara, 1994; Fok et al., 2000, 2001; Patti, Hartman, & Fok, 2001; Saraph, Benson, & Schroeder, 1989) that it can be measured by examining the perceived use of QM programs. In earlier research (Fok et al., 2000, 2001; Patti et al., 2001), we began the process of developing a measure of QM maturity. The instrument we developed dealt with perceived program *use* and asked respondents whether certain programs were in use in the organization, with a range from "not used" to "high usage."

In this study, the QM maturity instrument was used to gauge QM maturity. We conducted a factor analysis to identify the underlying dimensionality. Two factors emerged from the "Usage" items. The first factor appeared to include all the traditional quality management programs and was termed "Basic TQM Tools." The second factor was termed "Advanced TQM Tools" which includes programs like Six Sigma programs and Black Belt training. 67.39% of the variance was explained by these two factors.

Organizational Culture

Based on previous research (Fok et al., 2000, 2001; Harman, Fok, & Zee, 2009), we measure the Organizational Culture by constructing a series of paired opposite items which asked whether the organization's climate should be described as open vs. closed, soft vs. tough, competitive vs. collaborative, and the like. Table 3 below provides the items and shows the results of our factor analysis. We obtained a two-factor solution in case of the Organizational Culture items and have labeled Factor 1 as “Continuous Improvement Culture” and Factor 2 as “People-Friendly Culture.” 44.529% of the variance was explained by these two factors.

Organizational Performance

We used the Baldridge criteria in the form of a scale which asks respondents to provide perceptions about their organizations along Baldridge lines. The instrument included are items such as “Overall, my company is performing well,” “Overall, morale in my company is high,” “Overall, my company is productive,” and the like. Factor analysis in this study indicated that one factor was present. The results showed that 56.205% of the variance was explained by the factor and we named the factor as “Organizational Success.”

Impact of Green Movement

The instruments included are items such as “Have better relationship with suppliers,” “Have better reputation,” “Provide better working environment,” and “Improve productivity.” Factor analysis produced a single-factor solution and we named it “Impact of Green Movement.” 70.987% of the variance was explained by this factor.

RESULTS

Our first research question examines the relationship between Organizational Green Orientation and Organizational Culture. Table 2 provides the results of our correlation analysis. Both “Continuous Improvement Culture” and “People-Friendly Culture” have significant correlations with “Green Orientation” which implies that as organizations embrace a culture that focuses on quality, team, and being proactive and employee-friendly, they also report higher levels of green orientation.

Research Question 2 is supported by the results in Table 2. The findings support the premise that when organizations develop “green” products/services or use “green” material in the production, show more concern with avoiding negative consequences of not being green, and help their employees at all levels to be more green-oriented, the overall impact of these green initiatives is perceived to be more positive by the employees.

Research Question 3 is supported. The results are shown in Table 2. The relationship between “Green Orientation” and “Organizational Performance/Success” is significant at the .000 level. The relationship is negative which implies that as the organizations show more concern in helping their employees at all levels to be more green-oriented, and pay more attention to safety concerns, the organizational performance is perceived by the employees to be higher.

Research Question 4 is confirmed. The relationship is positive and imply that organizations with higher levels of performance would also be reported to have employees with more positive feelings about the impact of the green movement in their organizations.

Table 2 Pearson's Correlation Matrix – Organizational Green Orientation, Organizational Culture, Organizational Performance, and Impact of Green Movement

Correlations

	Green Orientation	Continuous Improvement Culture	People-Friendly Culture	Organizational Performance	Impact of Green Movement
Green Orientation	—	.127*	-.216**	.414**	.202**
Continuous Improvement Culture	.127*	—	—	.235**	.582**
People-Friendly Culture	-.216**	—	—	-.192**	NS
Organizational Performance	.414**	.235**	-.192**	—	.313**
Impact of Green Movement	.202**	.582**	NS	.313**	—

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
NS = not significant.

Our fifth research question examines the relationship between Organizational Culture and Impact of Green Movement. We found one pair of significant relationship in Table 3. "Continuous Improvement Culture" has a significant and positive correlation with "Impact of Green Movement".

Table 3 Pearson's Correlation Matrix – Organizational Culture, Impact of Green Movement, and Organizational Performance

Correlations

	Continuous Improvement Culture	People-Friendly Culture	Organizational Performance	Impact of Green Movement
Continuous Improvement Culture	—	—	.235**	.582**
People-Friendly Culture	—	—	-.192**	NS
Organizational Performance	.235**	-.192**	—	.313**
Impact of Green Movement	.582**	NS	.313**	—

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
NS = not significant.

Research Question 6 is supported by the results in Table 3. Two pairs of significant relationships are found. "Organizational Performance/Success" has significant correlations with both "Continuous Improvement Culture" and "People-Friendly Culture".

Our seventh research question examines the relationship between Organizational Green Orientation and QM Maturity and Table 4 provides the results. There are two pairs of significant relationships. The factor "Organizational Green Orientation" has significant and positive correlations with both "Use of Basic TQM Tools" and "Use of Advanced TQM Tools". The results confirm that when employees perceive that their organizations are more inclined to develop green products/services and have a green workplace will be reported to have used higher levels of usage of both traditional and advanced TQM tools.

Table 4 Pearson's Correlation Matrix – Organizational Green Orientation, QM Maturity and Impact of Green Movement

Correlations

	Green Orientation	Basic TQM Tools	Advanced TQM Tools	Impact of Green Movement
Green Orientation	—__	.316**	.211**	
Basic TQM Tools	.316**	—	—	.445**
Advanced TQM Tools	.211**	—	—	NS
Impact of Green Movement		.445**	NS	—

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
NS = not significant.

Research Question 8 examines the relationship between QM Maturity and Impact of Green Movement. We found one of significant relationship in Table 4. "Use of Basic TQM Tools" has a significant and positive correlation with "Impact of Green Movement".

Research Question 9 suggests that organizations which are described by employees as higher in QM Maturity will also have employees who report more positive feelings about the organization's performance. Both factors "Use of Basic TQM Tools" and "Use of Advanced TQM Tools" have significant and positive correlations with "Organizational Performance/Success" in Table 5.

Research Question 10 suggests that organizations with higher level of organizational culture would be reported by employees to have higher level of QM Maturity. We found three pairs of significant relationships in Table 5. The relationship between "Continuous Improvement Culture" and "Use of Traditional TQM Tools" is significant at the .000 level which implies that as organizations embrace culture that focuses on quality, team, and being proactive, they also use more traditional TQM tools. “People-Friendly Culture” has significant relationships with “Use of Traditional TQM Tools” and “Use of Advanced TQM Tools”.

Table 5 Pearson’s Correlation Matrix – QM maturity, Organizational Performance, and Organizational Culture (RQ9 and RQ10)

Correlations

	Continuous Improvement Culture	People-Friendly Culture	Basic TQM Tools	Advanced TQM Tools	Organizational Performance
Continuous Improvement Culture	—	—	.331**	NS	.235**
People-Friendly Culture	—	—	-.189**	-.271**	-.192**
Basic TQM Tools	.331**	-.189**	—	—	.378**
Advanced TQM Tools	NS	-.271**	—	—	.181**
Organizational Performance	.235**	-.192**	.378**	.181**	—

**. Correlation is significant at the 0.01 level (2-tailed).
NS = not significant.

DISCUSSION AND CONCLUSIONS

In this research, we find substantial support for relationships among employee perceptions of organizational green orientation and various outcomes include impact of green movement, organizational performance, QM maturity, and organizational culture.

Our results suggest that the more informed employees are about what the company is doing to meet the new societal expectations of eco-friendliness, the more likely they are going to feel positively about various aspects of the organization. Potentially those positive feelings will lead to higher satisfaction and productivity. It is then logical to conclude that more education on the companies' green initiatives should be provided to employees, because the dissemination of these information will be exceedingly helpful in building a strong organizational culture. A strong organizational culture that is supportive of the green movement, and in turn leads to positive organizational outcomes.

REFERENCES

Acharya, S. G., Vadher, J. A. and Acharya, G. D. (2014). A review on evaluating green manufacturing for sustainable development in foundry industries. *International Journal of Emerging Technology and Advanced Engineering.* Vol. 4, Issue 1, pp. 232-237.

Ahire, S. L. and Golhar, D. Y. (1996) 'Quality management in large versus small firms: An empirical investigation', *Journal of Small Business Management*, Vol. 34, No. 2, pp. 1-13.

Ahmad, S. and Schroeder, R. G. (2002) 'The importance of recruitment and selection process for sustainability of total quality management', *International Journal of Quality & Reliability Management*, Vol. 19, No. 5, pp. 540-550.

Ahn, S.H. (2014) 'An evaluation of green manufacturing technologies based on research databases', *International Journal of Precision Engineering and Manufacturing Green Technology,* Vol. 1, No. 1, pp. 5-9.

Apple Rubber (2016) '5 important reasons to embrace sustainable manufacturing', *Apple Rubber Blog.* Retrieved from https://www.applerubber.com/blog/5-important-reasons-to-embrace-sustainable-manufacturing/

Assarlind M. and Gremyr, I. (2014) 'Critical factors for quality management initiatives in small- and medium-sized enterprises', *Total Quality Management & Business Excellence*, Vol. 25, Issue 3-4, pp. 397-411.

Beatley, T. and Manning, K. (1998) *The ecology of place: planning for environment, economy and community*, Washington DC: Island Press.

Beckman C. M. (2006) 'The influence of founding team company affiliations on firm behavior', *Academy of Management Journal*, Vol. 49, No. 4, pp. 741-758.

Boundless (2016) 'Differences between strategic planning at small versus large firms', Boundless Management Boundless, 26 May. 2016. Retrieved 24 Aug. 2017 from https://www.boundless.com/management/textbooks/boundless-management-textbook/strategic-management-12/strategic-management-86/differences-between-strategic-planning-at-small-versus-large-firms-418-1407/

Child, J. (1972) 'Organization structure, environment and performance: The role of strategic choice', *Sociology: The Journal of the British Sociological Association*, Vol. 6, pp. 1-22.

Chin, K. S., Pun, K. F. and Hua, H. M. (2001) 'Consolidation of China's quality transformation efforts: a review', *International Journal of Quality & Reliability Management*, Vol. 18, No. 8, pp. 836-853.

Cho, K. (1994) 'Impact of Total Quality Management (TQM) on organizational performance in the U.S.: An empirical investigation of critical success factors', ETD *collection for University of Nebraska - Lincoln*. AAI9425277. http://digitalcommons.unl.edu/dissertations/AAI9425277

Daily, B. F. and Huang, S. (2001) 'Achieving sustainability through attention to human resource factors in environmental management', *International Journal of Operations& Production Manageement*, Vol. 21, Issue 12, pp. 1539-1552.
Deif, A. M. (2011) 'A system model for green manufacturing', *Journal of Cleaner Production*, Vol. 19, pp. 1553-1559.
Dervitsiotis, K. (2001) 'Emerging elements of a world view for sustainable quality', *Total Quality Management*, Vol. 12, No. 7, pp. 817-824.
Doszhanov, A. and Ahmad, Z. A. (2015) 'Customers' intention to use green products: The impact of green brand dimensions and green perceived value', https://www.researchgate.net/publication/281733955.
Flynn, B. B., Schroeder, R. G. and Sakakibara, S. (1994) 'A framework for quality management research and an associated measurement instrument', *Journal of Operations Management*, Vol. 11, pp. 339-366.
Fok, L. Y., Fok, W. M. and Hartman, S. J. (2001) 'Exploring the relationship between total quality management and information systems development', *Information & Management*, Vol. 38, pp. 355-371.
Fok, L. Y., Hartman, S. J., Patti, A. L. and Razek, J. R. (2000) 'The relationship between equity sensitivity, growth need strength, organizational citizenship behavior, and perceived outcomes in the quality environment: A study of accounting professionals', *Journal of Social Behavior and Personality*, Vol. 15, No. 1, pp. 99-120.
Gan, C., Wee, H. Y., Ozanne, L. and Kao, T (2008) 'Consumers' purchasing behavior towards green products in New Zealand', *Innovative Marketing*, Vol. 4, Issue 1, pp. 93-102.
Hartman, S. J., Fok, L. Y. and Zee, S. M. (2009) 'Linkage among employee perceptions of organizational commitment to the green movement and organizational culture, and their perceived impacts upon outcomes', *Journal of Organizational Culture, Communications and Conflict*, Vol. 13, No. 2, pp. 27-44.
Hayes, D. (1990) 'The green decade', *The Amicus Journal*, No. 12, pp. 24-29.
Hitchcock, D. and Willard, M. (2002) 'Sustainability: enlarging quality's mission', *Quality Progress*, Vol. 35, No. 2, pp. 43-47.
Isa, M. R., Zaroog, O. S. and Raju, S. (2016) 'Survey on total quality management in small and medium manufacturing industries in Shah Alam', *International Journal of Science and Research*, Vol. 5, No. 1, pp. 848-853.
Jonker, J. (2000) 'Organizations as responsible contributors to society: Linking quality, sustainability and accountability', *Total Quality Management*, Vol. 11, No. 4/5/6, pp. 741-746.
Kangovou (2017) 'The importance of eco-friendly products'. Retrieved from https://www.kangovou.com/the-importance-of-eco-friendly-products/
Klebnikoff, S. D. (1996) 'Eco-efficient leadership; the road to new business opportunities. *Proceeding of the APO World Conference on Green Productivity*, Manila, pp. 491-502.
Kober, R., Subraamanniam, T. and Watson, J. (2012) 'The impact of total quality management adoption on small and medium enterprises' financial performance', *Accounting & Finance*, Vol. 52, Issue 2, pp. 421-438.
Kreitner, R. & Kinicki, A. (2007) *Organizational Behavior*, Boston: McGraw-Hill.
Lagas, B. (2016) 'Five benefits of sustainability and green manufacturing', *Quality Digest*.
Li, J., Hartman, S. J. and Zee, S. M. L. (2009) 'A study of green movement perceptions and behavioral intentions', *International Journal of Sustainable Economy*, Vol. 1, No. 2, pp. 133-143.

Loffler, P. (1998) ‘Sustainable development in Europe’, *A Cause for Regional Environment*, Vol. 8, pp. 113-120.
Ma, C., Liu, X., Zhang, H. and Wu, Y. (2016) ‘A green production strategies for carbon sensitive products with a carbon cap policy’, *Advances in Production Engineering and Management*, Vol. 11, No. 3, pp. 216-226.
Maheshwari, S. P. (2014) ‘Awareness of green marketing and its influence on buying behavior of consumers: Special reference to Madhya Pradesh, India’, *AIMA Journal of Management & Research*, Vol. 8, Issue 1/4, pp. 497-505.
Matta, K., Davis, J., Mayer, R. and Conlon, E. (1996) ‘Research questions on the implementation of total quality management’*, Total Quality Management*, Vol. 7, Issue 1, pp. 39-49.
McAdam, R. and Leonard, D. (2003) ‘Corporate social responsibility in a total quality management context: Opportunities for sustainable growth’, *Corporate Goverance*, Vol. 3, Issue 4, pp. 36-45.
Melao, N. F. and Guia S. M. (2015) ‘Exploring the impacts of ISO 9001 on small- and medium-sized social service institutions: a multiple case study’, *Total Quality Management & Business Excellence*, Vol. 26, Issue 3-4, pp. 312-326.
Patti, A. L., Hartman, S. J. and Fok, L. Y. (2001) ‘Investigating organizational quality management maturity: An investment validation study’, *International Journal of Quality & Reliability Management*, Vol. 18, No. 9, pp. 882-899.
Paul, I. D., Bhole, G. P. and Chaudhari, J. R. (2014) ‘A review on green manufacturing: Its importance, methodology and its application’, *Procedia Materials Science,* Vol. 6, pp. 1644-1649.
Peattie, K. (2001) ‘Towards sustainability: The third age of green marketing’, *The Marketing Review*, Vol. 2, No. 2, pp. 129-146.
Rapp, C. and Eklund, J. (2002) ‘Sustainable development of improvement activities - the long-term operation of a suggestion scheme in a Swedish company’, *Total Quality Management*, Vol. 13, No. 7, pp. 945-969.
Roberts, M. (2013) ‘Large versus small business quality management strategies’, http://blog.lnsresearch.com/bid/179410/large-versus-small-business-quality-management-strategies-data
Sandu, R. (2014) ‘Green: Marketing, products, and consumers’, *SEA-Practical Application of Science*, Vol. 2, Issue 3(5), pp. 555-562.
Saraph, J. V., Benson, G. P. and Schroeder, R. G. (1989) ‘An instrument for measuring the critical factors of quality management’, *Decision Sciences*, Vol. 20 pp. 810-829.
Sen, P. K., Bohidar, S. K., Shrivas, Y., Sharma, C. and Modi, V. (2015) ‘Study on innovation, research and recent development in technology for green manufacturing’, *International Journal of Mechanical Engineering and Robotics Research*, Vol. 4, No. 1, pp. 185-194.
Sezen, B. and Cankaya, S. Y. (2013) ‘Effects of green manufacturing and eco-innovation on sustainability performance’, *Procedia – Social and behavioral Sciences,* Vol. 99, pp. 154-163.
Sujith, T.S. (2017) ‘Awareness of green marketing and its influence on buying behavior of consumers in Kerala’, *International Journal of Scientific Research and Management*, Vol. 5, No. 7, pp. 43.
Wilkinson, A., Hill, M. and Gollan, P. (2001) ‘The sustainability debate’, *International Journal of Operations & Production Management*, Vol. 21, No. 12, pp. 1492-1502.
World Commission on Environment and Development (WCED), 1987. Report: ‘Our Common Future’.

Yang, Y. C. (2017) ‘Consumer behavior towards green products’, *Journal of Economics, Business and Management*, Vol. 5, No. 4, pp. 160-167.

Yeow, K. Y. and Yazdanifard, R. (2014) ‘The concept of green marketing and green product development on consumer buying approach’, *Global Journal of Commerce & Management Perspective*, Vol. 3, pp. 33-38.

Zairi, M. (2002) ‘Beyond TQM implementation: the new paradigm of TQM sustainability’, *Total Quality Management*, Vol. 13, No. 8, pp. 1161-1172.

Zairi, M. and Liburd, I. M. (2001) ‘TQM sustainability – a roadmap for creating competitive advantage’, *Integrated Management, Proceeding of the 6th International Conference on ISO 9000 and TQM*, Paisley, Ayr, Scotland, 17-19 April, pp. 452-461.

Zee, S. M., Fok, L. Y., and Harman, S. J. (2011) ‘Exploring the relationships between organizational size and market focus and commitment to the green movement and impacts of organizational culture: A comparative study of Jamaica and the United States’, *International Journal of Business and Social Science*, Vol. 2, No. 22, pp. 19-34.

DECISION SCIENCES INSTITUTE

A Conceptual Model of ISO 14001 and Lean Integration for Evaluation and Improvement of Environmental Performance in Manufacturing Industry

Aqeel Ahmed
Massey University of New Zealand
A.Ahmed@massey.ac.nz

Sanjay Mathrani
Massey University of New Zealand
S.Mathrani@massey.ac.nz

ABSTRACT

Excessive emissions of greenhouse gases, production of wastes and inefficient utilization of resources have drastic impacts on environmental and operational performance in manufacturing industry. International organization for standardization (ISO) 14001 and lean are used as tools for wastes minimization which have a positive impact on environment. Although ISO 14001 and lean have been largely deployed in the manufacturing sector, little evidence in literature establishes synergies between both strategies in achieving environmental and operational benefits. In this research-in-progress article, an integrated model for ISO 14001 and lean is proposed to evaluate and improve environmental and operational performance in the manufacturing sector.

KEYWORDS: Lean, ISO 14001, integration, environmental performance and manufacturing industry.

INTRODUCTION

The increase in global warming and its effects have been discussed strategically in manufacturing companies since the 1950s. Generation of carbon dioxide, methane, nitrous oxides and ozone in manufacturing operations cause greenhouse gases (GHG) effect and ultimately contribute to global warming phenomena (Abas et al., 2018). Excessive consumption of energy and inefficient utilization of natural resources are considered as potential sources for emissions of GHG in energy-intensive industrial facilities (Roy et al., 2012). Furthermore, the consumption of nitrogen-rich fuels in transportation, fossil fuels in manufacturing processes and emissions of gases from energy-intensive appliances and equipment are also responsible for emissions of GHG (Langematz, 2018). The emissions of GHG are responsible for an increase in surface temperature, depletion of the ozone layer and melting of glaciers in the last few decades (Ding, Zhang, Zhao, Li, & Kang, 2019).

ISO 14001 standard emerges as a prominent internal environmental controlling tool and an appropriate way of expressing the environmental commitment of an organization to its internal and external stakeholders (Wong, Abdullah, Baini, & Tan, 2017). The practical implementation of ISO 14001 facilitates manufacturing operations in preventing wastes generation (air emission and solid wastes), and helps to fulfill the legal and operational compliance requirements through an environmental management system (EMS) (Mariotti, Kadasah, & Abdulghaffar, 2014; Massoud, Fayad, El-Fadel, & Kamleh, 2010). Alternatively, lean manufacturing philosophy has been adopted in manufacturing firms for identification and reduction of the wastes in production processes through an on-going improvement cycle (Verrier, Rose, & Caillaud, 2016). The inefficiencies of raw material utilization and energy consumption are also minimized through process improvement as per the lean perspective (Johansson & Sundin, 2014). The major aim for implementing waste reduction practices in manufacturing industry is to create a healthier environment, increase the profitability of the business, and boost the public image.

Many companies have implemented lean for improving the process flow (Puvanasvaran, Kerk Swee Tian, A.L. Vasu, & Muhamad, 2012). Lean manufacturing practices are uniform in nature, with a comprehensive set of strategies that provide tools and techniques for improving the processes by reduction of wastages (Habidin, Hibadullah, Mohd Fuzi, Salleh, & Md Latip, 2018). On the other hand, ISO 14001 is considered as a strategy for providing rules and roadmap to formulate and adopt an EMS framework in the manufacturing and service facilities (Fisher, 2003). The elements of ISO 14001 provide a clear framework for constructing an environmental program. The environmental objectives, plans and policies with operational procedures are defined through the framework by implementing ISO 14001. However, the current environmental measurement systems still fail in providing the relevant information needed for improvement (Puvanasvaran, Swee Tian, & Vasu, 2014). For instance, the environmental data sources currently used are basically designed for monitoring regulatory compliance only, and not measuring the environmental performance. Although, the measurement of environmental performance is an important parameter for identifying success or failure of an environmental strategy. Lean manufacturing is a strategy that provides value to customer by eliminating wastes (Verrier et al., 2016). The lean strategy includes tools and techniques for waste reduction which could provide additional support to the ISO 14001 EMS processes for enhancing environmental benefits. An integrated system by combining EMS and lean tools could be applied to streamline work processes in reducing wastes (environmental and operational) as well as evaluate performance (Puvanasvaran, Kerk Swee Tian, & Muhamad, 2011). Therefore, the research question for this study is *how can ISO 14001 and lean be integrated to evaluate and improve environmental performance in the manufacturing industry?* This article helps in developing a conceptual model by using different elements of ISO 14001 and lean tools and techniques for an integrated approach towards environmental and operational performance management by sustaining continual improvement through Dr. Deming's plan, do, check and act (PDCA) cycle.

LITERATURE REVIEW

ISO 14001 and Lean Manufacturing

ISO 14001 provides a framework for developing an EMS in an organization. EMS is a framework for monitoring the process for improving the environmental impacts by optimizing resource utilization and reduction of wastes generation in manufacturing facilities (Chattopadhyay, 2001). EMS is an instrumental tool for improving environmental performance and serves as a business strategy. Furthermore, it predicts the present deficiencies in the system for future improvement and management to maintain environmental performance proactively (Qi et al., 2011). The main benefits of an effective implementation of ISO 14001 include improvement in internal processes with reduction in environmental aspects (emissions of GHG and wastes generation) and an increase in production efficiency (Mariotti et al., 2014). One study stated that the World Bank initiated a program to provide technical assistance for implementation of ISO 14001 in 11 energy extensive facilities in Mexico (Wong et al., 2017). The majority of these companies experienced a reduction in wastes generation (solid and air), improvement in resource utilization (raw material and energy consumption) and increase in economic performance.

Since the start of the 1980s, environmental waste management and effective utilization of resources (raw materials and energy consumption) are treated as an integral part of strategy and policy deployment in the manufacturing industry after the introduction of Toyota Production System (TPS) concepts (Campos, de Melo Heizen, Verdinelli, & Cauchick Miguel, 2015). The mass production concept is substituted by lean manufacturing after the successful launch of TPS in the manufacturing industry (Aguado, Alvarez, & Domingo, 2013). The lean philosophy can be defined as: “To do more with less” (Azadeh, Yazdanparast, Zadeh, & Zadeh, 2017, p155). Lean manufacturing provides a favorable business environment for a company by achieving gains in production operations, creating cost-saving opportunities and maintaining customer satisfaction through operational processes (Alves & Alves, 2015). In a study conducted in the European context, it was found that the manufacturing industry consumes around 37% of total primary energy production in Europe. Europe Strategy 2020 aims to reduce 20% of the total energy consumption by implementing lean practices in their industries (May, Barletta, Stahl, & Taisch, 2015). Lean and environmental improvements are proactive tools for reducing heat waste, noise pollution and eliminating environmental impact by focusing on customer satisfaction (Pampanelli, Found, & Bernardes, 2014). The integration of environmental performance and lean is developed into an improvement strategy that achieves environmental and economic benefits simultaneously (Cherrafi, Elfezazi, Garza-Reyes, Benhida, & Mokhlis, 2017). Integration of lean strategy and ISO 14001 will further provide a holistic value-based method to achieve environmental performance in an on-going improvement cycle (Puvanasvaran et al., 2014).

Justification for Selection of ISO 14001 from Entire ISO 14000 Family

ISO 14000 family consists of several standards with their unique process and product-oriented attributes for improving environmental performance in manufacturing processes and service industry (Boudouropoulos & Arvanitoyannis, 1998). The main reasons for selecting ISO 14001 only in this study from the entire ISO 14000 family of standards are: (1) construction of EMS by ISO 14001, (2) high rate of adoption and flexibility of ISO 14001 and (3) low certification cost.

The fundamental standards for constructing an EMS within the ISO 14000 family are – ISO 14001 and ISO 14004 with some additional guidelines. However, only ISO 14001 is used as a base standard for obtaining ISO certification and conducting environmental performance audits because the other standards and guidelines of ISO 14000 family are auxiliary in nature (Fortuński, 2008). The high rate of adoption of ISO 14001 motivates an organization in obtaining ISO 14001 certification. ISO 14001 as an EMS framework has been adopted by more than 32000 international organizations since its launch in 1996 (Ferrón Vílchez, 2017). This certification can be applicable to all types and sizes of organizations since it is a generic standard, is flexible and adjustable to different cultures, geographical and social conditions (Neves, Salgado, & Beijo, 2017). Furthermore, ISO 14001 is a globally accepted international standard which improves the environmental performance of both service industry (e.g., hotel industry and adaptation in university campuses) (Chan & Wong, 2006; Fisher, 2003) and manufacturing sector (e.g., automobile, chemical and food industry) (Azadeh et al., 2017).

From a business perspective, adequate financial support is required for the implementation of the full ISO 14000 series because it is an expensive adoption for small organizations. ISO 14000 adaptation has two types of costs (direct and indirect). The direct cost for adoption of ISO 14000 series involves consultancy expenditures, certification fees and registration audits, and it varies from about \$10,000 (small organization) to \$30,000 (large organization). The indirect cost for implementation of ISO 14000 includes internal training programs, redesign of products and advertisement initiatives for modified environmental-friendly products. It ranges from $100,000 (small organization) to over $600,000 (large organization) (Miles, Munilla, & Russell, 1997) However, the adaption costs for ISO 14001 certification is approximately US$15,000 to US$20,000 in the manufacturing sector (Massoud et al., 2010).

Justification for Integration of ISO 14001 and Lean Strategies

Many different management strategies have been implemented for process improvement in manufacturing in recent decades. For example, (1) lean emphasizes on wastes minimization (Habidin et al., 2018), (2) ISO 14001 pivots on environmental waste minimization (Mariotti et al., 2014), (3) agile targets flexibility and speed (Greer & Hamon, 2011), (4) flexible manufacturing system (FMS) provides flexibility in operations with a minimum time and expenditure for changeover (ElMaraghy, 2005), (5) reconfigurable manufacturing system (RMS) offers flexibility in adjusting its production functionality and capacity by changing its structure as well as its hardware and software components (Koren & Shpitalni, 2010), and (6) dedicated manufacturing line (DML) focuses on manufacturing a single component of a product in high volume over a long period (ElMaraghy, 2005; Koren & Shpitalni, 2010). All above management strategies have a different focus. However, lean and ISO 14001 both focus on waste elimination and improvement of resource efficiency in manufacturing processes (Habidin et al., 2018). Therefore, the integration of ISO 14001 and lean approaches look promising in improving environmental performance because both philosophies have a common focus and some mutual operational objectives.

CONCEPTUAL MODEL OF ISO 14001 AND LEAN IMPLEMENTATION

A model is conceptualized for integration of ISO 14001 and lean strategies as shown in Figure 1. This model has been inspired from a framework by Dalrymple Bay Coal Terminal's EMS program (Dalrymple, 2019) which explains the process of implementing an environmental management plan using the PDCA cycle, however developed through a rigorous review of literature. The suggested structures consist of policies and procedures of ISO 14001 and its measurement tools as well as techniques of lean for improvement of environmental performance in the manufacturing industry. An integrated approach of ISO 14001 and lean implementation is conceptualized for monitoring, measuring and improving environmental performance. The proposed ISO 14001 and lean model is built upon on a continuous improvement process through the PDCA cycle.

The model illustrates that the environmental and waste reduction policies of ISO 14001 and lean implementation are formulated, documented and communicated to relevant stakeholders in the "plan" phase. These policies link with the integrated environmental waste management plan (IEWMP) through a bold arrow which represents the integration relationship (e.g., waste elimination policies of both) and dotted arrow which depicts non-combined impacts of both policies (e.g., legal implications from ISO and operational benefits of lean) in this model.

A clear definition of policy and its effective implementation is at the heart of every management strategy. The environmental policy can be defined as a short written statement which is based on the current status of the environmental performance of an organization by indicating possible future improvements (Uzumeri, Tabor, & Stanwick, 1996).

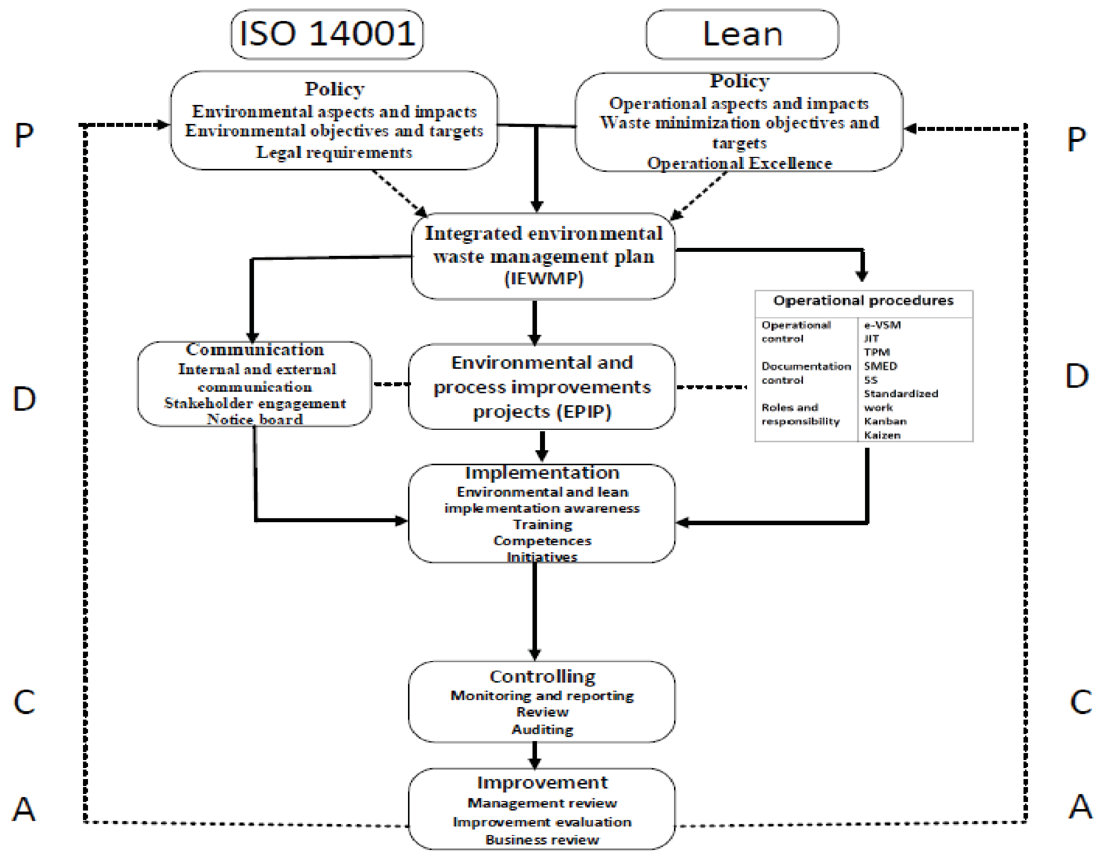


Figure. 1. A conceptual model for integration of ISO 14001 and lean in manufacturing industry

The foundation of the environmental policy lies in the initial environmental performance review of environmental aspects, legal requirements and regulations and the current environmental procedure. The waste minimization program statement includes the environmental objectives and quantifiable targets for EMS. The legal and regulatory requirements are an integral part of EMS policy to reduce or eliminate the potential fines, penalties and strategic constraints for an organization (Arimura, Darnall, Ganguli, & Katayama, 2016). The top management sets the lean and ISO 14001 policies according to the vision of an organization for achieving wastes minimization (Alefari, Salonitis, & Xu, 2017) and environmental performance objectives and targets (Uzumeri et al., 1996) to maintain continual improvements. The operational excellence (financial performance) is considered as an indicator for operational performance in the lean policy. Financial performance improves by reducing operational and environmental costs in the manufacturing sector (Belekoukias, Garza-Reyes, & Kumar, 2014). Once policies for ISO 14001 and lean are documented, it becomes possible to create a clear plan for IEWMP.

The “do” phase of PDCA’s cycle starts from IEWMP until the implementation stage in this conceptual model. The communication, environmental process improvement projects (EPIP) and operational procedures of both strategies start simultaneously indicated through bold arrows. The in-between dotted lines depict their interrelationships.

The communication process consists of internal and external communication, stakeholder engagement and notice board. The communication between the company and its stakeholders (internal and external) provides the fundamentals for the successful execution and maintenance of ISO 14001 (Sohal & Zutshi, 2004) and lean strategy (Worley & Doolen, 2006). An effective communication acts as an information medium for enabling all stakeholders to learn about the changes and exchange of ideas for achieving environmental objectives and targets. The communication medium remains active between an organization and its stakeholders in different implementation stages of EMS (Sohal & Zutshi, 2004). The waste reduction objectives and targets are clearly communicated to employees in different shifts and all stakeholders in producing quality products (Worley & Doolen, 2006). A clear management plan is communicated to all stakeholders (employees, suppliers and contractors) through newsletters, intranet and Internet websites, bulletin boards and trade journals (Savely, Carson, & Delclos, 2007). A continual smooth flow of information can also be managed by putting the relevant details of management initiatives (policy, objectives and targets) on the notice board (Wu & Low, 2012).

The operational procedures are divided into lean tools and techniques as well as process control initiatives (operational control, documentation control, and roles and responsibility). Operational planning and control include critical activities to prevent and minimize environmental accidents, reduce emissions to air and handle hazardous chemicals during operational processes (Foster, 2010). Document control is a mandatory requirement in EMS for documenting the scope of EMS, environmental policy and the core elements of environmental performance by maintaining EMS implementation records, evaluation and future improvement performance recommendations (Uzumeri et al., 1996). The leadership plays a key role in defining roles and responsibilities in an organization for execution of processes in implementing and improving EMS (Chattopadhyay, 2001). The different responsibilities are assigned to the senior, middle and lower workforce with sufficient financial and technical resources for improving the competence of the workforce through training and skill development programs (Muralikrishna & Manickam, 2017).

The lean tools provide solutions for improving environmental performance through implementation of its various methods and techniques. Environmental value stream mapping (e-VSM) is an essential lean tool for mapping consumption of resources (raw materials and energy consumption) as well as waste generation (gaseous emission, hazardous liquid and solid wastes) in manufacturing processing (Garza-Reyes, Kumar, Chaikittisilp, & Tan, 2018). The theme of just in time (JIT) is based on manufacturing the right products at the right time (Belekoukias et al., 2014). Furthermore, JIT helps an organization in managing inventory, decreasing wastes and reducing space utilization by reducing overproduction of goods in the manufacturing sector. Total productive maintenance (TPM) is an approach for maintenance of equipment to avoid breakdown, unnecessary stops and defects in operations. The operational benefits of TPM are controlling oil leakage, reduction of dust formation and emissions of fumes in manufacturing facilities (Belekoukias et al., 2014; Chiarini, 2014). Single-minute exchange of die (SMED) is a methodology practiced for decreasing the time to complete equipment changeovers (Oliveira, Sá, & Fernandes, 2017). Moreover, the reduction of lead time, lower inventory levels and process quality improvement can be achieved by effective implementation of SMED. The workplace is cleaned, organized and maintained through adequate consumption of resources as an effective strategy for waste disposal by sorting, setting in order, shining, standardizing and sustaining (5S) methods (Garza-Reyes et al., 2018). Standardized work is a set of rules and operational procedures to produce a product or execute service. This method is used for eliminating variation and inconsistency in processes by instructing the workforce to follow the defined rules and procedures in the implementation phase (Oliveira et al., 2017). Kanban is a lean tool for reducing idle time. The main philosophy behind kanban is to deliver what is necessary, when necessary. The meaning of kanban from the Japanese is a card or signal (Oliveira et al., 2017). Kaizen / continual improvement is a lean tool for improving programs and processes in a standardized manner. Kaizen is basically a tool for elimination of wastes by following a continual improvement strategy (Belekoukias et al., 2014).

The implementation stage starts on the basis of the outputs from the above three processes. This stage includes environmental and lean implementation awareness, competence development and training which are the integral processes for implementing new ISO and lean projects and their operational procedures. Awareness amongst the workforce is an important initiative to formulate environmental policy by comprehensively defining environmental aspects, impacts and objectives for environmental management (Mariotti et al., 2014). Competence of the workforce to improve environmental performance is judged on the basis of their education, training and experience for developing an environmental policy, implementing EMS and suggestions for improving environmental performance (Massoud et al., 2010). The knowledge creation, skill development and performance management of employees lay down the foundation for proactive lean implementation to achieve financial benefits and employee satisfaction (Bortolotti, Boscari, & Danese, 2015). EMS training helps the employees to understand policy and principles of environmental performance and their potential role in reducing environmental impacts by reducing air emissions, waste generation and effective utilization of raw materials (Jabbour, Jabbour, Govindan, Teixeira, & Freitas, 2013). The involvement of employees in training programs for lean implementation is an essential dimension towards a successful implementation by improving current skills and knowledge of the workforce in turn improving processes (Salonitis & Tsinopoulos, 2016).

The next phase of PDCA's cycle is "control" which is exemplified by a bold arrow indicating a link between implementation and controlling stage in the model. The controlling stage consists of monitoring and reporting on the implemented operational procedures as well as their review and audit. The effective implementation of the waste minimization program helps the organization in measuring and monitoring the environmental objectives and goals (Boudouropoulos & Arvanitoyannis, 2007).

In monitoring and measurement, the indicators of environmental performance constitute according to the organization's environmental policy and targets to achieve environmental objectives. These are achieved by introducing process controls such as calibration of equipment to reduce wastes and emissions, and maintaining ongoing evaluation of environmental performance to fulfill the operations and legal obligations (Muralikrishna & Manickam, 2017). The controlling of the lean initiatives include monitoring the level of success for lean implementation by comparing against the original plans (Mostafa, Dumrak, & Soltan, 2013). Internal audit is used as an evaluation tool to determine the current status of environmental performance against the policy, performance objectives and waste reduction targets in the organization (Foster, 2010).

The improvements are based on the outcomes of the "act" phase. Management review is an assessment initiative to propose future environmental improvements by introducing new performance operational requirements according to the current status of EMS (Muralikrishna & Manickam, 2017). General improvement includes the potential opportunities and actions (corrective, preventive and proactive) to identify the cause of environmental impacts, prevention strategies and elimination of environmental non-conformance during manufacturing processes for implementing EMS effectively (Muralikrishna & Manickam, 2017). Corrective actions are two types: (1) immediate actions which are taken as a follow-up of accidents and (2) revision actions which are made after the management review of EMS (Foster, 2010). The impacts of lean and environmental performance practices have different effects on business (financial and market) performance. The business and environmental performance are interrelated to each other in operational optimization of an organization. The targets are modified for business performance on the basis of lean and EMS implementation (Yang, Hong, & Modi, 2011). The continual improvement in this integrated ISO 14001 and lean model is achieved by the PDCA cycle through effective management of environmental audits, monitoring and measurement of environmental management by proposing improvement (proactive, preventive and corrective) actions.

BENEFITS OF ISO 14001 AND LEAN INTEGRATION

The integration of ISO 14001 and lean as a holistic approach could be utilized in achieving wastes minimization objectives and targets by monitoring, measuring and improving environmental and operational aspects in the manufacturing industry. The sustained integration of ISO 14001 and lean through the PDCA cycle for continual improvement can enable the incremental and on-going development of processes to achieve long-term environmental performance. The integration model will provide an opportunity to an organization to incorporate environmental considerations into the lean tools and techniques for reduction of environmental and operational wastes. A combined approach could be used for identification of environmental and operational deficiencies to provide solutions for mitigating inefficiencies as well as setting roles and responsibilities for all stakeholders. Training programs could be incorporated to create competencies in achieving EMS key performance indicators and implementation of related tools and techniques. These factors could achieve reduction of time and effort in improving the manufacturing processes, leading to an impact on the financial performance of an organization through process improvements such as reduction of environmental costs, scrap and waste generation, and treatment or control of wastes discharge. The integrated ISO 14001 and lean strategies would provide a clear vision for internal (management and employees) and external (suppliers and customers) stakeholders to achieve EMS related organizational goals, competitive advantage and profitability in the manufacturing industry.

CONCLUSION AND FUTURE RESEARCH DIRECTION

This study establishes a relationship between ISO 14001 and lean strategies to develop an integrated approach for evaluating and improving an organization's environmental performance. The various elements of ISO 14001 and lean tools provide the building blocks in the conceptual model for implementing, monitoring and measuring an EMS using the PDCA cycle. Further, the integrated approach of ISO 14001 and lean would help in streamlining the manufacturing processes to achieve long term competitiveness for an organization. This holistic model is generic in nature and could be implemented to any type and size of a manufacturing company for effective utilization of resources (raw materials and energy consumption) and reduction of wastes (solid, liquids and gaseous). In future, this study would be extended by an empirical evaluation of this integrated model through qualitative case studies to examine the impact of ISO 14001 and lean in achieving environmental benefits in the manufacturing industry.

REFERENCES

Abas, N., Kalair, A. R., Khan, N., Haider, A., Saleem, Z., & Saleem, M. S. (2018). Natural and synthetic refrigerants, global warming: A review. *Renewable and Sustainable Energy Reviews, 90*, 557-569. doi: https://doi.org/10.1016/j.rser.2018.03.099

Aguado, S., Alvarez, R., & Domingo, R. (2013). Model of efficient and sustainable improvements in a lean production system through processes of environmental innovation. *Journal of Cleaner Production, 47*, 141-148. doi: 10.1016/j.jclepro.2012.11.048

Alefari, M., Salonitis, K., & Xu, Y. (2017). The Role of Leadership in Implementing Lean Manufacturing. *Procedia CIRP, 63*, 756-761. doi: https://doi.org/10.1016/j.procir.2017.03.169

Alves, J. R. X., & Alves, J. M. (2015). Production management model integrating the principles of lean manufacturing and sustainability supported by the cultural transformation of a company. *International Journal of Production Research, 53*(17), 5320-5333. doi: 10.1080/00207543.2015.1033032

Arimura, T. H., Darnall, N., Ganguli, R., & Katayama, H. (2016). The effect of ISO 14001 on environmental performance: Resolving equivocal findings. *J Environ Manage, 166*, 556-566. doi: 10.1016/j.jenvman.2015.10.032

Azadeh, A., Yazdanparast, R., Zadeh, S. A., & Zadeh, A. E. (2017). Performance optimization of integrated resilience engineering and lean production principles. *Expert Systems with Applications, 84*, 155-170. doi: 10.1016/j.eswa.2017.05.012

Belekoukias, I., Garza-Reyes, J. A., & Kumar, V. (2014). The impact of lean methods and tools on the operational performance of manufacturing organisations. *International Journal of Production Research, 52*(18), 5346-5366. doi: 10.1080/00207543.2014.903348

Bortolotti, T., Boscari, S., & Danese, P. (2015). Successful lean implementation: Organizational culture and soft lean practices. *International Journal of Production Economics, 160*, 182-201. doi: https://doi.org/10.1016/j.ijpe.2014.10.013

Boudouropoulos, I. D., & Arvanitoyannis, I. (2007). *Potential and Perspectives for Application of Environmental Management System (EMS) and ISO 14000 to Food Industries* (Vol. 16(2)).

Boudouropoulos, I. D., & Arvanitoyannis, I. S. (1998). Current state and advances in the implementation of ISO 14000 by the food industry. Comparison of ISO 14000 to ISO 9000 to other environmental programs. *Trends in Food Science & Technology, 9*(11), 395-408. doi: https://doi.org/10.1016/S0924-2244(99)00008-4

Campos, L. M. S., de Melo Heizen, D. A., Verdinelli, M. A., & Cauchick Miguel, P. A. (2015). Environmental performance indicators: a study on ISO 14001 certified companies. *Journal of Cleaner Production, 99*, 286-296. doi: 10.1016/j.jclepro.2015.03.019

Chan, E. S. W., & Wong, S. C. K. (2006). Motivations for ISO 14001 in the hotel industry. *Tourism Management, 27*(3), 481-492. doi: https://doi.org/10.1016/j.tourman.2004.10.007

Chattopadhyay, S. P. (2001). Improving the speed of ISO 14000 implementation: a framework for increasing productivity. *Managerial Auditing Journal, 16*(1), 36-40. doi: doi:10.1108/02686900110363636

Cherrafi, A., Elfezazi, S., Garza-Reyes, J. A., Benhida, K., & Mokhlis, A. (2017). Barriers in Green Lean implementation: a combined systematic literature review and interpretive structural modelling approach. *Production Planning & Control, 28*(10), 829-842. doi: 10.1080/09537287.2017.1324184

Chiarini, A. (2014). Sustainable manufacturing-greening processes using specific Lean Production tools: an empirical observation from European motorcycle component manufacturers. *Journal of Cleaner Production, 85*, 226-233. doi: 10.1016/j.jclepro.2014.07.080

Dalrymple, B. C. T. (2019). Environmental Management System (ISO 14001) of Dalrymple Bay Coal Terminal. Retrieved on 13 April,, 2019, from http://www.dbct.com.au/iso14001-environmental-management-system

Ding, Y., Zhang, S., Zhao, L., Li, Z., & Kang, S. (2019). Global warming weakening the inherent stability of glaciers and permafrost. *Science Bulletin*. doi: https://doi.org/10.1016/j.scib.2018.12.028

ElMaraghy, H. A. (2005). Flexible and reconfigurable manufacturing systems paradigms. *International Journal of Flexible Manufacturing Systems, 17*(4), 261-276. doi: 10.1007/s10696-006-9028-7

Ferrón Vílchez, V. (2017). The dark side of ISO 14001: The symbolic environmental behavior. *European Research on Management and Business Economics, 23*(1), 33-39. doi: 10.1016/j.iedeen.2016.09.002

Fisher, R. M. (2003). Applying ISO 14001 as a business tool for campus sustainability: A case study from New Zealand. *International Journal of Sustainability in Higher Education, 4*(2), 138-150. doi: 10.1108/14676370310467159

Fortuński, B. (2008). Does the environmental management standard ISO 14001 stimulate sustainable development?: An example from the energy sector in Poland. *Management of Environmental Quality: An International Journal, 19*(2), 204-212. doi: doi:10.1108/14777830810856582

Foster, C. (2010). 17 - Establishing an environmental management system in the food industry. In U. Sonesson, J. Berlin & F. Ziegler (Eds.), *Environmental Assessment and Management in the Food Industry* (pp. 333-355): Woodhead Publishing.

Garza-Reyes, J. A., Kumar, V., Chaikittisilp, S., & Tan, K. H. (2018). The effect of lean methods and tools on the environmental performance of manufacturing organisations. *International Journal of Production Economics, 200*, 170-180. doi: https://doi.org/10.1016/j.ijpe.2018.03.030

Greer, D., & Hamon, Y. (2011). Agile Software Development. *Software: Practice and Experience, 41*(9), 943-944. doi: 10.1002/spe.1100

Habidin, N. F., Hibadullah, S. N., Mohd Fuzi, N., Salleh, M. I., & Md Latip, N. A. (2018). Lean manufacturing practices, ISO 14001, and environmental performance in Malaysian automotive suppliers. *International Journal of Management Science and Engineering Management, 13*(1), 45-53. doi: 10.1080/17509653.2017.1288587

Jabbour, C. J. C., Jabbour, A. B. L. d. S., Govindan, K., Teixeira, A. A., & Freitas, W. R. d. S. (2013). Environmental management and operational performance in automotive companies in Brazil: the role of human resource management and lean manufacturing. *Journal of Cleaner Production, 47*, 129-140. doi: 10.1016/j.jclepro.2012.07.010

Johansson, G., & Sundin, E. (2014). Lean and green product development: two sides of the same coin? *Journal of Cleaner Production, 85*, 104-121. doi: 10.1016/j.jclepro.2014.04.005

Koren, Y., & Shpitalni, M. (2010). Design of reconfigurable manufacturing systems. *Journal of Manufacturing Systems, 29*(4), 130-141. doi: https://doi.org/10.1016/j.jmsy.2011.01.001

Langematz, U. (2018). Future ozone in a changing climate. *Comptes Rendus Geoscience, 350*(7), 403-409. doi: https://doi.org/10.1016/j.crte.2018.06.015

Mariotti, F., Kadasah, N., & Abdulghaffar, N. (2014). Motivations and barriers affecting the implementation of ISO 14001 in Saudi Arabia: an empirical investigation. *Total Quality*

Management & Business Excellence, 25(11-12), 1352-1364. doi: 10.1080/14783363.2014.912038

Massoud, M. A., Fayad, R., El-Fadel, M., & Kamleh, R. (2010). Drivers, barriers and incentives to implementing environmental management systems in the food industry: A case of Lebanon. *Journal of Cleaner Production, 18*(3), 200-209. doi: https://doi.org/10.1016/j.jclepro.2009.09.022

May, G., Barletta, I., Stahl, B., & Taisch, M. (2015). Energy management in production: A novel method to develop key performance indicators for improving energy efficiency. *Applied Energy, 149*, 46-61. doi: 10.1016/j.apenergy.2015.03.065

Miles, M. P., Munilla, L. S., & Russell, G. R. (1997). Marketing and environmental registration/certification: What industrial marketers should understand about ISO 14000. *Industrial Marketing Management, 26*(4), 363-370. doi: https://doi.org/10.1016/S0019-8501(96)00144-7

Mostafa, S., Dumrak, J., & Soltan, H. (2013). A framework for lean manufacturing implementation. *Production & Manufacturing Research, 1*(1), 44-64. doi: 10.1080/21693277.2013.862159

Muralikrishna, I. V., & Manickam, V. (2017). ISO 9000, 14000 Series, and OHSAS 18001 *Environmental Management* (pp. 177-208).

Neves, F. O., Salgado, E. G., & Beijo, L. A. (2017). Analysis of the Environmental Management System based on ISO 14001 on the American continent. *J Environ Manage, 199*, 251-262. doi: 10.1016/j.jenvman.2017.05.049

Oliveira, J., Sá, J. C., & Fernandes, A. (2017). Continuous improvement through "Lean Tools": An application in a mechanical company. *Procedia Manufacturing, 13*, 1082-1089. doi: https://doi.org/10.1016/j.promfg.2017.09.139

Pampanelli, A. B., Found, P., & Bernardes, A. M. (2014). A Lean & Green Model for a production cell. *Journal of Cleaner Production, 85*, 19-30. doi: 10.1016/j.jclepro.2013.06.014

Puvanasvaran, P., Kerk Swee Tian, R., A.L. Vasu, S., & Muhamad, M. R. (2012). Lean principles adoption in environmental management system (EMS) - ISO 14001. *Journal of Industrial Engineering and Management, 5*(2). doi: 10.3926/jiem.486

Puvanasvaran, P., Kerk Swee Tian, R., & Muhamad, M. R. (2011, 27-30 June 2011). *Principles and business improvement initiatives of lean relates to Environmental Management System.* Paper presented at the First International Technology Management Conference.

Puvanasvaran, P., Swee Tian, R. K., & Vasu, S. A. L. (2014). Lean environmental management integration system for sustainability of ISO 14001:2004 standard implementation. *2014, 7*(5), 21. doi: 10.3926/jiem.907

Qi, G. Y., Zeng, S. X., Tam, C. M., Yin, H. T., Wu, J. F., & Dai, Z. H. (2011). Diffusion of ISO 14001 environmental management systems in China: rethinking on stakeholders’ roles. *Journal of Cleaner Production, 19*(11), 1250-1256. doi: 10.1016/j.jclepro.2011.03.006

Roy, P., Orikasa, T., Thammawong, M., Nakamura, N., Xu, Q., & Shiina, T. (2012). Life cycle of meats: An opportunity to abate the greenhouse gas emission from meat industry in Japan. *Journal of Environmental Management, 93*(1), 218-224. doi: https://doi.org/10.1016/j.jenvman.2011.09.017

Salonitis, K., & Tsinopoulos, C. (2016). Drivers and Barriers of Lean Implementation in the Greek Manufacturing Sector. *Procedia CIRP, 57*, 189-194. doi: 10.1016/j.procir.2016.11.033

Savely, S. M., Carson, A. I., & Delclos, G. L. (2007). An environmental management system implementation model for U.S. colleges and universities. *Journal of Cleaner Production, 15*(7), 660-670. doi: https://doi.org/10.1016/j.jclepro.2006.01.013

Sohal, A. S., & Zutshi, A. (2004). Adoption and maintenance of environmental management systems: Critical success factors. *Management of Environmental Quality: An International Journal, 15*(4), 399-419. doi: 10.1108/14777830410540144

Uzumeri, M. V., Tabor, R. H., & Stanwick, S. D. (1996). Setting up environmental management systems under ISO 14001. *Journal of Corporate Accounting & Finance, 7*(3), 19-33. doi: 10.1002/jcaf.3970070304

Verrier, B., Rose, B., & Caillaud, E. (2016). Lean and Green strategy: the Lean and Green House and maturity deployment model. *Journal of Cleaner Production, 116*, 150-156. doi: 10.1016/j.jclepro.2015.12.022

Wong, J. J., Abdullah, M. O., Baini, R., & Tan, Y. H. (2017). Performance monitoring: A study on ISO 14001 certified power plant in Malaysia. *Journal of Cleaner Production, 147*, 165-174. doi: https://doi.org/10.1016/j.jclepro.2017.01.088

Worley, J. M., & Doolen, T. L. (2006). The role of communication and management support in a lean manufacturing implementation. *Management Decision, 44*(2), 228-245. doi: doi:10.1108/00251740610650210

Wu, P., & Low, S. P. (2012). Lean Management and Low Carbon Emissions in Precast Concrete Factories in Singapore. *Journal of Architectural Engineering, 18*(2), 176-186. doi: doi:10.1061/(ASCE)AE.1943-5568.0000053

Yang, M. G., Hong, P., & Modi, S. B. (2011). Impact of lean manufacturing and environmental management on business performance: An empirical study of manufacturing firms. *International Journal of Production Economics, 129*(2), 251-261. doi: 10.1016/j.ijpe.2010.10.017

Authors Index