EVOLUTION OF PRODUCT QUALITY IN US AND JAPANESE AUTOMOTIVE FIRMS:

AN EXPLORATORY LONGITUDINAL ANALYSIS

Hemant V. Kher*
Department of Business Administration
Alfred Lerner College of Business and Economics
University of Delaware
Newark, DE 19716
Phone: (302) 831-1784
kher@udel.edu

Christine T. Kydd
Department of Business Administration
Alfred Lerner College of Business and Economics
University of Delaware
Newark, DE 19716
Phone: (302) 831-1783
chriskyd@udel.edu

ABSTRACT

For years we have been hearing that US automobile manufacturers have been losing market share to their Japanese rivals who are reputed to make better quality vehicles. Most such reports are based on the initial quality surveys on new automobiles. In this paper we address two exploratory questions: (1) how does the quality of an automobile change with its age, and, (2) can firm level variables help explain differences in the way quality changes with age? To answer these questions, we collected *Consumer Reports*' reliability ratings on approximately 200 automobile models made by the Big-6 firms. Our data includes model years 1998-2007; thus model ages range from 1-10 years old. We found that not only do automobiles made by

Japanese firms have higher initial quality, but, as automobiles get older the difference in the product quality between US and Japanese firms increases. We also found that the more generalist a US automobile firm, i.e., the wider is the US firm's product offering in the marketplace, the lower its overall automobile quality. Conversely, Japanese generalist firms were found to have higher quality than specialist firms. The result is partly explained by the fact that Japanese firms have taken a different path to broadening their product variety – they have ensured a high level of quality of their initial offerings before entering newer market segments. We conclude by noting limitations of our work and suggesting directions for future research.

Key words: *Automobile reliability, generalist / specialist, mixed effects models*

INTRODUCTION

How reliable is your automobile? A 2008 report by MSNBC indicated that Americans are keeping their cars / trucks for a longer period now than ever before. While one would suspect that the reasons behind this finding would be factors such soaring gas prices, collapsing housing market and a relatively poor economy, the MSNBC report suggested that the main reason is the increasing durability of automobiles. The report piqued our curiosity: how does automobile durability reliability change over time? Our search led to articles in the popular press (e.g. Jones, 2006) which pointed to a change in the market share; the Big 3 US firms (GM, Chrysler and Ford) are experiencing a declining market share, while the Big 3 Japanese automotive firms (Toyota, Honda and Nissan) are experiencing an increase in their market share in recent years. Reasons behind these changes in the market share were attributed to product quality (as gaged using initial quality surveys for new vehicles): while the reliability of automobiles made by Japanese firms is quite high, the reliability of automobiles made by US firms has been lagging.

Although Operations Management (OM) literature is replete with examples of how Toyota in particular, and Japanese automobile manufacturers in general have exemplified the use of lean manufacturing methods to gain competitive advantage in the marketplace, we did not however find any academic study that examines differences in automobile reliability over time.

The purpose of our study is to examine longitudinal changes in automobile reliability. In particular, we seek to understand how the reliability of automobile changes (deteriorates) with its age. To set the stage, we first briefly review extant OM literature that focuses on the automotive industry within the US; as noted before, most of this research focuses on the differences in operations strategies used by Japanese and US firms. Based on our literature review, we propose hypotheses concerning the influence of firm specific factors on the way reliability changes with age. We use data on automobile reliability published by *Consumer Reports* magazine to test our hypotheses. We discuss how the reliability measure we use is different and more informative relative to the way automotive quality has been measured in prior research. We discuss our results and its implications, and we conclude with directions for future research.

LITERATURE REVIEW

What are the challenges associated with making a high quality, reliable automobile? To begin with, putting the product together itself presents quite a challenge. In the words of Fine, LaFrance and Hillebrand (1996), the "*automobile is one of the most complex consumer products in existence*". Made from over a ton of steel, comprising of thousands of parts (both mechanical and electronic), the challenge of producing a high quality automobile can be daunting. Parts that are used in assembling automobiles are procured from hundreds of suppliers (Fujimoto and Takeishi, 2001), thereby further complicating the challenge of maintaining high quality.

Researchers have examined differences in both internal processes, e.g. how do shop floor practices differ across US and Japanese firms, as well as external processes, e.g. how do US and Japanese firms interact with suppliers that supply automotive components. MacDuffie (1997) reiterates the point made above in examining differences in shop floor practices related to problem solving at Ford, GM and Honda plants. He examined three specific types of problems – water leaks, paint defects and functional electrical defects and noted that these defects are "*ubiquitous – no assembly plant in the world has succeeded in permanently eliminating these defects*", and that "*all three have many sources*". He argued that while experimentation and gathering before / after data should be an important part of the problem solving process, the way these problems were resolved varied considerably across firms. At GM, he noted, the use of before / after data was haphazard; data gathering was done primarily for accountability and not necessarily for problem solving. At Ford, the emphasis was on reporting that action was being taken to remedy the problem which meant data was gathered after a solution was found, but there was no baseline to compare the solution against. In stark contrast, at Honda documentation of process improvement included gathering before / after data, and experimentation with different solutions was quite common.

The differences across auto firms are not confined to shop floor operations alone. Liker and Choi (2004) point to the differences between Japanese (Toyota and Honda) and American (Chrysler, Ford and GM) auto firms with respect to managing supplier relationships and note that while Japanese firms have created strategic partnerships with their suppliers, American firms have been "*more or less at war with their suppliers*". Dyer (1996) provided more insights on the differences in the way auto firms like Toyota, Nissan, Chrysler, GM and Ford build networks with their suppliers. He found that in general, Japanese firms' suppliers were located in close proximity (about 60 miles for Toyota, 115 miles for Nissan, 425 miles for Chrysler, and more than 500 miles each, on average, for Ford and GM). Suppliers' sales persons had more face-to-face interactions with Japanese firms than with US auto firms. Suppliers tended to trust Japanese firms more than they did US auto firms, and Japanese firms shared a lot more information with their suppliers than did US firms. Toyota topped on every performance measure – it had the best quality (lowest defects/100 vehicles), shortest new-model cycle time, lowest total inventory to sales ratio, and highest average pretax return on assets. Nissan was second on all these measures, Ford was third on every measure except for inventory, while GM and Chrysler generally occupied the 4^{th} and 5^{th} place in general.

There are two key themes that emerge from our brief literature review. First, the process of quality improvement across firms is an on-going one. Despite being in existence for several years, and despite having produced a large number of automobiles over this duration, automotive firms continue to learn (evidence of this organizational learning comes from studies by Haunschild and Rhee, 2004, and Levin, 2000); in other words, although firms are improving quality through experience, no firm has as yet perfected the manufacturing process to produce defect free automobiles (one evidence of this comes from the number of recalls issued by each firm, every year). In other words, some degree of defects will be inherent in automobiles made by all firms. Further, with the usual use-related wear and tear, as automobiles age, in general their reliability will decline over time.

The second key theme evident from many of the studies that focus on the differences between the way Big 3 US auto firms, and the Big 3 auto firms from Japan operating in the US is that not only are there firm-specific factors that lead to differences in performance, but there are also country-specific differences that lead to performance differences. Evidence we have reviewed from both the academic literature and the popular press suggests that while auto firms of Japanese origin have emphasized and maintained high levels of quality in automobiles they produce, their US counterparts have lagged. However, some researchers, e.g. MacDuffie and Pil, 1996, have suggested that US and European manufacturers are converging and approaching Japanese quality levels; Valdes-Depana, 2006, also makes similar observations "*GM and Ford cars are not that bad; depending on which survey you believe, they may even have become pretty good*". On the basis of these two observations, we propose our first two hypotheses concerning the effect of an automobile's age on its reliability:

H1: *Automobile reliability will decline with age regardless of the manufacturing firm*

H2: *Quality of the automobiles made by Big-3 Japanese firms will be higher than those made by Big-3 US firms, and the differences in quality ratings of the automobiles made by these two groups of firms will increase with car model age*

Impact of generalization versus specialization on automobile quality

Skinner (1974) highlighted the benefits of focusing on a narrow product mix aimed at a particular market niche, which would suggest that automobiles made by specialist auto firms offering a narrow product line will have higher quality levels. Lapre and Tsikriktsis (2006) comment that "*simplicity, repetition, experience, and task homogeneity breed competence resulting in better customer service and competitive position*" (p. 354). Manufacturers with a narrow product line would tend to specialize in a small set of vehicles. With respect to learning, Argote (1999) notes that “specialist” firms are more likely to benefit from their experience in comparison to “generalist” firms because it is more difficult for the generalist firms to transfer knowledge acquired in one unit across its other diverse units. Studies by Ingram and Baum (1997), and Carroll and Swaminathan (2000) have also shown that specialist firms tend to benefit more from their past experience than do generalist firms.

Two recent studies, however, find that the rate of learning may not vary across generalist and specialist firms as suggested above. Lapre and Tsikriktsis (2006) hypothesized that focused (specialist) airlines will learn to reduce customer dissatisfaction faster than full-service (generalist) airlines, but their data did not support this hypothesis. However, follow up analysis revealed that the best focused airline did indeed learn faster than the best full-service airline. Haunschild and Rhee (2004) found that, as hypothesized, generalist automakers have more voluntary recalls than specialist automakers. But in contrast to their hypothesis that specialists will learn faster to reduce recalls, they found that it was actually the generalist automakers that learned faster from voluntary recalls than specialists when the learning target was reduction in mandated recalls. The authors reasoned that generalist automakers are typically larger in size, and get more negative publicity from their mandated recalls than do specialist automakers. Generalists are therefore more concerned about avoiding mandated recalls and associated negative publicity, and thus have a greater incentive to learn faster from their voluntary recalls.

What are the implications of these studies on automobile quality across US and Japanese firms? In our view, the way in which a firm adopts a broader product range should impact the quality of products it offers. Specifically, US automotive firms typically have several different brand offerings (e.g. even after restructuring, GM has four brands – Buick, Cadillac, Chevrolet and GMC); in contrast, most Japanese firms have about 2-3 brands (Honda and Acura, Nissan and

Infiniti, and Toyota, Lexus and its more recent brand Scion). Further, while the US firms have a wide range of offerings across its brands, Japanese firms, especially Honda and Toyota have entered some markets (e.g. heavy trucks) at a relatively later stage compared to the US firms. Several different brands, coupled with a wide range of offerings have two important quality related implications. First, there is much product similarity across brands (e.g. Welch, 2005, noted that at one time GM had 6 similar family sedans across its different brands); thus, if components are found defective to warrant a recall, it is likely that a large number of different automobiles are affected. Second, with a large vehicle offering, it is difficult to devote sufficient time and resources in redesigning existing vehicles. Welch (2005) notes that while it took GM 9 years to replace Cavalier with Cobalt, Honda completely redesigns Civic every 5 years. In summary, we believe that differences in the way the generalist approach is adopted across firms has very different quality implications. Specifically, we believe that generalist US-based automotive firms will have lower quality (after controlling for the effects of time, as noted in H1), and that there will not be any difference in the quality for generalist and specialist Japanese auto firms (after controlling for effects of time as noted in H2).

H3a: *After controlling for the effect of age on ratings, rate of quality decline will vary between generalist and specialist automakers such that generalist US firms will have lower quality ratings than specialist firms*

H3b: *After controlling for the effect of age on ratings, rate of quality decline will not vary between Japanese generalist and specialist firms*

DATA COLLECTION

We test hypotheses concerning automobile reliability using data published by *Consumer Reports* (CR) during 2008 (permission on file). In all of our hypotheses, an automobile model (e.g. Ford Mustang, Toyota Camry, etc.) is the unit of analysis. Data was collected on a total of n=196 1998-2007automobile (cars / trucks) models made by the Big-6 firms. Of these, 65 models are from Japanese auto firms and the remaining 131 models are from US auto firms. Together these six firms accounted for an average of 85% of the US vehicle market in each of the ten years (1998-2007) covered in our study. We now describe the dependent and independent variables used to test the hypotheses.

Dependent Variable: Automobile Quality

How does one assess automobile quality? Conlon et al. (2001) note that while "*there are many dimensions to quality* … *the one dimension that has received the most attention in the automotive industry* … [is] … *vehicle reliability*" (p. 1194). But in prior studies automotive quality has been measured in different ways. As noted before, McDuffie (1997) studied quality using assembly related defects like water leaks, paint defects, and functional electrical defects are addressed by US and Japanese automakers (specifically, Ford, GM and Honda). Levin (2000) also focused on three assembly related defects -- body rust, paint and exterior trim, and body integrity. Barber and Darrough (1996) measured quality in terms of product recalls in the 6 largest auto firms in the US, while Haunschild and Rhee (2004) and later Rhee and Haunschild (2006) studied involuntary (mandated) and voluntary (self-initiated) recalls as quality surrogates for all auto makers selling vehicles in the US in the 1966-1999 time span. Conlon et al. (2001)

and Devaraj et al. (2001) used the composite rating from CR to portray quality, while Rhee and Haunschild (2006) used it as a proxy for automaker "reputation". The composite rating, labeled as "used car verdict," is calculated by weighting expensive to repair items more than other items. The used car verdict thus classifies all cars into 1 of 5 groups (1=worst to 5=best).

In our view, the measures used to capture automobile quality in past studies are diverse and collectively do not present a comprehensive measurement of this complex construct. For instance, Levin (2000) considered only three different types of defects that may arise during the assembly process. However, customers may weigh problems in other key areas such as engines, transmissions, drive systems, etc. even more heavily in judging the quality of automobiles. Similarly, Haunschild and Rhee (2004) documented reduction in recalls made by automakers. However, recalls may not completely reflect automobile quality or reliability. As noted by Barber and Darrough (1996), while recalls may be viewed by customers as a signal regarding the credibility of the production process, customers observe many other factors (e.g. such as those studied by Levin, and the others listed above) that characterize the quality of vehicles. We should also point out that recalls are not always mandated (i.e. some problems are such that the automaker voluntarily decides whether to announce a recall), and that recalls may or may not affect future automotive quality (e.g. a recall related problem may be addressed as a "one-time fix" implemented at the dealership; this may not necessarily result in "learning" for the automaker). It is also possible that an auto maker might be producing lower quality vehicles despite having reduced recalls over time. Recalls imply potentially catastrophic failures that may cause injury and liability. Quality defects that may not fall into such a category are still a problem for customers even when they do not generate recalls. Even the measure used by Conlon et al. (2001) and Devaraj et al. (2001) is just a composite quality score that does not average many sub-dimensions of quality that cumulatively could result in a defective or an under-performing vehicle. Such a composite rating does not have enough variation in it to fully capture learning effects or firm level differences in automobile quality.

TABLE 1: Automobile reliability dimensions used by Consumer Reports (2008):

	Item
1	Engine Major
2	Engine Minor
3	Engine Cooling
4	Transmission Major
5	Transmission Minor
6	Drive System
7	Fuel System
8	Electrical
9	Climate System
10	Suspension
11	Brakes
12	Exhaust
13	Paint/Trim
14	Body Integrity
15	Body Hardware
16	Power Equipment
17	Audio System

NOTE: Each item was scored from 1=poor to 5=best.

We use a comprehensive measure that encompasses a more robust and multi-dimensional picture of automotive quality. Specifically, we use 17 different factors ("trouble spots", listed in Table 1) to comprehensively gage automobile reliability. For each problem spot, CR assigns a rating in relation to the "average" for that problem spot for that year. As an example, consider transmission as the problem spot. Automobiles with the fewest reported transmission problems relative to the average reported number of transmission problems for all used automobiles for that year receive a rating of 5 (excellent). Automobiles falling in the second lowest number of transmission problems would receive a rating of 4 (good); while those reporting transmission problems near the average for that year receive a rating of 3 (average). Finally, automobiles receiving ratings of 2 or 1 have the second highest and highest number of reported transmission problems relative to the average reported number of transmission problems for all used automobiles for that year, respectively. For each model we created a summated index of CR's ratings for the 17 problem spots; this index varies from a low of 17 to a high of 85.

Independent Variable: Model Age

A model year tells us how old it is; e.g. a 1998 model is 10 years old, whereas a 2007 model is 1 year old in our data. We use the age of the model (coded from 0 to 9) in testing our hypotheses. Our dataset is unbalanced in that the number of annual weighted composite quality scores for the models in our sample range from a low of 1 to a high of 10, with the average being 6. Only about a fourth of the models in our sample have data for all ten years (e.g. Toyota Camry); some models were discontinued at some point within our time window (e.g. Acura CL was discontinued after 2003), while many new models were introduced within our window of analysis (e.g. Honda Ridgeline was introduced in 2006).

Independent Variable: Generalism / Specialism

Haunschild and Rhee (2004) note that the measurement of generalist / specialist varies across different studies. Fortunately, the issue of measuring generalism / specialism within the automotive industry has precedence. Specifically, Dobrev et al. (2001) measured this variable for automakers using the range of engine capacity in horse power. Their rationale is that specialist firms tend to offer a limited range of engine capacities (i.e. vehicles with more or less similar engines), whereas generalist firms offer a wider variety of vehicles and this is reflected in a wider range of engine horsepower (specifically, smaller vehicles will typically have smaller engine capacity, whereas larger SUVs and heavy trucks will use engines with larger capacity – a wide range is thus reflective of a firm offering a wider product variety). We follow this approach and calculate the difference between a firm's largest and smallest engine, measured in horsepower, and use the variable to represent the firm's degree of specialization. Summary statistics for the dependent and independent variables are given in Table 2.

TABLE 2: Summary statistics for variables used in the study.

Firm	US or Japan?	No. of Models	Avg. Rating*	Avg. Age*	Engine HP Range
Chrysler	US	23	53	4.54	360
Ford	US	35	59	4.29	330
GM	US	73	53	4.53	395
Honda	Japan	18	72	4.29	122
Nissan	Japan	19	64	4.11	87
Toyota	Japan	28	75	4.02	195

*Note: Average model rating and average model age is over the 1998-2007 period of the models in our data.

Analytic Strategy

Because we have repeated measures for each model, using ordinary least squares (OLS) regression to analyze the data would lead to biased standard errors for the estimates (Raudenbush and Bryk 2002). With repeated measures, the appropriate analytical approach would be the use of a multilevel (hierarchical) regression model, also termed as the random coefficients model (RCM), wherein we assume that repeated measures on reliability ratings are nested (correlated) within each model. The RCM is flexible in that allows us to specify both fixed and random effects – in other words, results of the model help understand whether the rate of decline in reliability is the same (fixed) for all models, or whether there is a significant difference (random effects) in the rate of decline across models.

For all of our hypotheses, the model's reliability rating score is the dependent variable (DV). To test H1 (reliability will decline with model age for all models), we use the model's age as the independent variable. We also introduce non-linearity by including a quadratic term (Age^2) in our model to test whether the decline in ratings tapers off. To test H2, which posits that rate of ratings decline will vary across US and Japanese firms, we introduce a dummy variable J (coded 0 for US firms and 1 for Japanese firms). To capture different rates of decline for the two groups of firms, we introduce interaction terms between J and Age (J*Age) and J and Age^2 ($J*Age^2$). Finally, to test H3a and H3b, we enhance the model used to test H2 by including the specialist / generalist variable described above (HP-Diff), along with its interaction with J.

RESULTS

All of our models are fit in STATA, version 12.0, using the XTMIXED (mixed models) framework. To determine whether the use of random coefficients model (RCM) is appropriate for the data, we first fit a baseline model for automobile reliability. An ICC value of 0.70, indicated a strong correlation among reliability ratings within each model (i.e. ratings values are not independent), and justified the use of RCM framework.

TABLE 3: Results of multilevel regression models.

	MODEL 1 H1	MODEL 2 H2	MODEL 3 H3a & H3b
<u>Fixed Effects</u>			
Intercept	72.82	71.10	101.15
Age	-4.93	-6.34	-6.28
Age^2	0.30	0.39	0.38
J	-	6.00	-35.11
J*Age	-	3.79	3.79
J*Age^2	-	-0.22	-0.22
HP-Diff	-	-	-0.08
J*HP-Diff	-	-	0.16
<u>Random Effects</u>			
Intercept (Variance)	1.73	0.82	0.81
Age (Variance)	48.23	36.49	34.09
Covariance (Int., Age)	5.76	0.92	-0.46
Residual (Variance)	24.34	23.73	23.81

Note: All fixed effects are significant at the 1% level (z > 2.33 in each case). Except for the covariance between the intercept and age in Models 2 and 3 all random effects are significant at 1% level.

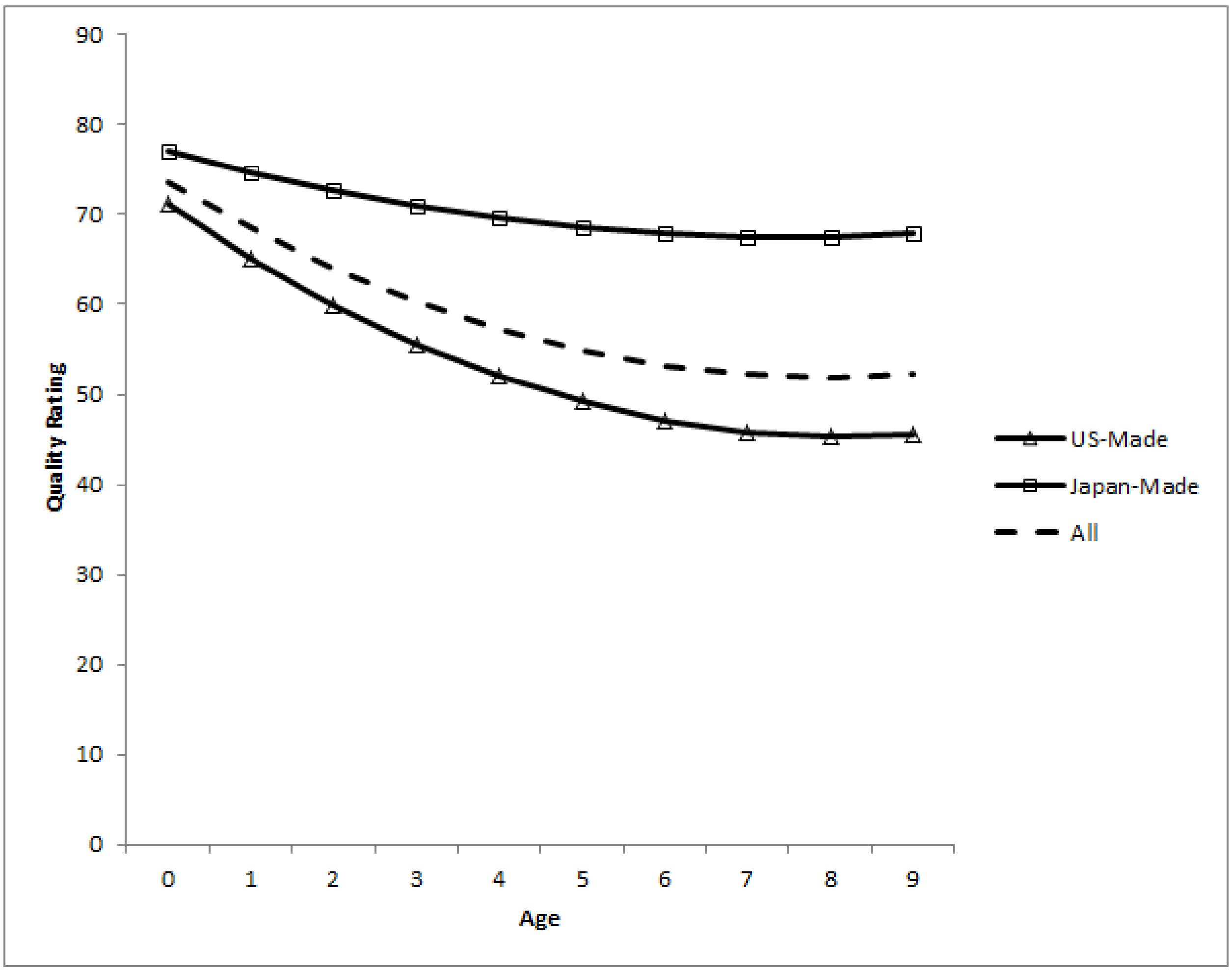
We then followed steps outline by Bliese and Ployhart (2002) on determining whether our RCM model should utilize fixed and random intercepts, and fixed and random slopes for the main dependent variable age. A series of likelihood ratio tests indicated that (1) fixed and random intercepts for reliability ratings greatly enhanced our model (i.e. even in the initial year, 2007, there was significant variability in reliability ratings across models), and (2) fixed and random slopes for age enhanced our model (i.e. rate of decline in reliability varied significantly across models). Results for testing hypotheses H1, H2, H3a and H3b are presented in Table 3. Note that with a few minor exceptions (as noted in Table 3), the fixed and random effect coefficients are highly significant ($p<0.01$).

We begin with Model 1, which addresses H1. The coefficient for Age is negative, suggesting that reliability rating declines with age for all models in our data, whereas the coefficient for Age^2 is positive, indicating a tapering effect – in other words, there is a steep decline in reliability ratings initially, however, the rate of this decline levels off beyond a certain point (i.e. age). The covariance between the intercept and slope is positive; to properly interpret this result, note that the slope for age is negative. The positive covariance thus indicates that models that have above average initial ratings (i.e. ratings larger than the intercept value of 72.82) tend to have an above average rate of decline (i.e. these models have slope values that are greater than -4.93), and vice-versa.

The above finding concerning the relationship between initial quality ratings, and the rate of decline in ratings with age become clearer when we examine the results for Model 2, which addresses H2 (ratings for models made by US firms will be lower than the models from Japanese firms, and the difference in ratings will increase over time). While the average initial quality rating for US models is 71, the average initial rating for Japanese models is higher by 6 points. Whereas on average rating for US models declines at a rate of 6.34 points annually, the average rate of linear decline in Japanese models is lower (-6.34+3.79=-2.55). The quadratic terms for both groups are positive, indicating a tapering effect over time. A graph of model implied quality trajectories in Figure 1 shows that the difference in summated ratings for US and Japanese automobiles increases with age.

H3a and H3b posit that the effect of generalist / specialist will vary across the two groups of auto firms. While the coefficient for the generalism / specialism is negative (-0.08, $p<0.01$) for US firms, implying the more generalist a firm the lower its automobile rating on average, the coefficient for Japanese firms is positive and significant (+0.16-0.08 = +0.08, $p<0.01$) indicating that generalist Japanese firms actually have higher quality ratings. The latter result is different from our hypothesis that across Japanese auto firms, whether or not a firm is more generalist than other firm will not make a difference. As seen in Table 2, within Japanese firms Toyota has the largest range (195) while Honda has the second largest range for engine horsepower (122); models made by Toyota also have the highest average quality (75), while Honda has the second largest quality (72), which helps explain the result for H3b.

FIGURE 1: Model implied change in automobile quality as a function of automobile age.



Discussion

Before we discuss our results, we hope the reader will bear in mind the aggregation issue. Our analysis discusses quality at the firm level. Although our results show that automobiles made by Japanese firms have better quality than those made by US firms, the results cannot be interpreted to mean that every Japanese automobile is of high quality, and that every US automobile is of a lower quality as this is certainly not true (and nor is this a correct interpretation of our work).

Automobiles are complex products, and despite years of experience, no firm has as yet mastered the technique of building a totally defect free automobile on a repeated basis. The source of complexity of the automobile as a product has to do with the fact that each firm has its own unique approach to designing, manufacturing and assembling these parts and each firm must work with hundreds of parts suppliers. Over the past few decades, a number of scholars have examined differences between the firms in each of these areas and have attributed these differences to a variety of performance measures including quality. Several popular press articles have attributed the decline in market share for US firms to differences in vehicle quality, however, this has been examined in the context of the quality of a new car, using, e.g. JD Powers initial quality survey. While our work is still in early stages, we found that the differences in quality between US and Japanese automobile models get larger with automobile age.

Another finding that emerges from our study is that firm-level differences in product variety strategy impacts automobile quality. US firms have a presence in every automobile market segment, and their strategy of maintaining “product freshness” (entering a “hot” segment and discontinuing less popular models) comes at the cost of quality. Some (e.g. Welch, 2005) have reasoned that with a large offering, it becomes difficult for US firms to engage in activities aimed at continually improving automobile quality such as vehicle redesigns. Our results supported the hypothesis that for US firms, the more generalist a firm, the lower its automobile quality. Surprisingly, we found the opposite result for Japanese firms. One explanation for this unexpected finding may have to do with the way Japanese firms enter newer market segments – continue to do well what you already do, and enter new segments gradually.

LIMITATIONS AND FUTURE RESEARCH

A major limitation of our study is that our data did not capture a true ageing effect even though the models in our data varied in age from 1-10 years. A better way to capture the ageing effect would be to follow models from a particular year (e.g. 2001) over time (e.g. for 10 years, with ratings captured annually). Perhaps an even better approach may be to analyze data that would allow us to separate out the effects of age (A), period (P) and cohort (C) using an APC design (O'Brien, 2000). Here we would track models from different years (e.g. models made in 2001, 2002, 2003, etc., would be the different cohorts) over time (the age variable), using ratings from CR's annual Auto Issue which is published every April (the period variable). Such a design would allow the researcher to better understand how model's quality ratings change with age.

There are several ways in which the research reported in our study can be extended. For example, we discussed how the frequency of redesigns varies across Japanese and US auto firms; it would be interesting to examine a link between redesign frequency and automobile quality. Automobile firms routinely discontinue older models and introduce new models. A number of hypotheses could be examined here – e.g., is it the poor quality of a model and the presence of better quality models within its market segment that forces a firm to discontinue a model? Does the quality of newly introduced models differ across Japanese and US firms? Is there a link between changes in a model's quality and the changes in its market share over time? We also believe that our quality measure is more objective than the measures used in prior studies and thus adds to our findings. Given the 17 dimensions to the measure (Table 1), it is also possible to focus on quality changes (or more broadly, quality differences) in key areas – e.g. engine problems, transmission problems, etc. Such an analysis may help determine the major sources of quality differences across firms.

REFERENCES

Argote, L. 1999. Organizational Learning:Cr eating, Retaining and Transferring Knowledge. Kluwer Academic Publishers, Norwell, MA.

Barber, B. M., and Darrough, M. N., 1996. Product reliability and firm value: The experience of American and Japanese automakers, 1973-1992. Journal of Political Economy, Vol. 104, No. 5, pp. 1084-1099.

Bliese, P. D., and Ployhart, R. E. 2002. "Growth Modeling Using Random Coefficients Models: Model Building, Testing and Illustrations", *Organizational Research Methods,* (5: 4), pp. 362-387.

Carroll, G. R., and Swaminathan, A., 2000. Why the microbrewery movement? Organizational dynamics of resource partitioning in the US brewing industry. American Journal of Sociology, Vol. 106, No. 3, pp. 715-762.

Conlon E., Devaraj, S., and Matta, K. F, 2001. The relationship between initial quality perceptions and maintenance behavior: The case of automotive industry.

Devaraj, S., Matta, K. F., and Conlon, E., 2001. Product and service quality: The antecedents of customer loyalty in the automotive industry. Production and Operations Management Journal, Vol. 10, No. 4, pp. 424-439.

Dobrev, S., Kim, T.-Y., and Hannan, M. T., 2001. Dynamics of niche width and resource partitioning. American Journal of Sociology, Vol. 106, pp. 1299–1337.

Dyer, J. H., 1996. Specialized networks as a source of competitive advantage: Evidence from the auto industry. Strategic Management Journal, Vol. 17, pp. 271-291.

Fine, C. H., Lafrance, J. C., and Hillebrand, D., 1996. The US automobile manufacturing industry. US Department of Commerce, Office of Technology Policy.

Fujimoto, T., and Takeishi, A., 2001. Automobiles: Strategy-based lean production system. Working paper.

Haunschild, P. R., and Rhee, M. (2004). The role of volition in organizational learning: The case of automotive product recalls. Management Science, Vol. 50, No. 11, pp. 1545-1560.

Ingram, P., J. A. C. Baum. 1997. Opportunity and constraint: Organizational learning from operating and competitive experience. Strategic Management Journal, Volume 18, (Summer Special Issue), 75–98.

Jones, R., 2006. Japanese automakers sweep US rivals. MSNBC.com.

Lapre, M. A., and Tsikriktsis, N., 2006. Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. Management Science, Vol. 52, No. 3, pp. 353-366.

Levin, D. Z., 2000. Organizational learning and the transfer of knowledge: An investigation of quality improvement. Organization Science, Vol. 11, No. 6, pp. 630-647.

Liker, J. K., and Choi, T. Y., 2004. Building deep supplier relationships. Harvard Business Review, December 2004, pp. 104-113.

MacDuffie, J. P., 1997. The road to "root cause": Shop-floor problem-solving at three auto assembly plants. Management Science, Vol. 43, No. 4, pp. 479-502.

MacDuffie, J. P., and Pil, F. K., 1996. Performance findings of the international assembly plant. International Motor Vehicle Program, MIT Press.

O'Brien, R. M., 2000. Age period cohort characteristics models. Social Science Research, Vol. 29, pp. 123-139.

Raudenbush, W. and Bryk, S. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*, Thousand Oaks, CA: Sage Publications.

Rhee, M., and haunschild, P. R., (2006). The liability of good reputation: A study of product recalls in the US automobile industry. Oraganization Science, Vol. 17, No. 1, pp. 101-117.

Skinner, W. 1974. The focused factory. Harvard Business Review, Vol. 52, No. 3, 113–121.

Valdes-Depana, P., 2006. Are American Cars Really That Bad? CNN.com.Autos.

Welch, D., 2005. Running out of gas. Business Week (March 28, 2005).