Supply Contracts for On-Time Delivery: the Case of U.S. Influenza Vaccine Market

(Authors' names blinded for peer review)

This paper examines a supply chain consisting of one manufacturer and one retailer that faces three sources of uncertainty: delivery, demand, and design. If the manufacturer starts production after product design has been finalized, the delivery of products to consumers can be late, leading to loss of demand. One option for the manufacturer to improve its delivery performance is to produce prior to the completion of product design, which is vulnerable to the risk that the final design can differ from the projected design. Without a properly designed supply contract, however, the manufacturer lacks the motivation to improve the on-time delivery performance, which leads to shrinkage of the effective demand, and, in turn, incentivises the retailer to choose a low ordering quantity that further discourages the manufacturer's delivery-improving efforts. To break this vicious cycle that erodes the supply chain's delivery performance, a well-designed supply contract should motivate the manufacturer to choose an optimal mix of early and regular production by trading off delivery advantage of early production against informational advantage of regular production. In this work, we evaluate different forms of contracts, and construct new forms of coordinating contracts that are reported in practice but not documented in the extant literature. We also show how a menu of screening contracts enables the retailer to screen the manufacturer type and maximize its profit when there is information asymmetry about on-time delivery performance.

Last updated: April 13, 2012

There is no such uncertainty as a sure thing.
— Robert Burns (1759–1796)

1. Introduction

Behind many conventional products are the myriad unconventional business challenges. When reflecting on the vaccine industry, James Matthews of Sanofi Pasteur observes, "Even though the seasonal influenza vaccine is considered a conventional vaccine by the industry, new challenges with respect to timing and availability of strains and the composition of the influenza vaccine are the rule" (Matthews 2006). These challenges have profound influence on the supply chain performance, and in particular, can lead to major shortage can occur even when the supply is abundant. As an illustrative example, influenza coverage recorded a decline to 41% in the 2000–2001 influenza season, compared to 57% in the previous season; in the meanwhile, 7.5 million vaccine doses, or 10.6% of the total supply, remained unused by the end of the season (Nowalk et al 2005; O'Mara et al 2003). Fukuda et al (2002) explains this seemingly paradoxical situation as follows,

> "The availability of influenza vaccine [in 2000 and 2001] was significantly lower during [October and November] than in previous years, which left many clinicians and patients unable to find vaccine and led to the cancellation of many vaccination campaigns. Ironically, in both years, increasing supplies of vaccine became available in December, but the waning levels of demand resulted in substantial surpluses of unused vaccine."

One option for vaccine manufacturers to improve their delivery performance is to produce prior to the announcement of the influenza vaccine composition. This option, however, is vulnerable to the risk that the final design can differ from manufacturers' projected design. Without a properly designed supply contract, however, manufacturers lack the motivation to improve the on-time delivery performance, which leads to shrinkage of the effective demand, and, in turn, induces health care providers to choose low ordering quantities that further discourages manufacturers' delivery-improving efforts. To break this vicious cycle that erodes the supply chain's delivery performance, vaccine manufacturers and health care providers have adopted various supply contracts. Table 1 provides a representative sample of contract terms used by two major influenza vaccine manufacturers during three recent consecutive influenza seasons. We use fictitious company names for protection of anonymity but all the contract terms are real.

Table 1 Sample contract terms

Manufacturer	Season	Contract terms
X	2010-2011	A proportion of unused doses can be returned for full credit: • doses shipped before October 15: up to 25% of the doses; • doses shipped after October 15: up to 50% of the doses.
	2009-2010	A proportion of unused doses can be returned for full credit: • doses shipped before October 15: up to 25% of the doses; • doses shipped after October 15: up to 50% of the doses.
	2008-2009	A proportion of unused doses can be returned for full credit: • doses shipped before November 15: up to 25% of the doses; • doses shipped after November 15: up to 50% of the doses.
Y	2010-2011	No returns are allowed; No rebate for late-delivered items.
	2009-2010	A 10% rebate is provided for prebook orders shipped after September 30.
	2008-2009	No returns are allowed; No rebate for late-delivered items.

The contract terms used by Manufacturer X resemble the well-studied Quantity Flexibility (QF) contract, except that the return allowance depends on the timing of delivery. The type of contract is unconventional in the sense that, to our best knowledge, it has not been studied in the literature. In contracts used by Manufacturer Y during the 2009–2010 season, by contract, a rebate is provided for late-delivered items. These supply contracts motivate us to build an analytical framework to explain the prevalence of such complex supply contract terms in the influenza vaccine industry. Furthermore, in analyzing various supply contracts, we observe that manufacturers (e.g., Manufacturer Y) often change the type and specification of their contract terms over different years, suggesting that they are adjusting their incentive contracts for better outcomes. Therefore, another aim of this research is to shed insights on improving the existing supply contracts. In examining the business environment associated with these supply contracts, we identify these key sources of uncertainties: delivery, demand, and design, and our model captures the following aspects, as illustrated by the case of the U.S. influenza vaccine market.

- The product design is exogenous to the manufacturer. The composition of influenza vaccine is updated year after year based on "which influenza virus strains are circulating, how they are spreading, and how well current vaccine strains protect against newly identified strains" (CDC 2011b). Thus, the timing for determining the composition cannot be moved ahead to the will of individual vaccine manufacturers.
- The demand is time-sensitive. There exists an ideal demand period during which consumers are willing to make return visits in case that the product becomes temporarily unavailable; after the ideal demand period, however, a proportion of customers no longer desire the product. The ideal period for influenza vaccine administration is traditionally between early October and late November, after which the demand rapidly declines, as vaccine providers struggle to convince patients to receive vaccine despite the invariant benefits of late season vaccination (CDC 2011a).
- The leadtime required for manufacturing and distributing the product is long and uncertain. In the influenza vaccine industry, due to the long and complex process of production, testing, releasing and distribution, manufacturers have to make their production decisions way in advance of the demand season, but their delivery of vaccine can still be delayed.
- Uncertainty in demand is not necessarily lessened over time. In the case of influenza vaccine, neither the manufacturer nor the retailer gains advantage in demand information over time.

Demand for the influenza vaccine is largely influenced by various vaccination guidelines that are usually determined the availability of the vaccine.

- Manufacturers have the option to produce before the design uncertainty is resolved. As a common practice, influenza vaccine manufacturers often produce ahead of the first VRBPAC meeting in attempting to meet their tight delivery schedules (VRBPAC 2007, pp. 102–103). Despite its evident benefits, early production can mean that a proportion, if not the whole, of the strains will have to be discarded, in case that a manufacturers' predicted vaccine composition differs from the finalized one.

While incorporating the above considerations, we also evaluate the performance of various supply contracts. We list the abbreviations of a few contracts in Table 2. We show that three well-know conventional contracts, namely wholesale contract, buyback contract, and QF contract, fails to coordinate the supply chain before they do not adequately penalize the manufacturer for late delivery. A few complex contracts, including BCS, BR, and D-QF contracts, have not been reported in the literature but can achieve satisfactory performance in coordinating the supply chain.

Table 2 Abbreviation of Supply Contracts

Abbreviation	Contract
QF	Quantity flexibility contract
BCS	Buyback and cost-sharing contract
BR	Buyback and late rebate contract
D-QF	Delivery-time-dependent quantity flexibility contract

We provide an analytical framework that explains the prevalence of such complex supply contracts. When the demand uncertainty is the only source of uncertainty, the literature has shown that the simple wholesale contract or its variation such as buyback performs well. However, when there are uncertainties in design and delivery in conjunction with demand uncertainty as in the flu vaccine industry, we show that simple conventional contracts do not work as well as complex coordinating contracts; for those complex contracts, we aim to determine their conditions for coordination, and derive the optimal contract parameters. Our analysis helps explain, while extant literature often focuses on one specific form of supply contracts, supply contracts for products like influenza vaccine often exhibits as a hybrid of simple contracts.

Although our study is motivated by the influenza vaccine industry, our model and analysis apply to other industries of similar characteristics. For example, the apparel and footwear industry is described by Chad Jackson of Aberdeen Group as "a difficult place to run a profitable business" in which "the trends shift quickly and unpredictably, the deadlines are short and there are harsh consequences for delays." On-time delivery is a crucial success factor in this industry because "apparel and footwear buying cycles are transitioning into shorter seasons. The right product must be on the rack to be sold in the right season" (Jackson 2009). A manufacturer, often located overseas, may start producing a certain product early to ensure on-time delivery, but the design or style may not be on trend for the selling season; or the manufacturer may product at a later stage to make sure the product will be trendy, but the delivery time may be uncertain. These similar tradeoffs can be addressed by the analytical framework built in this paper.

2. Literature

The design of supply contracts is one central issue in supply chain management. For the state of the art of this field, readers are referred to Cachon (2003) and the papers discussed therein. Here we briefly discuss several papers that are closely relevant to ours. Tsay (1999), among the first to analyze the QF contract, characterizes the impact of QF contract parameters on the production and forecasting decisions of a supplier and a buyer in a supply chain faced by uncertain demand. Donohue (2000) models the problem of coordinating a manufacturer and distributor of fast fashion products through wholesale contracts. Donohue's work is closely related to ours in that the manufacturer has two production modes with three key differences. First, in our paper, the motivation of early production comes from design uncertainty, whereas in Donohue's case it is for cost advantage. Second, different from one of Donohue's key assumptions, our model does not require early production to be more costly than the regular production, which is consistent with our observation from the influenza vaccine industry. Third, our paper focuses on analyzing a variety of complex contracts observed in the practice but studied in the literature. Cachon and Zhang (2006) consider the sourcing problem of a buyer whose operating costs are affected by both the procurement price and delivery lead time. They characterize the optimal procurement mechanisms and identify two simpler but effective strategies. We also aim to design contracts to improve the delivery performance, but under a different contractual environment.

The design of supply contracts is shaped by various dynamics of the business environment. Taylor (2001) shows that, in a declining price environment, the joint usage of price protection, midlife returns, and end-of-life returns contracts leads to both coordination and a win-win outcome for a supply chain consisting of one manufacturer and one retailer. Taylor (2002) considers a channel coordination problem and shows that a properly-designed target rebate and returns contract coordinates the supply chain while using linear rebates, buyback, or target rebates alone cannot. The two papers are closely related to our study in that we also analyze various complex contracts adopted in practice. Taylor and Plambeck (2007) consider a scenario in which the supplier must commit to the production capacity of a product under development due to a long leadtime. They characterize the conditions under which the buyer should commit to the price only, or both the price and the buying quantity. Taylor and Xiao (2009) examine a channel coordination problem wherein the retailer can improve the sales forecasting by exerting costly efforts and shows that the buyback contract is always superior to the rebate contract because the former is better at separating retailers observing different signals.

As our work is motivated by supply contracts we collect from the U.S. influenza vaccine industry, it adds to a new perspective compared to the thin but growing literature on influenza vaccine operations management. Chick et al. (2008) develop the first influenza vaccine supply chain model with two players: a social-welfare-maximizing government, and a profit-maximizing manufacturer. They show that a variation of the cost-sharing contract can coordinate the supply chain. Cho (2010) models the VRBPAC's problem of updating the vaccine composition under dynamic information, and characterize the optimal decision rule that significantly improves the social welfare. Deo and Corbett (2009) build a model of the Cournot competition between influenza vaccine manufacturers with endogenous entry, which allows them to calibrate its impact on consumers' welfare. All of the three papers focus on the decisions of either the government or the manufacturer, or both. Our paper, by contrast, investigate the channel coordination problem between a manufacturer and a retailer and, to our best knowledge, the first to highlight the impact of uncertainties in delivery and composition on the manufacturer's production decisions.

The issue of on-time delivery has been approached in the literature from various angles. Grout and Chisty (1993) analyze purchasing contracts in a just-in-time production setting where on-time delivery is an essential performance metric to the buyer but can be directly controlled by the supplier. Under delivery time uncertainty only, they show that a bonus scheme improves the on-time delivery performance without analyzing the efficiency of the buyer-supplier relationship. Stecke and Zhao (2007) consider the problem of jointly managing production and transportation that is faced by a make-to-order manufacturer with a commit-to-delivery business mode. Their proposed mixed-integer programming model minimizes the total shipping cost for accepted orders when there exist multiple shipping modes with varying costs. Katok et al. (2008) investigate service-level agreements where the supplier commits to a predetermined fraction of demand met per review period. They analyze the effect of two key contract parameters: the length of the review period, and the bonus for meeting the service-level target, and find that longer review periods can be more effective in improving the service performance.

3. Modeling Framework

In this section, we describe the basic assumptions, settings, and sequence of events of our model in §3.1, and then characterize the first-best solution in §3.3.

3.1 Model Description

We consider a supply chain, consisting of one manufacturer and one retailer, that faces three sources of uncertainties: design, delivery, and demand. Due to uncertainties in demand and delivery, the retailer is subject to shrinking total demand due to unreliable delivery. We refer to the loss of demand due to delay in delivery as *functional shortage* (cf. O'Mara et al 2003). The manufacturer, on the other hand, is faced with uncertainties in design and delivery, and in response, chooses to operate in a hybrid of two production modes.

The retailer faces uncertain, time-sensitive demand for a single product that is realized during an *ideal period*. During the ideal demand period, consumers are willing to make return visits in case that the product is temporarily unavailable. After the ideal demand period, however, a proportion of unfulfilled demand will be permanently lost. The manufacturer has an uncertain production leadtime and thus cannot guarantee to deliver the products during the ideal period. The demand for the product is denoted by d, which has a probability density of $f(\cdot)$ and a cumulative density of $F(\cdot)$. Finally, we define $\gamma \in [0,1]$ as the proportion of consumers that will not return to the retailer if their requests for the product are not satisfied before the ideal period ends. Thus, γ captures the time-sensitivity of demand.

The manufacturer operates in dual production modes: regular and early. Under the *regular production mode*, the manufacturer has an uncertain production leadtime and cannot always deliver the product in a timely fashion. With probability $\alpha \in [0,1]$, the delivery is on time (during the ideal period); with probability $(1-\alpha)$, the delivery is late (after the ideal period). It incurs c_r per unit for the manufacturer to operate in the regular production mode. The manufacturer also has an *early production mode*, in which the manufacturer starts production before the product design is finalized. The early production mode guarantees on-time delivery but is subject to design uncertainty: with probability $\beta \in [0,1]$, the early production is successful; with probability $(1-\beta)$, the early production fails. It incurs c_e per unit to operate under the regular production mode. Throughout the paper, we assume that that $c_e \geqslant \beta c_r$, which includes the special case that $c_e = c_r$. Under the dual-production-mode setting, the manufacturer needs to decide the early production quantity, denoted by Q_e, as well as the regular production quantity, denoted by Q_r, where r and e are the respective subscripts for the regular and the early production modes. Table 3 highlights the key differences between the regular and the early production modes.

Table 3 Comparison of the Two Production Modes

	Regular production mode	Early production mode
Delivery uncertainty	On-time with prob. α	Always on-time
Design uncertainty	Always successful	Successful with prob. β
Unit Production cost	c_r	c_e

The dual production mode in our model is motivated by the influenza vaccine industry. The early production mode enables vaccine manufacturers to eliminate delay in delivering influenza vaccine at the cost of introducing design uncertainty: manufacturers might have to abandon their whole batches in the early production if the finalized vaccine composition differs from their prediction. Due to the inherent risk of uncertain vaccine composition, manufacturers only allocate a small proportion of chicken eggs for early production, and reserve the majority of their eggs to the regular production.

We consider the following sequence of events that specifies the order of various production, delivery, and distribution operations:

- $t = 1$ (before the production starts): the retailer places an order Q to the manufacturer, which incurs an administrative cost of c_o per unit on the retailer's side.
- $t = 2$ (right before the early production): the manufacturer begins to produce $Q_e \leq Q$ units in the early production mode.
- $t = 3$ (right before the regular production): The product design is finalized. If the predicted design matches the finalized design, the manufacturer proceeds to begin producing $Q_r = Q - Q_e$ by the regular mode; otherwise, the manufacturer chooses $Q_r = Q$. Here we assume that the manufacturer operates under a *forced compliance regime* such that the manufacturer must supply exactly as the retailer orders (cf. Cachon and Lariviere 2001), although the delivery might be delayed.
- $t = 4$ (the end of the ideal period): If the design used in the early production coincides with the finalized one (with probability β), then the products from the early mode are shipped to the retailer to meet the demand before the end of the ideal period. In addition, with probability α, the products from the regular production mode are delivered during the ideal period to meet the demand.
- $t = 5$ (after the ideal period): Of the customers customers whose demand is unfulfilled during the ideal period, if any, $(1 - \gamma)$ of them return to the retailer. The products from regular production mode, with probability $(1 - \alpha)$, are delivered after the ideal period to meet the remaining unsatisfied demand.

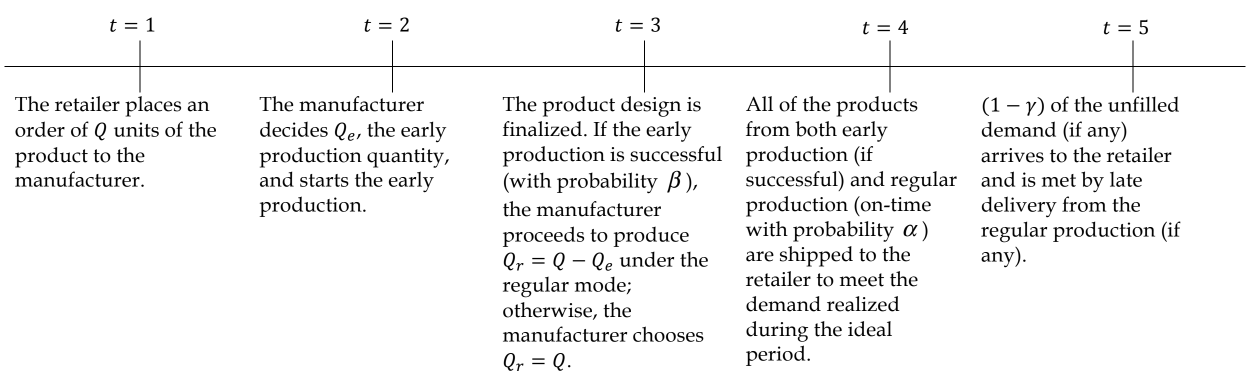

Figure 1 Sequence of events in the dual-production-mode setting

In the above sequence of operations, as depicted by Figure 1, two decisions are pivotal: the retailer's order size Q, and the manufacturer's early production quantity Q_e. Given Q and Q_e, the regular production quantity Q_r is still uncertain depending on whether the predicted design matches the finalized one:

$$Q_r = \begin{cases} Q - Q_e & \text{with probability } \beta \\ Q & \text{with probability } (1 - \beta), \end{cases}$$

and, given the demand d, the sales quantity falls into one of the following three possible cases:

$$Z(Q,Q_e|d) := \begin{cases} \min\{Q,d\} & \text{with probability } \alpha \\ \min\{Q_e,d\} & \\ \quad +\min\{Q-Q_e,(1-\gamma)(d-Q_e)^+\} & \text{with probability } (1-\alpha)\beta \\ \min\{Q,(1-\gamma)d\} & \text{with probability}(1-\alpha)(1-\beta), \end{cases} \tag{1}$$

where $(d-Q_e)^+ = \max\{d-Q_e,0\}$ denotes the positive part of $d-Q_e$. In the first case, the regular production is on-time (with probability α); in the second case, the regular production is late and the early production matches the finalized design; in the last case, the regular production is late and the early production does not match the finalized design. The expected sales quantity is thus:

$$\begin{aligned} E[Z(Q,Q_e)] = E[E[Z(Q,Q_e|d)]] & \\ = \alpha \left[\int_0^Q \xi dF(\xi) + \int_Q^\infty Q dF(\xi)\right] & \\ + (1-\alpha)\beta \left[\int_0^{Q_e} \xi dF(\xi) + \int_{Q_e}^{\frac{Q-\gamma Q_e}{1-\gamma}} [\gamma Q_e + (1-\gamma)\xi] dF(\xi) + \int_{\frac{Q-\gamma Q_e}{1-\gamma}}^\infty Q dF(\xi)\right] & \\ + (1-\alpha)(1-\beta) \left[\int_0^{\frac{Q}{1-\gamma}} (1-\gamma)\xi dF(\xi) + \int_{\frac{Q}{1-\gamma}}^\infty Q dF(\xi)\right]. & \end{aligned} \tag{2}$$

It is worthy noting that Q_e only appears in the second term of the right-hand side of (2). This shows that the early production is most valuable when the regular production yields late delivery but the firm's predicted design coincides with the finalized one.

LEMMA 1. *(i)* $\partial E[Z(Q,Q_e)]/\partial Q_e \geq 0$.
(ii) $\partial E[Z(Q,Q_e)]/\partial Q_e$ *increases in* Q.
(iii) $\partial^2 E[Z(Q,Q_e)]/\partial Q_e^2 \leq 0$.

Proof. We prove Part (i) of the lemma by noting that

$$\partial E[Z(Q,Q_e)]/\partial Q_e > 0 = (1-\alpha)\beta\gamma \left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right] \geq 0$$

because

$$\frac{Q-\gamma Q_e}{1-\gamma} \geq \frac{Q_e-\gamma Q_e}{1-\gamma} = Q_e.$$

We then prove Part (ii) by taking the derivative of $\partial E[Z(Q,Q_e)]/\partial Q_e$ with respect to Q:

$$\partial^2 E[Z(Q,Q_e)]/\partial Q_e \partial Q = \frac{\beta\gamma(1-\alpha)}{1-\gamma} \cdot f\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) \geq 0.$$

Finally, we take the second-order derivative of $E[Z(Q,Q_e)]$ in terms of Q_e:

$$\partial^2 E[Z(Q,Q_e)]/\partial Q_e^2 = -\frac{\beta\gamma(1-\alpha)}{1-\gamma} \cdot \left[(1-\gamma)\cdot f(Q_e) + \gamma \cdot f\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right] \leq 0,$$

which proves Part (iii). □

Lemma 1 suggests that the expected selling quantity $E[Z(Q,Q_e)]$ increases in the early production quantity Q_e, and the marginal return is diminishing in Q_e. Interestingly, the marginal benefit from early production is amplified when the ordering quantity Q increases. In other words, a lower ordering quantity Q not just directly lowers the revenue due to the reduced supply level, but also has a second-order effect in that it hurts the delivery advantage of the early production quantity.

3.2 Production Yield

Throughout the paper we suppress the consideration of random production yield. In the influenza vaccine industry, the phenomenon of random yield is well-documented; see Chick et al. (2008) and Cho (2010) for examples. The issue is further complicated by the fact that, as opposed to most of the extant literature, different strains can have varying random yield distributions: Cho (2010) estimate that two strains share a mean yield rate of 0.9, yet one strain has a standard deviation of 0.1, another 0.4. By suppressing random yield, we assume that delivery uncertainty is a more interesting predictor for the supply chain performance than random yield. In fact, this consideration is supported by the vast vaccine policy literature (e.g., O'Mara et al. 2003 Fukuda et al. 2002). Therefore, incorporating delivery uncertainty, in conjunction with design and demand uncertainties, captures the essence of the influenza vaccine supply chain. By doing so, our model also naturally extends to other markets, for example, the apparel and footwear industry, where there is no yield uncertainty.

3.3 First-Best Solution

We now consider the first-best scenario in which supply chain is vertically integrated such that it total expected profit is maximized by jointly determining the ordering quantity Q and the early production quantity Q_e. Let p denote the unit retail price of the product to end consumers. The first-best solution, denoted by (Q^{FB}, Q_e^{FB}), maximizes the supply chain profit:

$$\pi_S(Q,Q_e) = pE[Z(Q,Q_e)] - [(c_r + c_o)Q + (c_e - \beta c_r)Q_e], \tag{3}$$

where the subscript S refers to the supply chain consisting of the manufacturer and the retailer, $pE[Z(Q,Q_e)]$ is the expected proceeds from selling the products, and $(c_r + c_o)Q + (c_e - \beta c_r)Q_e$ is the expected production costs, which is derived from

$$E[c_eQ_e + (c_r + c_o)Q_r] = c_oQ + \beta \underbrace{[c_eQ_e + c_r(Q - Q_e)]}_{\text{Early mode succeeds}} + (1-\beta)\underbrace{[c_eQ_e + c_rQ]}_{\text{Early mode fails}}.$$

In the first-best solution, it might be the case that the manufacturer operates under regular production mode only, i.e., $Q_e^{FB} = 0$. The next proposition gives the condition for the first-best early-production quantity Q_e^{FB} to be positive:

LEMMA 2. *$Q_e^{FB} > 0$ if and only if*

$$\alpha F(Q_s^{FB}) + \frac{1-\alpha}{1-\gamma} \cdot Q_s^{FB} \cdot f\left(\frac{Q_s^{FB}}{1-\gamma}\right) + \frac{c_e - \beta c_r}{\beta\gamma p} < 1 - \frac{c_r + c_o}{p}, \tag{4}$$

where Q_s^{FB} is the first-best ordering quantity when there is only a single production mode (cf. Lemma 5 in Appendix A).

Proof. We start with the single-production mode scenario (c.f. Appendix A) in which the only decision is the retailer's ordering quantity. The first-best ordering quantity Q_s^{FB} is unique and satisfies (50) by the first-order condition.

Now we introduce the early production mode. When the early production quantity Q_e is very small, the marginal benefit is close to zero. We evaluate the first-order derivative of the supply chain profit in terms of Q_e close to zero:

$$\begin{aligned}\partial\pi_S(Q,Q_e)/\partial Q_e|_{Q_e\to 0} &= -(c_e - \beta c_r) + p\left[\gamma Q_e - \frac{\gamma(Q-\gamma Q_e)}{1-\gamma} f\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) + \gamma F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right]\Bigg|_{Q_e\to 0} \\ &= p\left[\gamma \cdot F\left(\frac{Q}{1-\gamma}\right) - \frac{\gamma Q}{1-\gamma} \cdot f\left(\frac{Q}{1-\gamma}\right)\right] - (c_e - \beta c_r).\end{aligned}$$

It is necessary to introduce early production mode, i.e., to set $Q_e > 0$ if and only if the $\partial \pi_S(Q, Q_e)/\partial Q_e$ is positive as $Q_e \to 0$ and $Q = Q_s^{FB}$. The sufficient and necessary condition for $Q_e > 0$ is

$$\gamma \cdot F\left(\frac{Q_s^{FB}}{1-\gamma}\right) - \frac{\gamma Q_s^{FB}}{1-\gamma} \cdot f\left(\frac{Q_s^{FB}}{1-\gamma}\right) > \frac{c_e - \beta c_r}{p}. \tag{5}$$

We rewrite (51) in Appendix A as

$$F\left(\frac{Q_s^{FB}}{1-\gamma}\right) = 1 - \frac{\frac{c_r + c_o}{p} - \alpha \bar{F}\left(Q_s^{FB}\right)}{1-\alpha}. \tag{6}$$

Substituting (6) into (5) yields (4). $\square$

The risk of the early production mode outweighs its benefit when (4) in Lemma 2 is violated. This is likely to happen under (1) a low γ, meaning that the product is not highly time-sensitive, (2) a high expected production cost difference $c_e - \beta c_r$, (3) a low β, which means that uncertainty in product design is high, and (4) a high α, which means that the regular production mode yields satisfactory on-time delivery performance. In other words, it is beneficial to the supply chain to introduce the early production mode when the delivery from the regular production mode is unreliable, the early production mode is inexpensive, and the manufacturer's prediction for the finalized design is sufficiently accurate.

In the rest of the analysis, we focus on the interesting case wherein $Q_e^{FB} > 0$. We consider the decentralized supply chain model under different supply contracts, where the manufacturer and the retailer maximize their own profits. The question of interest is whether the first-best profit can be achieved under each contract. A contract is said to *coordinate* the supply chain when the first-best profit is achieved. We start with analyzing three representative conventional contracts in the next section, followed by their derivatives that coordinate the supply chain under different conditions.

4. Evaluation of Conventional Contracts

In this section, we evaluate the performance of several conventional contracts. In §4.1, we show that under the wholesale contract, the manufacturer has no incentive to operate in early production mode. In §4.2, we show that the buyback contract gives the manufacturer incentive to operate in the early mode, but still cannot coordinate the supply chain. In §4.3, we consider the quantity flexibility (QF) contract that has been extensively studied in the literature but not applied in the influenza vaccine industry. While the literature establishes that QF contract can conditionally coordinate the supply chain with a single production mode, we prove that it cannot coordinate the supply chain with dual production mode.

4.1 Wholesale Contract

We first consider a wholesale contract that specifies the unit wholesale price w that the retailer pays to the manufacturer. The manufacturer's profit under the wholesale contract is the revenue from selling Q units of products less the costs of producing Q_r units in the regular mode and Q_e units in the early mode:

$$\begin{aligned}\pi_M^W(Q, Q_e) &= wQ - c_r E[Q_r] - c_e Q_e \\ &= wQ - c_r[\beta(Q - Q_e) + (1-\beta)Q] - c_e Q_e \\ &= (w - c_r)Q - (c_e - \beta c_r)Q_e.\end{aligned} \tag{7}$$

The proposition below evaluates the wholesale contract.

Proposition 1. *Under the wholesale contract and dual production modes, the manufacturer has no incentive to operate in the early production mode, that is, the manufacturer always chooses* $Q_e = 0$ *and* $Q_r = Q$.

Proof. Under the wholesale contract, the manufacturer's expected profit given Q is given by (7). We see that $\partial\pi_M^W(Q, Q_e)/\partial Q_e = -(c_e - \beta c_r) < 0$. Hence the manufacturer chooses $Q_e = 0$. □

Proposition 1 says that, for any ordering quantity Q, the manufacturer has no incentive to operate in the early production mode because the marginal net surplus from early production is negative. In consequence, the manufacturer only operates in the regular production mode, and the problem is reduced to the single-production-mode scenario (Appendix A).

4.2 Buyback Contract

We have shown that the wholesale contract does not coordinate the supply chain, which is not surprising given the phenomenon of double marginalization. Now we evaluate the buyback contract in which the manufacturer is partially responsible for the costs of the left-over inventory.[1] We show in Appendix A that a simple buyback contract can overcome double marginalization and thus coordinate the supply chain when the manufacturer can only operate in the regular production mode. Can the buyback contract still coordinate the supply chain when the manufacturer has the option to produce before the design is finalized? To answer the question, we note that when there is a single production mode, the pivotal decision is the retailer's ordering quantity Q. As long as the retailer places the first-best order (Q_s^{FB}), the whole supply chain is coordinated. When there are two production modes, however, the manufacturer's choice of the early production quantity Q_e also impacts the supply chain performance. To coordinate the supply chain, the buyback contract must induce (Q^{FB}, Q_e^{FB}) in equilibrium.

Under the buyback contract, the manufacturer pays the retailer b for each unit of unsold product. The manufacturer's total buyback cost is therefore $b \cdot (Q - E[Q_e)])$. The manufacturer's expected profit under this contract is

$$\begin{aligned}\pi_M^B(Q, Q_e) &= (w - c_r)Q - (c_e - \beta c_r)Q_e - b \cdot \{Q - E[Z(Q, Q_e)]\} \\ &= (w - c_r - b)Q - (c_e - \beta c_r)Q_e + bE[Z(Q, Q_e)]. \end{aligned} \tag{8}$$

The retailer's expected profit under this contract is

$$\begin{aligned}\pi_R^B(Q, Q_e) &= pE[Z(Q, Q_e)] - (w + c_o)Q + b \cdot \{Q - E[Z(Q, Q_e)]\} \\ &= (p - b)E[Z(Q, Q_e)] - (w + c_o - b)Q. \end{aligned} \tag{9}$$

Let us suppose that under the buyback contract, the retailer orders Q^{FB}, and derive the manufacturer's early production quantity. Given $Q = Q^{FB}$, the manufacturer's decision is to choose Q_e that maximizes

$$\pi_M^B(Q^{FB}, Q_e) = (w - c_r - b)Q^{FB} - (c_e - \beta c_r)Q_e + bE[Z(Q^{FB}, Q_e)]. \tag{10}$$

Since Q^{FB} is given, only the last two terms in (10) impact the manufacturer's decision. To be more specific, the manufacturer weighs between the risk of production failure, captured by the term $(c_e - \beta c_r)Q_e$, and the benefit of reducing buybacks, captured by the term $bE[Z(Q^{FB}, Q_e)]$.

On the other hand, in the first-best scenario, given Q^{FB}, the manufacturer sets Q_e that maximizes $pE[Z(Q^{FB}, Q_e)] - (c_e - \beta c_r)Q_e$ (cf. (3)), which differs from $bE[Z(Q^{FB}, Q_e)] - (c_e - \beta c_r)Q_e$ under

[1] The analysis of revenue-sharing contract is of equal theoretic interest, but is ignored here due to its difficulty of implementation of the influenza vaccine market. Nevertheless, the analysis of buyback as well as BCS and BR contracts can extend to the revenue-sharing contract and its variants.

the buyback contract. Since $b < p$,[2] the manufacturer's marginal benefit from operating in the early production mode is lower than in the first-best scenario, although the marginal cost from operating in the early production mode is the same. Therefore, the manufacturer would set $Q_e < Q_e^{FB}$. This proves the following proposition.

PROPOSITION 2. *A buyback contract cannot coordinate the supply chain with two production modes.*

REMARK 1. Under a buyback contract, the first-best cannot be achieved because the manufacturer is not provided with adequate incentive to operate under the early production mode and hence improve the delivery performance. A suboptimal delivery performance leads to a shrunken demand and induces the retailer to choose an ordering quantity Q that is below Q^{FB}. This, by Lemma 1, again hurts the manufacturer's incentive to operate under the early production mode because the manufacturer's marginal return is lower due to a lower ordering quantity. In consequence, a vicious cycle emerges under the buyback contract.

4.3 QF Contract

As another practical way to overcoming double marginalization, the QF contract is a widely adopted contract form in which the manufacturer provides the retailer with full credit for the leftover inventory up to a pre-determined threshold. While there exist different specifications of the threshold, we focus on the QF contract where the threshold is a proportion of the ordering quantity, a practical setting commonly assumed in the QF literature (e.g., Tsay 1999, Cachon and Lariviere 2001, and Plambeck and Taylor 2002). We show in this section that the QF contract cannot coordinate the supply chain with dual production modes.

We denote by $\kappa \in (0,1)$ the proportion of the retailer's ordering quantity Q that is allowed to return after the sales season. The returning quantity, denoted by $R(Q,Q_e)$, is equal to the left-over inventory, or κ of the total ordering quantity Q, whichever is lower, that is,

$$R(Q,Q_e) = \min\{\kappa Q, Q - Z(Q,Q_e)\},$$

where the sales quantity $Z(Q,Q_e)$ is presented in (1), and $Q - Z(Q,Q_e)$ is the leftover inventory. The total transfer payment from the manufacturer to the retailer, denoted by $T_c(Q,Q_e)$, is thus

$$T_c(Q,Q_e) := w \cdot R(Q,Q_e),$$

since the manufacturer provides full credit for all the returnable products.

To facilitate our subsequent analysis, we define an indicator I such that

$$I = \begin{cases} 1 & \text{if products from the regular production are delivered on time,} \\ 0 & \text{otherwise.} \end{cases}$$

In our setting, $\Pr(I=1) = \alpha$ and $\Pr(I=0) = 1-\alpha$. Similarly, we define another indicator J such that

$$J = \begin{cases} 1 & \text{if the early production matches the finalized design,} \\ 0 & \text{otherwise.} \end{cases}$$

We have $\Pr(J=1) = \beta$ and $\Pr(J=0) = 1-\beta$. Depending on the realizations of I and J, the sales quantity $Z(Q,Q_e)$ and the returning quantity $R(Q,Q_e)$ take different values as shown below.

[2] Note that it is infeasible to set $b = p$ because the retailer can then profit from leftover inventory.

Case 1. $I=1$ and $J=0$ or 1, that is, products from the regular mode are delivered on time. In this case, Q units of products are shipped during the ideal vaccination period. Hence the total sales quantity is

$$Z(Q,Q_e|I=1)=\begin{cases}\xi & \text{if } 0\le\xi<Q\\ Q & \text{if } \xi\ge Q.\end{cases}$$

By comparing the leftover inventory $Q-Z(Q,Q_e)$ with the return allowance κQ, we represent the total returning quantity as

$$R(Q,Q_e|I=1)=\begin{cases}\kappa Q & \text{if } 0\le\xi<(1-\kappa)Q\\ Q-\xi & \text{if } (1-\kappa)Q\le\xi<Q\\ 0 & \text{if } \xi\ge Q.\end{cases} \tag{11}$$

Case 2. $I=0$ and $J=1$, i.e., the regular-mode products are not delivered on time, and the design used in early production mode coincides with the finalized product specification. In this case, Q_e units of products are shipped during the ideal vaccination period, and $Q-Q_e$ units after the ideal period. The sales quantity is

$$Z(Q,Q_e|I=0,J=1)=\begin{cases}\xi & \text{if } 0\le\xi<Q_e\\ Q_e+(1-\gamma)(\xi-Q_e) & \text{if } Q_e\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ Q & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases}$$

The realization of the returning quantity, however, depends on the range of κ:
(a) If $0<\kappa<1-Q_e/Q$, then $Q_e<(1-\kappa)Q$, or $\kappa Q<Q-Q_e$, meaning that the late-delivered products outnumbers the return allowance. By comparing κQ and $Q-Z(Q,Q_e)$, we have

$$R(Q,Q_e|I=0,J=1)=\begin{cases}\kappa Q & \text{if } 0\le\xi<\frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}\\ Q-\gamma Q_e-(1-\gamma)\xi & \text{if } \frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases} \tag{12}$$

(b) If $1-Q_e/Q\le\kappa<1-\gamma Q_e/Q$, then $Q_e\ge(1-\kappa)Q$. When $Q_e\le\xi\le(Q-\gamma Q_e)/(1-\gamma)$, we have

$$\begin{aligned}\kappa Q-[Q-Z(Q,Q_e)]&=\gamma Q_e+(1-\gamma)\xi-(1-\kappa)Q\\&\ge Q_e-(1-\kappa)Q\ge 0,\end{aligned}$$

that is,

$$\min\{\kappa Q,Q-Z(Q,Q_e)\}=Q-Z(Q,Q_e)=Q-\gamma Q_e-(1-\gamma)\xi.$$

The returning quantity can therefore be written as

$$R(Q,Q_e|I=0,J=1)=\begin{cases}\kappa Q & \text{if } 0\le\xi<(1-\kappa)Q\\ Q-\xi & \text{if } (1-\kappa)Q\le\xi<Q_e\\ Q-\gamma Q_e-(1-\gamma)\xi & \text{if } Q_e\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases} \tag{13}$$

(c) If $1-\gamma Q_e/Q\le\kappa<1$, then $[(1-\kappa)Q-\gamma Q_e]/(1-\gamma)<0$, which follows

$$R(Q,Q_e|I=0,J=1)=\begin{cases}\kappa Q & \text{if } 0\le\xi<(1-\kappa)Q\\ Q-\xi & \text{if } (1-\kappa)Q\le\xi<Q_e\\ Q-\gamma Q_e-(1-\gamma)\xi & \text{if } Q_e\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases} \tag{14}$$

Case 3. $I=0$ and $J=0$. This is the worst case in which all the units produced in the regular mode are not delivered on time, and the predicted design does not match the finalized one. In this case, no products are shipped during the ideal vaccination period, and all of the ordered products are shipped after the ideal period to meet $(1-\gamma)$ of the total demand. The sales quantity is

$$Z(Q,Q_e|I=0,J=0)=\begin{cases}(1-\gamma)\xi & \text{if } 0\le\xi<\frac{Q}{1-\gamma}\\ Q & \text{if } \xi\ge\frac{Q}{1-\gamma}.\end{cases}$$

The returning quantity is

$$R(Q,Q_e|I=0,J=0)=\begin{cases}\kappa Q & \text{if } 0\le\xi<\frac{(1-\kappa)Q}{1-\gamma}\\ Q-(1-\gamma)\xi & \text{if } \frac{(1-\kappa)Q}{1-\gamma}\le\xi<\frac{Q}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q}{1-\gamma}.\end{cases} \tag{15}$$

Next, we show that the QF contract cannot coordinate the supply chain because the manufacturer is not provided with adequate incentive to operate in the early production mode.

PROPOSITION 3. *The QF contract cannot coordinate the supply chain.*

Proof. Under the QF contract, the manufacturer's objective function is

$$\pi_M^{QF}(Q,Q_e)=(w-c_r)Q-(c_e-\beta c_r)Q_e-E[T_c(Q,Q_e)].$$

In the above equation, $(c_e-\beta c_r)$ is the manufacturer's marginal production cost in the early production mode, while the term $E[T_c(Q,Q_e)]$ is the expected transfer payment the retailer receives from the manufacturer. Next, we analyze how the contract influences the manufacturer's production decisions when κ takes vales from different ranges.

Case (a). $0\le\kappa<1-Q_e/Q$: The expected units of returns, from (11), (12), and (15), can be represented as

$$\begin{aligned}E[R(Q,Q_e)]=&\alpha\left[\int_0^{(1-\kappa)Q}\kappa Q dF(\xi)+\int_{(1-\kappa)Q}^{Q}(Q-\xi)dF(\xi)\right]\\&+(1-\alpha)\beta\left[\int_0^{\frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}}\kappa Q dF(\xi)+\int_{\frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}}^{\frac{Q-\gamma Q_e}{1-\gamma}}\left[Q-\gamma Q_e-(1-\gamma)\xi\right]dF(\xi)\right]\\&+(1-\alpha)(1-\beta)\left[\int_0^{\frac{(1-\kappa)Q}{1-\gamma}}\kappa Q dF(\xi)+\int_{\frac{(1-\kappa)Q}{1-\gamma}}^{\frac{Q}{1-\gamma}}\left[Q-(1-\gamma)\xi\right]dF(\xi)\right].\end{aligned}$$

The first-order derivative of $E[T_c]$ is

$$\partial E[T_c(Q,Q_e)]/\partial Q_e=w\cdot\partial E[R(Q,Q_e)]/\partial Q_e=w(1-\alpha)\beta\gamma\left[F\left(\frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}\right)-F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right],$$

which is negative since $\kappa\le1$, meaning that the manufacturer's transfer payment decreases in Q_e. In other words, the manufacturer expects a lower returning quantity by increasing the production quantity under the early production mode. Hence the marginal benefit from producing in the early mode can be captured by $(-\partial E[T_c(Q,Q_e)]/\partial Q_e)$. We note from (3) that the marginal benefits of operating in the early production mode to the supply chain is

$$p\cdot\partial E[Z(Q,Q_e)]/\partial Q_e=p(1-\alpha)\beta\gamma\left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)-F(Q_e)\right]. \tag{16}$$

Note that $0 \le \kappa < 1 - Q_e/Q$ is equivalent to $Q_e < (1-\kappa)Q$, which gives

$$Q_e < (1-\kappa)Q = \frac{(1-\kappa)Q - \gamma(1-\kappa)Q}{1-\gamma} < \frac{(1-\kappa)Q - \gamma Q_e}{1-\gamma},$$

and hence

$$F(Q_e) < F\left(\frac{(1-\kappa)Q - \gamma Q_e}{1-\gamma}\right).$$

Therefore we have

$$\begin{aligned} p \cdot \partial E[Z(Q,Q_e)]/\partial Q_e > & \, w\partial E[Z(Q,Q_e)]/\partial Q_e \\ & = w(1-\alpha)\beta\gamma \left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right] \\ & > w(1-\alpha)\beta\gamma \left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F\left(\frac{(1-\kappa)Q-\gamma Q_e}{1-\gamma}\right)\right] \\ & = -\partial E[T_c(Q,Q_e)]/\partial Q_e \text{ for any combination}(Q,Q_e). \end{aligned} \tag{17}$$

It follows from (17) that the marginal benefit of operating in the early production mode to the manufacturer is always strictly lower than that to the supply chain. Therefore, the QF contract cannot provide the manufacturer with adequate incentive to operate in the early production mode even if the retailer places an order that matches the first-best quantity.

Cases (b) and (c). We combine the two cases (Case (b): $1 - Q_e/Q \le \kappa < 1 - \gamma Q_e/Q$ and Case (c): $1 - \gamma Q_e/Q \le \kappa < 1$) because the returning quantity, enumerated by (13) and (14), are exactly the same under the two scenarios. We have from $1 - Q_e/Q \le 1$ that $Q_e \ge (1-\gamma)Q$. Using (11), (13), and (15), the expected units of returns can be expressed as

$$\begin{aligned} & E[R(Q,Q_e)] \\ & = \alpha\left[\int_0^{(1-\kappa)Q} \kappa Q dF(\xi) + \int_{(1-\kappa)Q}^{Q} (Q-\xi) dF(\xi)\right] \\ & + (1-\alpha)\beta\left[\int_0^{(1-\kappa)Q} \kappa Q dF(\xi) + \int_{(1-\kappa)Q}^{Q_e} (Q-\xi) dF(\xi) + \int_{Q_e}^{\frac{Q-\gamma Q_e}{1-\gamma}} \left[Q - \gamma Q_e - (1-\gamma)\xi\right] dF(\xi)\right] \\ & + (1-\alpha)(1-\beta)\left[\int_{Q_e}^{\frac{(1-\kappa)Q}{1-\gamma}} \kappa Q dF(\xi) + \int_{\frac{(1-\kappa)Q}{1-\gamma}}^{\frac{Q}{1-\gamma}} \left[Q - (1-\gamma)\xi\right] dF(\xi)\right]. \end{aligned}$$

The first-order derivative of $E[T_c]$ in terms of Q_e is

$$\partial E[T_c(Q,Q_e)]/\partial Q_e = w \cdot \partial E[R(Q,Q_e)]/\partial Q_e = w(1-\alpha)\beta\gamma\left[F(Q_e) - F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right]. \tag{18}$$

We now compare the marginal benefit of early production to the manufacturer, captured by (18), with that to the supply chain, captured by (16):

$$-\partial E[T_c(Q,Q_e)]/\partial Q_e = w/p \cdot \partial E[Z(Q,Q_e)]/\partial Q_e < \partial E[Z(Q,Q_e)]/\partial Q_e,$$

meaning that the contract does not provide enough incentive to the manufacturer to operate in the early production mode.

To summarize Cases (a), (b), and (c), the QF contract cannot coordinate the supply chain. □

Remark 2. Similar to the buyback contract (cf. Remark 1), the QF contract does not coordinate the supply chain not just because the manufacturer is not provided with adequate incentive to operate in the early production mode, but also because the retailer's order size decreases, which further reduces the manufacturer's incentive to exert delivery-performance-improving efforts.

The extant supply chain literature establishes that the QF contract can coordinate a supply chain with a single production mode when a set of conditions are met (Cachon 2003, §6.2.5). In the presence of the dual production mode, Proposition 3 states that the QF contract can *never* coordinate the supply chain. This loosely explains the rare application of the QF contract in the vaccine market.

5. Coordinating Contracts

In the preceding section, we have shown that under the three sources of uncertainties and dual production modes, it no longer feasible to apply conventional contracts such as buyback and QF contracts. The underlying reason is that although these contracts are effective in overcoming double marginalization under a single production mode, they fail to break the vicious cycle formed as a result of a complex web of incentives. In this section, we describe three different types of contracts, namely, Buyback-and-Cost-Sharing (BCS) contract (§5.1), Buyback-and-Rebate (BR) contract (§5.2), and Delivery-time-dependent Quantity Flexibility (D-QF) contract (§5.3), that coordinate the supply chain under varying conditions. We derive the optimal parameters for each of these contracts. §5.4 compares and contrasts these coordinating contracts. §5.5 numerically verifies our main insights.

5.1 BCS Contract

We have shown that the buyback contract cannot coordinate the supply chain because it does not provide the manufacturer with adequate incentive to operate in the early production mode. One way to address the supply coordination problem is to increase the retailer's share in the risks due to design uncertainty. In this section, we propose a BCS contract that complements the buyback contract with cost sharing. Under the BCS contract, the manufacturer provides the retailer with partial credits for unsold items as in a buyback contract, while the retailer shares a proportion of the manufacturer's costs incurred in the early production.[3] To be more specific, the manufacturer pays the retailer b for each unit of leftover inventory; in the meantime, the retailer subsidizes the manufacturer k for each unit produced in the early production mode, regardless of whether the production is successful or not. Here we make the underlying assumption that the manufacturer's early production quantity Q_e is observable to the retailer; we will relax this assumption in §5.2. Under the BCS contract, the retailer's expected profit is

$$\begin{aligned}\pi_R^{BCS}(Q,Q_e) &= p\cdot E[Z(Q,Q_e)] - (w+c_o)\cdot Q + b\cdot\{Q - E[Z(Q,Q_e)]\} - k\cdot Q_e \\ &= (p-b)\cdot E[Z(Q,Q_e)] - (w+c_o-b)\cdot Q - k\cdot Q_e,\end{aligned} \tag{19}$$

and the manufacturer's expected profit is

$$\begin{aligned}\pi_M^{BCS}(Q,Q_e) &= (w-c_r)\cdot Q - (c_e-\beta c_r)\cdot Q_e - b\cdot\{Q - E[Z(Q,Q_e)]\} + k\cdot Q_e \\ &= (w-c_r-b)Q - (c_e-\beta c_r-k)\cdot Q_e + b\cdot E[Z(Q,Q_e)].\end{aligned} \tag{20}$$

By comparing (20) with (10), one can observe that, under the BCS contract, while the manufacturer's marginal benefit from operating in the early mode remains the same as in the buyback

[3] Chick et al. (2008) consider a cost-sharing contract under which the buyer (government) subsidizes the influenza vaccine manufacturer for each egg used in the production, and show that the cost-sharing contract coordinates the supply chain under a single production mode. However, one can show that the cost-sharing contract fails to coordinate the supply chain under the dual production mode.

contract, the manufacturer's marginal risk from early production is reduced by k per unit of the product.

The derivation of the optimal BCS contract consists of the following two steps:
Step 1. As an initial step, we suppose that the retailer orders exactly Q^{FB}, and choose the contract parameters to ensure that the manufacturer would respond by setting $Q_e = Q_e^{FB}$. To implement the first-best solution, we need to make sure that the manufacturer's risk-benefit tradeoff is equivalent to that under the first-best scenario, i.e.,

$$\frac{b}{c_e - \beta c_r - k} = \frac{p}{c_e - \beta c_r} \implies k = \frac{p-b}{p} \cdot (c_e - \beta c_r), \tag{21}$$

where b and $c_e - \beta c_r - k$ are the coefficients of $E[Z(Q,Q_e)]$ and Q_e in (20), and p and $c_e - \beta c_r$ are are the coefficients of $E[Z(Q,Q_e)]$ and Q_e in (3). Therefore, (21) ensures that the manufacturer weighs costs and benefits of the early production as if in the first-best scenario.
Step 2. Once we have set the cost-sharing parameter k as a function of the buyback price b according to (21), the manufacturer would then choose Q_e^{FB} as long as the retailer orders Q^{FB}. We then need to check that the retailer chooses an ordering quantity of Q^{FB}. Comparing the coefficients of the terms involving Q in (19) against those in (3), we see that the buyback price b needs to satisfy

$$\frac{p-b}{p} = \frac{w + c_o - b}{c_r + c_o}. \tag{22}$$

Expressions (21) and (22) jointly lead to the following proposition.

Proposition 4. *A Buyback-and-Cost-Sharing (BCS) contract with*

$$b^*_{BCS} = \frac{w - c_r}{p - c_r - c_o} \cdot p \text{ and } k^*_{BCS} = \frac{p - w - c_o}{p - c_r - c_o} \cdot (c_e - \beta c_r)$$

coordinates the supply chain with two production modes.

Note that the optimal BCS contract parameters are independent of the time-sensitivity of demand (γ). The intuition follows from our procedures of deriving the optimal contract, which are based on comparisons of cost parameters associated with each quantity items (e.g., the ordering quantity Q, the expected selling quantity $E[Z(Q,Q_e)]$). While γ affects the expected selling quantity, it does not affect the cost structure. Hence the optimal parameters of BCS contract is independent of the time-sensitivity parameter γ. Based on Proposition 4, we obtain the following corollary that details the impact of α and β on the contract parameters.

Corollary 1. *Under the BCS contract,*

	b^*_{BCS}	k^*_{BCS}	*R's profit share*	*M's profit share*
α	-	-	-	-
β	-	$\downarrow$	-	-
γ	-	-	-	-
c_e	-	$\uparrow$	-	-

The optimal BCS contract uses exactly the same buyback price as in the single-production-mode scenario (cf. Lemma 6 in Appendix A), and the buyback price b^*_{BCS} is independent of both α and β. The cost-sharing parameter k^*_{BCS}, tied to the early production mode only, depends on β but not on α. As β increases, the manufacturer's risk of losing early production batch is lower, so the early production becomes a more viable option, and the required cost-sharing level from the retailer is thus lower.

5.2 BR Contract

Although a BCS contract can coordinate the supply chain, it requires the manufacturer's early production quantity to be observable to the retailer. To satisfy this condition requires the retailer's additional monitoring efforts, without which the retailer can observe Q_e only with probability of $(1-\alpha)\beta$ (when the early production is successful but the regular production is late). This somewhat restrictive assumption partially explains why BCS has not been reportedly adopted in the influenza vaccine industry.

We now propose another variant of the buyback contract, BR contract, that combines buyback contract and rebate contract. This contract is based on two quantities observable to both the manufacturer and the retailer: the leftover inventory, and the late-delivery quantity. Thus, under the BR contract the manufacturer's early production decision does not have to be observable to the retailer. The manufacturers provides the retailer with a rebate for each late-delivered unit of product, in addition to providing the retailer with a buyback credit for each unsold unit. We use ρ to denote the proportion of wholesale price that manufacturer provides to the retailer for late delivery, and $\rho \cdot w$ is the rebate for each unit of product not delivered on time. Thus, the expected transfer payment from the manufacturer to the retailer is represented as

$$\begin{aligned} & b\cdot\{Q-E[Z(Q,Q_e)]\}+\rho w\left[(1-\alpha)\beta(Q-Q_e)+(1-\alpha)(1-\beta)Q\right] \\ &= b\cdot\{Q-E[Z(Q,Q_e)]\}+\rho w(1-\alpha)(Q-\beta Q_e). \end{aligned} \tag{23}$$

In (23), the first term is the manufacture's expected buyback credit to the retailer; the second term is the manufacturer's expected rebate, in which the quantity of late delivery is $Q-Q_e$ with probability $(1-\alpha)\beta$, is Q with probability $(1-\alpha)(1-\beta)$, and is 0 with probability α. The manufacturer's profit is its revenue from selling to the retailer less the transfer payment to the retailer, including both credits for leftover inventory and rebate for late-delivered items:

$$\begin{aligned} \pi_M^{BR}(Q_e,Q) &= (w-c_r)Q-(c_e-\beta c_r)Q_e-b\{Q-E[Z(Q,Q_e)]\}-\rho w(1-\alpha)(Q-\beta Q_e) \\ &= b\cdot E[Z(Q,Q_e)]+[w-c_r-b-\rho w(1-\alpha)]\cdot Q-[c_e-\beta c_r-\rho w\beta(1-\alpha)]\cdot Q_e, \end{aligned} \tag{24}$$

and the retailer's profit is its revenue plus the transfer payment to the retailer:

$$\begin{aligned} \pi_R^{BR}(Q,Q_e) &= p\cdot E[Z(Q,Q_e)]-(w+c_o)Q+b\cdot\{Q-E[Z(Q,Q_e)]\}+\rho w(1-\alpha)(Q-\beta Q_e) \\ &= (p-b)\cdot E[Z(Q,Q_e)]-[w+c_o-b-\rho w(1-\alpha)]Q-\beta\pi(1-\alpha)Q_e. \end{aligned} \tag{25}$$

The procedure of determining the optimal contract parameters is similar to that in §5.1. Under the following two conditions the proposed BR contract coordinates the supply chain:

$$\frac{b}{c_e-\beta c_r-\rho w\beta(1-\alpha)}=\frac{p}{c_e-\beta c_r}, \text{ and} \tag{26}$$

$$\frac{p-b}{w+c_o-b-\rho w(1-\alpha)}=\frac{p}{c_r+c_o}. \tag{27}$$

The condition given in (26) ensures that the manufacturer responds with choosing an early production quantity of Q_e^{FB} whenever the retailer chooses an ordering quantity of Q^{FB}, and the condition given in (27) ensures that it is optimal for the retailer to order Q^{FB}. Jointly solving (26) and (27) yields the optimal supply contract presented in the following proposition:

PROPOSITION 5. *A Buyback-and-Rebate (BR) coordinates the supply chain with two production modes when*

$$b_{BR}^*=\frac{\beta w-c_e}{\beta(p-c_o)-c_e}\cdot p, \text{ and } \rho_{BR}^*=\frac{(p-w-c_o)(c_e-\beta c_r)}{w(1-\alpha)[\beta(p-c_o)-c_e]}.$$

The following corollary provides comparative statics to show the impact of α and β on the optimal BR contract parameters.

COROLLARY 2. *Under the BR contract,*

	b^*_{BR}	ρ^*_{BR}	*R's profit share*	*M's profit share*
α	-	$\uparrow$	-	-
β	$\uparrow$	$\downarrow$	$\uparrow$	$\downarrow$
γ	-	-	-	-
c_e	$\uparrow$	$\uparrow$	$\uparrow$	$\downarrow$

Corollary 2 has three interesting implications. First, as the manufacturer's on-time delivery performance (measured by α) improves, one might expect that the manufacturer would promise a lower rebate for late delivery to the retailer; Corollary 2 states that the opposite is true. To understand this counterintuitive result, note that the manufacturer provides a rebate level that is positively tied to the production quantity from the regular mode (Q), which is subject to delay in delivery, and negatively tied to the the production quantity from the early mode (Q_e). As α increases, regular production becomes a more viable option. The manufacturer chooses a higher regular production quantity and raises the rebate rate ρ^*_{BR} at the same time. Second, as β increases, the manufacturer's predicted design becomes more accurate. Thus, it makes sense for the manufacturer to increase the early production quantity (Q_e) and lower the rebate rate (ρ^*_{BR}) accordingly. Third, the buyback price (b^*_{BR}) is in the contract primarily to motivate the retailer to choose the first-best ordering quantity, and is independent of the manufacturer's on-time delivery performance, as captured by α. As β increases, early production becomes less risky, and a higher total production quantity is beneficial to the supply chain. The manufacturer should therefore increase the buyback price b^*_{BR} to encourage the retailer to increase the order size.

5.3 D-QF Contract

We now examine the D-QF contract in which the return allowance depends on the timing of delivery as well as the total ordering quantity. This type of contract has not been reported in the supply chain coordination literature but is adopted in influenza vaccine industry. Our analysis helps reveal the benefits of the D-QF contract relative to the well-studied QF contract. The contract is specified as follows. Let Y_1 be the shipping quantity by the end of the ideal period, and Y_2 the shipping quantity after the ideal period. The maximum returning quantity, referred to as the returning allowance, is thus equal to $\kappa_1 Y_1 + \kappa_2 Y_2$, where $\kappa_1, \kappa_2 \in [0,1]$ are the respective returnable proportions of the delivery quantities for on-time and late-delivered units of products. Hence the return allowance is a random variable and can be represented as

$$\kappa_1 Y_1 + \kappa_2 Y_2 = \begin{cases} \kappa_1 Q & \text{if } I = 1 \\ \kappa_1 Q_e + \kappa_2 (Q - Q_e) & \text{if } I = 0 \text{ and } J = 1 \\ \kappa_2 Q & \text{if } I = 0 \text{ and } J = 0. \end{cases}$$

We denote by $R_d(Q, Q_e)$ the total returning quantity under the D-QF contract:

$$R_d(Q, Q_e) = \min\{\kappa_1 Y_1 + \kappa_2 Y_2, Q - Z(Q, Q_e)\}.$$

The transfer payment from the manufacturer to the retailer at the end of the demand season is given by $T_d(Q, Q_e) := w \cdot R_d(Q, Q_e)$. Comparing the return allowance $\kappa_1 Y_1 + \kappa_2 Y_2$ with the leftover inventory $Q - Z(Q, Q_e)$, we can express the returning quantity for different production and demand realizations as follows.

Case 1. $I=1$ and $J=0$ or 1. In this case, $Y_1=Q$ and $Y_2=0$. Therefore,

$$R_d(Q,Q_e|I=1)=\begin{cases}\kappa_1 Q & \text{if } \xi<(1-\kappa_1)Q\\ Q-\xi & \text{if } (1-\kappa_1)Q\le\xi<Q\\ 0 & \text{if } \xi\ge Q.\end{cases} \tag{28}$$

Case 2. $I=0$ and $J=1$. We have $Y_1=Q_e$ and $Y_2=Q-Q_e$. The returning quantity depends on different combinations of (κ_1,κ_2). Two critical values, $[(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e]/(1-\gamma)$ and $(Q-\gamma Q_e)/(1-\gamma)$, are important in the following derivations. The former is the demand level at which the retailer uses up the level of allowable returns, i.e., $Q-\gamma Q_e-(1-\gamma)\xi=\kappa_1 Q_e+\kappa_2(Q-Q_e)$, and the latter is the demand level at which the total supply exactly matches the effective demand, i.e., $Q_e+(1-\gamma)(\xi-Q_e)=Q$.
(a) If $\kappa_2-\kappa_1>1-(1-\kappa_2)Q^{FB}/Q_e^{FB}$, then

$$R_d(Q,Q_e|I=0,J=1)=\begin{cases}\kappa_1 Q_e+\kappa_2(Q-Q_e) & \text{if } 0\le\xi<\frac{(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\\ Q-\gamma Q_e-(1-\gamma)\xi & \text{if } \frac{(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases} \tag{29}$$

We use the inequality $[(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e]/(1-\gamma)>Q_e$ in deriving the above expression. The two subcases $0\le\xi<Q_e$ and $Q_e\le\xi<[(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e]/(1-\gamma)$ can be combined as one because when $0<\xi<Q_e$, $Z(Q,Q_e)=\xi$, which gives

$$\begin{aligned}\kappa_1 Y_1+\kappa_2 Y_2-[Q-Z(Q,Q_e)]&=\kappa_1 Q_e+\kappa_2(Q-Q_e)-(Q-\xi)\\ &<\kappa_1 Q_e+\kappa_2(Q-Q_e)-(Q-Q_e)\\ &=[1-(\kappa_2-\kappa_1)]Q_e-(1-\kappa_2)Q<0.\end{aligned}$$

We arrive at the representation of the returning quantity for the remaining subcases by comparing the leftover inventory and the maximum allowable returning quantity.
(b) If $\kappa_2-\kappa_1\le1-(1-\kappa_2)Q/Q_e$, then

$$R_d(Q,Q_e|I=0,J=1)=\begin{cases}\kappa_1 Q_e+\kappa_2(Q-Q_e) & \text{if } 0\le\xi<(1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e\\ Q-\xi & \text{if } (1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e\le\xi<Q_e\\ Q-\gamma Q_e-(1-\gamma)\xi & \text{if } Q_e\le\xi<\frac{Q-\gamma Q_e}{1-\gamma}\\ 0 & \text{if } \xi\ge\frac{Q-\gamma Q_e}{1-\gamma}.\end{cases} \tag{30}$$

Note that when $Q_e\le\xi<(Q-\gamma Q_e)/(1-\gamma)$,

$$\begin{aligned}\kappa_1 Y_1+\kappa_2 Y_2-(Q-Z(Q,Q_e))&=\kappa_1 Q_e+\kappa_2(Q-Q_e)-[Q-Q_e-(1-\gamma)(\xi-Q_e)]\\ &=[1-(\kappa_1-\kappa_2)]Q_e-(1-\kappa_2)Q+(1-\gamma)(\xi-Q_e)\\ &>[1-(\kappa_1-\kappa_2)]Q_e-(1-\kappa_2)Q\ge0.\end{aligned}$$

The other three subcases follow directly from comparing the leftover inventory and the maximum allowable returning quantity.
Case 3. $I=0$ and $J=0$. In this case, $Y_1=0$ and $Y_2=Q$. Therefore,

$$R_d(Q,Q_e|I=0,J=0)=\begin{cases}\kappa_2 Q & \text{if } I=0,J=0 \text{ and } 0\le\xi<\frac{(1-\kappa_2)Q}{1-\gamma}\\ Q-(1-\gamma)\xi & \text{if } I=0,J=0 \text{ and } \frac{(1-\kappa_2)Q}{1-\gamma}\le\xi<\frac{Q}{1-\gamma}\\ 0 & \text{if } I=0,J=0 \text{ and } \xi\ge\frac{Q}{1-\gamma}.\end{cases} \tag{31}$$

PROPOSITION 6. *A D-QF contract with* $0 \le \kappa_1 < \kappa_2 \le 1$ *can coordinate the supply chain only if one of the following two conditions is satisfied:*

a) $\kappa_2 - \kappa_1 > 1 - (1-\kappa_2)Q^{FB}/Q_e^{FB}$, *and*
$$w(\kappa_2 - \kappa_1 - \gamma) \cdot F\left(\frac{(1-\kappa_2)Q^{FB} + (\kappa_2 - \kappa_1 - \gamma)Q_e^{FB}}{1-\gamma}\right) = (p-w)\gamma \cdot F\left(\frac{Q^{FB} - \gamma Q_e^{FB}}{1-\gamma}\right) - p\gamma \cdot F(Q_e^{FB});$$
b) $\kappa_2 - \kappa_1 \le 1 - (1-\kappa_2)Q^{FB}/Q_e^{FB}$, *and*
$$w(\kappa_2 - \kappa_1) \cdot F\left((1-\kappa_2)Q^{FB} + (\kappa_2 - \kappa_1)Q_e^{FB}\right) = (p-w)\gamma \cdot \left[F\left(\frac{Q^{FB} - \gamma Q_e^{FB}}{1-\gamma}\right) - F(Q_e^{FB})\right].$$

Proof. In each of following two cases, we first derive and evaluate the manufacturer's marginal benefit from engaging in early production for an arbitrary combination of (Q, Q_e), and then apply the analysis to the case that $Q = Q^{FB}$ and $Q_e = Q_e^{FB}$.

(a) If $\kappa_2 - \kappa_1 > 1 - (1-\kappa_2)Q/Q_e$, then the expected returning quantity, from (28), (29), and (31), can be expressed as

$$\begin{aligned}
E[R_d(Q, Q_e)] = \alpha & \left[\int_0^{(1-\kappa_1)Q} \kappa_1 Q dF(\xi) + \int_{(1-\kappa_1)Q}^{Q} (Q-\xi) dF(\xi)\right] \\
& + (1-\alpha)\beta \left[\int_0^{\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}} \left[\kappa_1 Q_e + \kappa_2(Q-Q_e)\right] dF(\xi)\right. \\
& \left. + \int_{\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}}^{\frac{Q-\gamma Q_e}{1-\gamma}} \left[Q - \gamma Q_e - (1-\gamma)\xi\right] dF(\xi)\right] \\
& + (1-\alpha)(1-\beta) \left[\int_0^{\frac{(1-\kappa_2)Q}{1-\gamma}} \kappa_2 Q dF(\xi) + \int_{\frac{(1-\kappa_2)Q}{1-\gamma}}^{\frac{Q}{1-\gamma}} \left[Q - (1-\gamma)\xi\right] dF(\xi)\right].
\end{aligned}$$

The first-order derivative of $E[T_d]$ in terms of Q_e is

$$\begin{aligned}
\partial E[T_d(Q, Q_e)]/\partial Q_e = & w \cdot \partial E[R_d(Q, Q_e)]/\partial Q_e & (32) \\
= & w(1-\alpha)\beta\gamma \left[F\left(\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\right) - F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right] \\
& - w(1-\alpha)\beta(\kappa_2-\kappa_1) F\left(\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\right). & (33)
\end{aligned}$$

If κ_1 is chosen to be no lower than κ_2, then the second term in the right-hand side of (33) is negative or zero. In this case, the marginal benefit from operating in the early mode to the manufacturer, captured by $(-\partial E[T_d(Q, Q_e)]/\partial Q_e)$, is lower than that to the supply chain, captured by $p \cdot \partial E[Z(Q, Q_e)]/\partial Q_e$. To see this, note that

$$\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma} > \frac{[1-(\kappa_2-\kappa_1)]Q_e + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma} = \frac{(1-\gamma)Q_e}{1-\gamma} = Q_e,$$

which gives

$$\begin{aligned}
-\partial E[T_d(Q, Q_e)]/\partial Q_e \le & w(1-\alpha)\beta\gamma \left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F\left(\frac{(1-\kappa_2)Q + (\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\right)\right] \\
< & p \cdot \partial E[Z(Q, Q_e)]/\partial Q_e = p(1-\alpha)\beta\gamma \left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right].
\end{aligned}$$

Therefore, any D-QF contracts with $\kappa_1 \ge \kappa_2$ cannot coordinate the supply chain. However, when we choose $\kappa_2 > \kappa_1$, the second term in (33) is positive, meaning that the D-QF contract provides additional incentives for the manufacturers to operate in the early production mode. To the end, (κ_1, κ_2) must satisfy

$$-\partial E[T_d(Q,Q_e)]/\partial Q_e = p\cdot \partial E[Z(Q,Q_e)]/\partial Q_e, \text{ or}$$

$$w(\kappa_2-\kappa_1-\gamma)\cdot F\left(\frac{(1-\kappa_2)Q+(\kappa_2-\kappa_1-\gamma)Q_e}{1-\gamma}\right) = (p-w)\gamma\cdot F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - p\gamma\cdot F(Q_e) \quad (34)$$

to be able to coordinate the supply chain.

(b) If $\kappa_2-\kappa_1 \le 1-(1-\kappa_2)Q/Q_e$, then the expected returning quantity, from (28), (30), and (31), can be expressed as

$$\begin{aligned} E[R_d(Q,Q_e)] = \alpha &\left[\int_0^{(1-\kappa_1)Q} \kappa_1 Q dF(\xi) + \int_{(1-\kappa_1)Q}^{Q} (Q-\xi) dF(\xi)\right] \\ &+ (1-\alpha)\beta \left[\int_0^{(1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e} \left[\kappa_1 Q_e + \kappa_2(Q-Q_e)\right] dF(\xi)\right. \\ &\left. + \int_{(1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e}^{Q_e} (Q-\xi) dF(\xi) + \int_{Q_e}^{\frac{Q-\gamma Q_e}{1-\gamma}} \left[Q-\gamma Q_e-(1-\gamma)\xi\right] dF(\xi)\right] \\ &+ (1-\alpha)(1-\beta)\left[\int_0^{\frac{(1-\kappa_2)Q}{1-\gamma}} \kappa_2 Q dF(\xi) + \int_{\frac{(1-\kappa_2)Q}{1-\gamma}}^{\frac{Q}{1-\gamma}} \left[Q-(1-\gamma)\xi\right] dF(\xi)\right]. \end{aligned}$$

The first-order derivative of $E[T_d]$ in terms of Q_e is

$$\begin{aligned} \partial E[T_d(Q,Q_e)]/\partial Q_e = & w(1-\alpha)\beta\gamma\left[F(Q_e) - F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right)\right] \\ & - w(1-\alpha)\beta(\kappa_2-\kappa_1)F\left((1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e\right). \end{aligned} \quad (35)$$

If $\kappa_1 \ge \kappa_2$, then the second term in the right-hand side of (35) is negative or zero. In this case, the marginal benefits from operating in the early mode $(-\partial E[T_d(Q,Q_e)]/\partial Q_e)$ would be lower than that to the supply chain $(p\cdot\partial E[Z(Q,Q_e)]/\partial Q_e)$ because

$$\begin{aligned} -\partial E[T_d(Q,Q_e)]/\partial Q_e \le & w(1-\alpha)\beta\gamma\left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right] \\ & < p\cdot\partial E[Z(Q,Q_e)]/\partial Q_e = p(1-\alpha)\beta\gamma\left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right]. \end{aligned}$$

Therefore, a delivery-time-dependent contract with $\kappa_1 \ge \kappa_2$ cannot coordinate the supply chain. Similar to Case (a), we need $\kappa_2 > \kappa_1$ so that the contract can provide extra incentive for the manufacturer to operate under the early production mode. In addition, (κ_1, κ_2) must satisfy

$$-\partial E[T_d(Q,Q_e)]/\partial Q_e = p\cdot \partial E[Z(Q,Q_e)]/\partial Q_e, \text{ or}$$

$$w(\kappa_2-\kappa_1)F\left((1-\kappa_2)Q+(\kappa_2-\kappa_1)Q_e\right) = (p-w)\gamma\left[F\left(\frac{Q-\gamma Q_e}{1-\gamma}\right) - F(Q_e)\right]$$

so that the contract can coordinate the supply chain. □

In contrast to Proposition 3, which shows that the QF contract can never coordinate the supply chain under the dual production mode, Proposition 6 shows that the D-QF contract can conditionally coordinate the supply chain. The underlying reason is that, under the D-QF contract, the manufacturer's incentive to produce in the early mode is positively tied to the return allowance difference $\kappa_2 - \kappa_1$: late delivery incurs economic losses to the manufacturer not only because of lost sales due to the time-sensitiveness of the demand as in the QF contract, but also because the retailer is allowed to return a higher proportion of late-delivered units than that of on-time-delivered units since $\kappa_2 > \kappa_1$.

The intuition behind the proof of Proposition 6 is as follows. There are two contract parameters, κ_1 and κ_2, to specify in the D-QF contract. The optimal contract choice should support the equilibrium where the retailer chooses an ordering quantity of Q^{FB} units, and the manufacturer chooses an early production quantity of Q_e^{FB}. One of the two conditions (i) and (ii) in Proposition 6 must be satisfied to guarantee that the manufacturer choose Q_e^{FB} when the retailer orders Q^{FB}. The condition (i) corresponds to the scenario that the return allowance difference $\kappa_2 - \kappa_1$ is high. The condition (ii) corresponds to the scenario that the allowance difference is relatively low. Note that the QF contract satisfies the first part of the condition (ii), i.e., $\kappa_2 - \kappa_1 = 0$ is small enough, but does not satisfy the second part; the QF contract therefore cannot coordinate the supply chain by Proposition 6.

Proposition 6 provides the necessary conditions for the D-QF contract to coordinate the supply chain; it does not guarantee the existence of a coordinating D-QF contract. To coordinate the supply chain, the contract should also motivate the retailer to choose an ordering quantity (Q) that matches the first-best level (Q^{FB}). Nevertheless, as illustrated by the numerical study in §5.5, the D-QF contract coordinates the supply chain in almost all of the cases except for a few wherein Q_e^{FB} is equal to or close to zero.

5.4 Discussions of BCS, BR, and D-QF Contracts

We have analyzed three contract forms, namely, BCS, BR, and D-QF contracts, that coordinate the supply chain under different conditions. Our findings can be summarized as follows: (i) the BCS contract is coordinating only when the manufacturer's early production is observable to the retailer; (ii) the BR contract can coordinate the supply chain regardless of whether the manufacturer's early production is observable to the retailer or not; (iii) the D-QF contract provides the manufacturer with better incentives than the QF contract to operate in the early production mode, and can coordinate the supply chain under certain conditions; (iv) under all of these three contracts, any arbitrary profit division between the manufacturer and the retailer can be realized by adjusting the wholesale price w.

Table 4 Comparison of BCS and BR Contracts

	BCS contract	BR contract
Optimal parameters	$b^*_{BCS} = \frac{w-c_r}{p-c_r-c_o} \cdot p$, $k^*_{BCS} = \frac{p-w-c_o}{p-c_r-c_o} \cdot (c_e - \beta c_r)$	$b^*_{BR} = \frac{\beta w - c_e}{\beta(p-c_o)-c_e} \cdot p$, $\rho^*_{BR} = \frac{(p-w-c_o)(c_e-\beta c_r)}{w(1-\alpha)[\beta(p-c_o)-c_e])}$
Retailer's profit share	$\frac{p-w-c_o}{p-c_r-c_o}$	$\frac{\beta(p-w-c_o)}{\beta(p-c_o)-c_e}$
Pros	The optimal parameters are independent of α and β	Q_e does not need to be observable to the retailer
Cons	Q_e needs to be observable to the retailer	The optimal parameters depends on α and β

Table 4 compares the two variants of the buyback contracts, namely the BR and BCS contracts, and provides insights into understanding the lack of implementation of BCS contracts in the U.S.

vaccine market. First, it is practically hard for retailers to observe the manufacturer's early production decision. Second, a manufacturer often serves multiple retailers, making it a daunting task to allocate the costs incurred in the early production appropriately among retailers of varying sizes. Third, we can show that under any given contractual environment the retailer's expected profit under the optimal BR contract is always no lower than under the BCS contract by noticing that

$$\frac{\beta(p-w-c_o)}{\beta(p-c_o)-c_e}-\frac{p-w-c_o}{p-c_r-c_o}=\frac{(p-w-c_o)(c_e-\beta c_r)}{(p-c_r-c_o)[\beta(p-c_o)-c_e]}\geq 0$$

as long as $\beta p-c_e>0$, which is true because otherwise, even if the demand is very high, the sales of one unit of product from the early production mode still generates a lower expected revenue (βp) than the its unit cost (c_e). This gives the following corollary:

COROLLARY 3. *The retailer prefers the BR contract to the BCS contract, and therefore has no incentive to collect information about the manufacturer's early production activities.*

The D-QF contract—which has never been studied before in the analytical literature to our best knowledge—finds its extensive application in the U.S. influenza vaccine market. Our analysis confirms its advantage over C-DF contract, which is well-studied but seldom applied in the influenza vaccine industry. We also show D-QF contract's restrictiveness in coordinating the supply chain. Nevertheless, it is even more convenient to implement compared to the BR contract—the manufacturer only needs to specify the two return allowance ratios, namely, κ_1,κ_2, in its contract.

5.5 Numerical Experiments

We obtain managerial insights from numerical experiments. The cost and price parameters, as listed in Table 5, are loosely based on the U.S. influenza vaccine market; see Appendix B for details about the source of the data. Without loss of generality, the demand is assumed to be uniformed distributed between 0 and 1, i.e., $f(\xi)=1$, and $F(\xi)=\xi$, for $0\leq\xi\leq 1$.

For each scenario specified by (α,β,γ), we search over the contract parameter space in the following way: first, for each contract parameter, we solve the Nash equilibrium formed between the manufacturer and the retailer, and obtained the expected profit of the supply chain; second, we compare the supply chain's expected profits under each contract parameters, and identify the optimal contract parameter.

Tables 6 and 7 summarize the respective experimental results for coordinating and conventional contracts. In addition to providing the optimal contract parameters in different environments, we also provide 1) the manufacturer's and the retailer's expected profits, 2) the manufacturer's and the retailer's profit shares relative to the supply chain's total expected profit, 3) the supply chain's efficiency, which is the supply chain's expected profit under each incentive contract divided by that in the first-best scenario, and 4) the supply chain's expected profit in the first-best scenario. We highlight several implications from the experiments results:

- The wholesale contract performs consistently the worst because, in this case, the retailer faces an overage cost ($w=\$12$) than its underage cost ($p-w=\$6$), and therefore always chooses a negative safety stock. The buyback contract eases this issue but can never coordinate the supply chain.
- BCS and BR contracts always coordinate the supply chain, and the BR contract always gives the retailer a larger profit share than the BCS contract does.
- Even though the QF contract has been proven unable to coordinate the supply chain, it can achieve performance comparable to D-QF contract; similar results are observed under other demand distributions, e.g., normal distribution, and triangle distribution. This surprising phenomenon is due to the high margin of the supply chain: with unit production costs of $c_r=c_e=\$3$, the retail price is \$18; the gross margin is above 80%. We also assume that $c_o=0$. An insight from this

Table 5 Cost and price parameters used in numerical experiments

Unit cost, regular (c_r)	Unit cost, early (c_e)	Wholesale price (w)	Retail price (p)
\$3	\$3	\$12	\$18

observation is that the QF contract can work as well as D-QF contract for high-profit-margin products, even though it theoretically cannot coordinate the supply chain.

To verify the advantage of the D-QF contract over QF contract, we conduct sensitivity analysis in terms of the unit production cost. We choose to vary the unit product cost because the retail and wholesale prices can often be obtained through public channels, but the unit production cost is part of the manufacturer's sensitive information; we previously chose \$3 to be consistent with the literature but the actual production cost can be higher due to various overheads and transactions. As shown in Figure 2, the QF contract can achieve near-first-best performance when the unit production cost is low; when the unit cost increases to a certain point, however, under the QF contract the efficiency of the supply chain significantly declines, while Q-DF contract can still coordinate the supply chain.

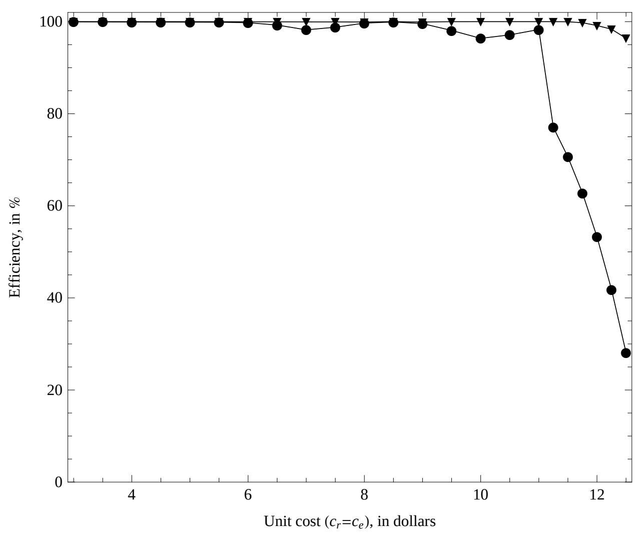

Figure 2 Effect of unit production cost on the supply chain's efficiency under QF and D-QF contracts. We fix $p = \$18, \alpha = 70\%, \beta = 90\%$, and $\gamma = 50\%$. The unit production cost $c_r = c_e$ is chosen from the interval between $\$3$ and $\$12.5$, and the wholesale price is chosen as $w = (3p + 2c_r)/5$.

6. Asymmetric Information About Delivery Performance

Our analyses in the previous sections are based on the assumption that neither the manufacturer nor the retailer has any informational advantage about the production process. In this section, by contrast, we assume that the manufacturer is better informed than the retailer about the probability of on-time delivery when operating under the regular production mode.[4] The delivery performance of the regular production mode is unobservable to the retailer but it is common knowledge that the probability of on-time delivery α belongs to the set $\{\alpha_h, \alpha_l\}$. The manufacturer can be either

[4] We focus on the asymmetric information on the probability of on-time delivery because other information is relatively transparent in the case of U.S. influenza vaccine market: the production costs remain steady over different years and is accessible via multiple sources; the success probability of early production is common to both the manufacturer and the retailer; the retailer does not have any informational advantage about the demand distribution due to the long production cycle.

Table 6 Performance of Coordinating Supply Contracts. Cost and price parameters are from Table 5

Contract Type	α	β	γ	Optimal Parameters		Manufacturer Profit (% share)	Retailer Profit (% share)	Supply Chain Profit (% efficiency)	Supply Chain First-Best Profit
	α	β		b^*_{BCS}	k^*_{BCS}				
BCS	70%	90%	30%[a]	10.8	12%	3.569 (60%)	2.379 (40%)	5.948 (100%)	5.948
	70%	95%		10.8	6%	3.655 (60%)	2.437 (40%)	6.092 (100%)	6.092
	70%	99%		10.8	1.2%	3.730 (60%)	2.487 (40%)	6.217 (100%)	6.217
	80%	90%		10.8	12%	3.590 (60%)	2.394 (40%)	5.984 (100%	5.984
	80%	95%		10.8	6%	3.664 (60%)	2.443 (40%)	6.107 (100%)	6.107
	80%	99%		10.8	1.2%	3.732 (60%)	2.490 (40%)	6.220 (100%)	6.220
	90%	90%		10.8	12%	3.625 (60%)	2.417 (40%)	6.042 (100%)	6.042
	90%	95%		10.8	6%	3.677 (60%)	2.451 (40%)	6.127 (100%)	6.128
	90%	99%		10.8	1.2%	3.734 (60%)	2.489 (40%)	6.223 (100%)	6.223
	α	β		b^*_{BR}	ρ^*_{BR}				
BR	70%	90%	30%[a]	10.63	3.788%	3.515 (59.1%)	2.433 (40.9%)	5.948 (100%)	5.948
	70%	95%		10.72	1.773%	3.629 (59.6%)	2.463 (40.4%)	6.092 (100%)	6.092
	70%	99%		10.79	0.3374%	3.725 (59.9%)	2.492 (40.1%)	6.217 (100%)	6.217
	80%	90%		10.64	5.682%	3.536 (59.1%)	2.448 (40.9%)	5.984 (100%)	5.984
	80%	95%		10.72	2.660%	3.638 (59.6%)	2.469 (40.4%)	6.107 (100%)	6.107
	80%	99%		10.79	0.5061%	3.727 (59.9%)	2.493 (40.1%)	6.220 (100%)	6.220
	90%	90%		10.64	11.36%	3.570 (59.1%)	2.472 (40.9%)	6.042 (100%)	6.042
	90%	95%		10.72	5.319%	3.651 (59.6%)	2.477 (40.4%)	6.128 (100%)	6.128
	90%	99%		10.79	1.012%	3.729 (59.9%)	2.494 (40.1%)	6.223 (100%)	6.223
	α	β	γ	κ^*_1	κ^*_2				
D-QF	70%	90%	10%	66%	78%	3.603 (59.4%)	2.462 (40.6%)	6.065 (100.0%)	6.066
	70%	90%	30%	67%	72%	3.513 (59.1%)	2.433 (40.9%)	5.946 (100.0%)	5.948
	70%	90%	50%	66%	77%	3.454 (59.2%)	2.382 (40.8%)	5.836 (100.0%)	5.837
	70%	95%	10%	68%	69%	3.655 (59.6%)	2.476 (40.4%)	6.131 (99.9%)	6.135
	70%	95%	30%	67%	72%	3.640 (59.8%)	2.452 (40.2%)	6.091 (100.0%)	6.092
	70%	95%	50%	67%	73%	3.605 (59.7%)	2.432 (40.3%)	6.037 (100.0%)	6.037
	70%	99%	10%	68%	69%	3.734 (60.0%)	2.490 (40.0%)	6.224 (100.0%)	6.224
	70%	99%	30%	68%	69%	3.730 (60.0%)	2.487 (40.0%)	6.217 (100.0%)	6.217
	70%	99%	50%	68%	70%	3.724 (60.0%)	2.482 (40.0%)	6.206 (100.0%)	6.206
	80%	90%	10%	100%	100%	2.951 (50.1%)	2.935 (49.9%)	5.886 (96.3%)	6.114[b]
	80%	90%	30%	68%	71%	3.525 (59.0%)	2.453 (41.0%)	5.978 (99.9%)	5.984
	80%	90%	50%	68%	75%	3.453 (58.6%)	2.444 (41.4%)	5.897 (100.0%)	5.899
	80%	95%	10%	69%	70%	3.646 (59.4%)	2.496 (40.6%)	6.142 (99.9%)	6.148
	80%	95%	30%	68%	71%	3.628 (59.4%)	2.478 (40.6%)	6.106 (100.0%)	6.107
	80%	95%	50%	67%	75%	3.624 (59.7%)	2.444 (40.3%)	6.068 (100.0%)	6.068
	80%	99%	10%	68%	69%	3.735 (60.0%)	2.490 (40.0%)	6.225 (100.0%)	6.225
	80%	99%	30%	68%	70%	3.732 (60.0%)	2.488 (40.0%)	6.220 (100.0%)	6.220
	80%	99%	50%	68%	70%	3.727 (60.0%)	2.485 (40.0%)	6.213 (100.0%)	6.213
	90%	90%	10%	100%	100%	3.011 (50.4%)	2.966 (49.6%)	5.977 (96.7%)	6.181[b]
	90%	90%	30%	67%	81%	3.591 (59.4%)	2.451 (40.6%)	6.042 (100.0%)	6.042
	90%	90%	50%	68%	72%	3.515 (58.9%)	2.450 (41.1%)	5.965 (99.9%)	5.971
	90%	95%	10%	49%	49%	3.951 (66.9%)	1.952 (33.1%)	5.903 (95.5%)	6.181[b]
	90%	95%	30%	69%	70%	3.624 (59.2%)	2.500 (40.8%)	6.124 (99.9%)	6.128
	90%	95%	50%	68%	72%	3.625 (59.4%)	2.476 (40.6%)	6.101 (100.0%)	6.102
	90%	99%	10%	68%	68%	3.738 (60.0%)	2.488 (40.0%)	6.226 (100.0%)	6.227
	90%	99%	30%	68%	69%	3.734 (60.0%)	2.489 (40.0%)	6.223 (100.0%)	6.223
	90%	99%	50%	68%	72%	3.731 (60.0%)	2.488 (40.0%)	6.219 (100.0%)	6.219

[a] In the BCS and BR contracts, the optimal contract parameters do not depend on γ. We set $\gamma = 30\%$ to illustrate the profit division between the manufacturer and the retailer.

[b] In these cases, $Q_e^{FB} = 0$.

Table 7 Performance of Conventional Supply Contracts. Cost and price parameters are from Table 5

Contract Type	α	β	γ	Optimal Parameters	Manufacturer Profit (% share)	Retailer Profit (% share)	Supply Chain Profit (% efficiency)	Supply Chain First-Best Profit
Wholesale	α	β						
	70%	90%			2.658 (75%)	0.8861(25%)	3.544 (59.6%)	5.948
	70%	95%			2.658 (75%)	0.8861 (25%)	3.544 (58.2%)	6.092
	70%	99%			2.658 (75%)	0.8861 (25%)	3.544 (57.0%)	6.217
	80%	90%			2.763 (75%)	0.9210 (25%)	3.684 (61.6%)	5.984
	80%	95%	$\gamma = 30\%$	N/A	2.763 (75%)	0.9210 (25%)	3.684 (60.3%)	6.107
	80%	99%			2.763 (75%)	0.9210 (25%)	3.684 (59.2%)	6.220
	90%	90%			2.877 (75%)	0.9589 (25%)	3.836 (63.5%)	6.042
	90%	95%			2.877 (75%)	0.9589 (25%)	3.836 (62.6%)	6.128
	90%	99%			2.877 (75%)	0.9589 (25%)	3.836 (61.6%)	6.223
Buyback[a]	α	β		b_B^*				
	70%	90%		12	3.113 (54.7%)	2.579 (45.3%)	5.692 (95.7%)	5.948
	70%	95%		12	3.215 (55.0%)	2.632 (45.0%)	5.847 (96.0%)	6.092
	70%	99%		12	3.309 (55.4%)	2.662 (44.6%)	5.970 (96.0%)	6.217
	80%	90%		12	3.139 (55.0%)	2.572 (44.0%)	5.711 (95.4%)	5.984
	80%	95%	$\gamma = 30\%$	12	3.225 (55.0%)	2.634 (45.0%)	5.859 (95.9%)	6.107
	80%	99%		12	3.310 (55.4%)	2.663 (44.6%)	5.973 (96.0%)	6.220
	90%	90%		12	3.229 (54.1%)	2.744 (45.9%)	5.973 (98.9%)	6.042
	90%	95%		12	3.240 (55.2%)	2.626 (44.8%)	5.866 (95.7%)	6.128
	90%	99%		12	3.312 (55.4%)	2.663 (44.6%)	5.975 (96.0%)	6.223
QF	α	β	γ	κ^*				
	70%	90%	10%	88%	3.042 (51.6%)	2.852 (48.4%)	5.894 (97.2%)	6.066
	70%	90%	30%	69%	3.476 (58.5%)	2.467 (41.5%)	5.943 (99.9%)	5.948
	70%	90%	50%	69%	3.400 (58.3%)	2.434 (41.7%)	5.834 (100.0%)	5.837
	70%	95%	10%	69%	3.633 (59.3%)	2.497 (40.7%)	6.130 (99.9%)	6.135
	70%	95%	30%	69%	3.559 (59.0%)	2.477 (41.0%)	6.036 (100.0%)	6.092
	70%	95%	50%	67%	3.605 (59.0%)	2.432 (41.0%)	6.037 (100.0%)	6.037
	70%	99%	10%	68%	3.734 (60.0%)	2.489 (40.0%)	6.223 (100.0%)	6.224
	70%	99%	30%	68%	3.730 (60.0%)	2.487 (40.0%)	6.217 (100.0%)	6.217
	70%	99%	50%	68%	3.724 (60.0%)	2.482 (40.0%)	6.206 (100.0%)	6.206
	80%	90%	10%	100%	2.951 (50.1%)	2.935 (49.9%)	5.886 (96.3%)	6.114[b]
	80%	90%	30%	69%	3.507 (58.7%)	2.469 (41.3%)	5.976 (99.9%)	5.984
	80%	90%	50%	69%	3.440 (58.4%)	2.455 (41.6%)	5.895 (99.9%)	5.899
	80%	95%	10%	70%	3.604 (58.7%)	2.536 (41.3%)	6.140 (99.9%)	6.148
	80%	95%	30%	69%	3.606 (59.1%)	2.499 (40.9%)	6.105 (100.0%)	6.107
	80%	95%	50%	69%	3.578 (59.0%)	2.489 (41.0%)	6.067 (100.0%)	6.068
	80%	99%	10%	68%	3.735 (60.0%)	2.489 (40.0%)	6.224 (100.0%)	6.225
	80%	99%	30%	68%	3.732 (60.0%)	2.487 (40.0%)	6.219 (100.0%)	6.220
	80%	99%	50%	68%	3.727 (60.0%)	2.485 (40.0%)	6.212 (100.0%)	6.213
	90%	90%	10%	100%	3.011 (50.4%)	2.966 (49.6%)	5.977 (96.7%)	6.181[b]
	90%	90%	30%	75%	3.397 (56.6%)	2.608 (53.4%)	6.005 (99.4%)	6.042
	90%	90%	50%	69%	3.495 (58.6%)	2.469 (41.4%)	5.964 (99.9%)	5.971
	90%	95%	10%	49%	3.951 (66.9%)	1.952 (33.1%)	5.903 (95.5%)	6.181[b]
	90%	95%	30%	69%	3.625 (59.2%)	2.499 (40.8%)	6.124 (99.9%)	6.128
	90%	95%	50%	69%	3.601 (59.0%)	2.499 (41.0%)	6.010 (100.0%	6.102
	90%	99%	10%	68%	3.738 (60.0%)	2.488 (40.0%)	6.226 (100.0%)	6.227
	90%	99%	30%	68%	3.734 (60.0%)	2.489 (40.0%)	6.223 (100.0%)	6.223
	90%	99%	50%	68%	3.731 (60.0%)	2.487 (40.0%)	6.219 (100.0%)	6.219

[a] In determining the optimal contract, we restrict that the buyback price b_B^* cannot exceed w to be consistent with practice; maximizing the supply chain's expected profit without this restriction will leads to a trivial solution $b_B^* = p$, which is clearly unrealistic in that it always gives the manufacture a negative expected profit.

[b] $Q_e^{FB} = 0$ in these cases.

type h (α_h) or type l (α_l) with respective probabilities δ and $(1-\delta)$. We continue to assume that the retailer cannot observe the manufacturer's early production decision (Q_e) so that the BCS contract—immune from asymmetric information because its optimal parameters are independent of the probability of on-time delivery (cf. Proposition 4)—is an infeasible option.

Symmetric-Information Benchmark. As a benchmark, we start by analyzing the symmetric-information scenario and deriving the optimal supply contract. We focus on BR contracts for its convenience of mapping profit division ratios to specific contract parameters.

We denote by $Z^i(Q,Q_e)$ the sales quantity when the manufacturer is of type i, and $\pi_S^i(Q,Q_e)$ the supply chain's expected profit such that

$$\pi_S^i(Q,Q_e) = pE[Z^i(Q,Q_e)] - [(c_r+c_o)Q + (c_e-\beta c_r)Q_e], s = h, l.$$

Furthermore, we define by π_h^{FB} and π_l^{FB} the respective supply chain profit under the first-best scenario when the manufacturer's type is h and l; we also denote by (Q_h^{FB}, Q_{eh}^{FB}) and (Q_l^{FB}, Q_{el}^{FB}) the respective first-best production quantities under the regular and early modes that maximize the supply chain profit. That is,

$$\pi_i^{FB} = \max_{(Q,Q_e)} \pi_S^i(Q,Q_e), \text{ and}$$
$$(Q_i^{FB}, Q_{ei}^{FB}) = \arg\max_{(Q,Q_e)} \pi_S^i(Q,Q_e), i = h, l.$$

In the following, we define BR(Δ,α) contract as a contract in which the manufacturer has profit share of $0 \le \Delta \le$ and the probability of on-time delivery under regular production mode is α.

Definition 1. A BR(Δ,α) contract is a BR contract for the supply chain with an on-time delivery probability of α, and has the following parameters

$$\begin{aligned} w &= c_r + (p-c_o-c_r)\Delta, \\ b &= \frac{\beta\Delta(p-c_o-c_r)-(c_e-\beta c_r)}{\beta(p-c_o)-c_e}\cdot p, \text{ and} \\ \rho &= \frac{(1-\Delta)(p-c_o-c_r)(c_e-\beta c_r)}{(1-\alpha)[\beta(p-c_o)-c_e][\Delta(p-c_o-c_r)+c_r]}, \end{aligned}$$

Under a BR contract, the manufacturer's profit share is $0 \le \Delta \le 1$.

The retailer offers two different contracts depending on the manufacturer's type. When the manufacturer is of type $i \in \{h,l\}$, the retailer pays a wholesale price of w_i for each unit of product; meanwhile, the manufacturer provides a credit b_i for each unit of leftover inventory, as well as a rebate ρ_i for each unit of late delivered product. The lemma below shows that the supply chain can always be coordinated and the profit division between the manufacturer and the retailer is flexible. Note that the contract does not specify the manufacturer's early production quantity (Q_e) since it is an unobservable quantity to the retailer.

Lemma 3. *Under symmetric information, the retailer can offer two different BR contracts, that is, BR$(\Delta,\alpha_i), i = h, l$, for the two types of manufacturers, respectively, where Δ can be arbitrarily chosen. The contract coordinates the supply chain, that is, in equilibrium, the retailer orders Q_i^{FB} and the manufacturer chooses an early production quantity of Q_{ei}^{FB}.*

Under symmetric information, the supply chain can be coordinated using different wholesale prices and BR contract parameters; in the meantime, the contract can implement any profit division between the manufacturer and the retailer, and a profit division ratio corresponds to a unique supply contract.

Contract Design Under Asymmetric Information. Under asymmetric information, the manufacturer's type is unobservable to the retailer, and the supply contract specified in Lemma 3 is no longer feasible. Instead, we propose a screening mechanism that is implemented by a menu of contracts wherein 1) the first contract is a BR(Δ_h, α_h) contract plus a fixed transfer payment, that is, the manufacturer pays the retailer a fixed amount S, and 2) the second contract is is a BR(Δ_l, α_l) contract. Note that the contract menu does not specify the manufacturer's early production quantity because it is not observable to the retailer.

Following the standard information-economics approach (see, e.g., Laffont and Martimort 2002), we need the following four constraints that are necessary to implement the revelation principle:

$$\begin{aligned} \text{(IC-}h\text{)}: &\quad \Delta_h \pi_h^{FB} - S \geq \Delta_l \pi_{hl}^*, \\ \text{(IC-}l\text{)}: &\quad \Delta_l \pi_l^{FB} \geq \Delta_h \pi_{lh}^* - S, \\ \text{(IR-}h\text{)}: &\quad \Delta_h \pi_h^{FB} - S \geq 0, \\ \text{(IR-}l\text{)}: &\quad \Delta_l \pi_l^{FB} \geq 0, \end{aligned}$$

where

$$\begin{aligned} \pi_{hl}^* &= \max_{Q_e} \pi_S^h(Q_l^{FB}, Q_e), \\ \pi_{lh}^* &= \max_{Q_e} \pi_S^l(\tilde{Q}, Q_e), \text{ and} \\ \tilde{Q} &= \arg\max_{Q} \, [\delta \max_{Q_e} \pi_S^h(Q, Q_e) + (1-\delta) \max_{Q_e} \pi_S^l(Q, Q_e)]. \end{aligned}$$

Constraint (IC-h) is the type h manufacturer's incentive compatibility constraint and ensures that the type h manufacturer is not better off by choosing the second contract; the right-hand side of Constraint (IC-h) is the type h manufacturer's maximum expected profit when choosing the second contract. Constraint (IC-l) is imposed to ensure that the type l manufacturer does not have the incentive to mimic the type h manufacturer by choosing the first contract; without this constraint, the retailer cannot distinguish the two types of manufacturers and therefore chooses an ordering quantity of $\tilde{Q}$ that maximizes the supply chain's expected profit when the manufacturer can be either type h or type l with respective probabilities δ and $(1-\delta)$. Constraint (IR-h) is the type h's individual rationality constraint and can be safely ignored because it follows from Constraint (IC-h). Constraint (IR-l) is the type l's individual rationality constraint and can also be ignored because $\Delta_h \geq 0$. To the end, the above constraints can be can be rewritten as

$$\begin{aligned} \text{(IC-}h\text{)}: &\quad \Delta_h \pi_h^{FB} - S \geq \Delta_l \pi_{hl}^*, \\ \text{(IC-}l\text{)}: &\quad \Delta_l \pi_l^{FB} \geq \Delta_h \pi_{lh}^* - S. \end{aligned}$$

We have the following proposition that establishes the existence of the proposed screening mechanism.

PROPOSITION 7. *Under asymmetric information, there always exists a coordinating screening mechanism—characterized by (Δ_h, S, Δ_l)—that satisfies Constraints (IC-h) and (IC-l) and coordinates the supply chain: in the first contract the manufacturer pays the retailer a fixed amount S in addition to a BR(Δ_h, α_h) contract; the second contract is a BR(Δ_l, α_l) contract.*

Proof. It is sufficient to show that there exists (Δ_h, S, Δ_l) that satisfies

$$\Delta_h \pi_{lh}^* - \Delta_l \pi_l^{FB} \leq S \leq \Delta_h \pi_h^{FB} - \Delta_l \pi_{hl}^*.$$

We note that $\pi_{hl}^* < \pi_h^{FB}$ and $\pi_{lh}^* < \pi_l^{FB}$, which gives

$$\Delta_h \pi_h^{FB} - \Delta_l \pi_{hl}^* > \Delta_h \pi_{hl}^* - \Delta_l \pi_l^{FB}.$$

Therefore, we conclude that one can always construct a mechanism (Δ_h, S, Δ_l) that satisfy both Constraints (IC-h) and (IC-l); as a result, the first-best can always be achieved. $\square$

Proposition 7 establishes the existence of a menu of screening contracts that fully restores the efficiency of the supply chain. Next, we consider a specific objective function and provide the optimal contract parameters. Different from the extant literature where the objective is often assumed to be the retailer's profit, we incorporate here one key consideration that a reasonable division of the supply chain's profit between the manufacturer and the retailer is key to maintaining their relationship. Therefore, the objective of the mechanism is two-fold: on the one hand, the mechanism coordinates the supply chain by motivating the manufacturer to be truth-telling; on the other hand, the mechanism guarantees reasonable profit shares for both the manufacturer and the retailer.

PROPOSITION 8. *Under asymmetric information, if the manufacturer's minimum desirable profit share is Δ, then the retailer's optimal menu of contracts is as follows:*

1. *Under the first contract, the wholesale price $w_1 = p - c_o$, and the manufacturer provides full credits for any left-over inventory; the manufacturer also provides a fixed transfer payment $S = (1-\Delta)\cdot\pi_h^{FB}$ to the retailer.*
2. *The second contract is a $BR(\Delta, \alpha_l)$ contract.*

Proof. We write the retailer's contract design problem as

$$\max_{\Delta_h,\Delta_l,S} \quad \left\{\delta[(1-\Delta_h)\pi_h^{FB} + S] + (1-\delta)(1-\Delta_l)\pi_l^{FB}\right\} \tag{36}$$

subject to

$$\Delta_h\pi_h^{FB} - S \geq \Delta_l\pi_{hl}^*, \tag{37}$$
$$\Delta_l\pi_l^{FB} \geq \Delta_h\pi_{lh}^* - S, \tag{38}$$
$$\Delta_h \leq 1, \tag{39}$$
$$\Delta_l \geq \Delta, \tag{40}$$

which, by substituting $\Delta_l = \Delta$ into the objective function, is equivalent to

$$\max_{\Delta_h,S} \quad \left\{\delta[(1-\Delta_h)\pi_h^{FB} + S] + (1-\delta)(1-\Delta)\pi_l^{FB}\right\} \tag{41}$$

subject to

$$\Delta_h\pi_h^{FB} - S \geq \Delta_l\pi_{hl}^*, \tag{42}$$
$$\Delta_l\pi_l^{FB} \geq \Delta_h\pi_{lh}^* - S, \tag{43}$$
$$\Delta_h \leq 1. \tag{44}$$

We now show that $\Delta_h = 1$ and $S = (1-\Delta)\pi_h^{FB}$ solve the above program. First, when $\Delta_h = 1$ and $S = (1-\Delta)\pi_h^{FB}$, the manufacturer's effective profit share is Δ, which meets the manufacturer's minimum profit share requirement and hence constitutes the retailer's maximum profit. Second, we check that Constraints (42)–(43) are satisfied because

$$\Delta_h\pi_h^{FB} - S = \Delta\pi_h^{FB} > \Delta\pi_{hl}^* = \Delta_l\pi_{hl}^*, \text{ and}$$
$$\Delta_l\pi_l^{FB} = \Delta\pi_l^{FB} > \Delta\pi_{lh}^* = \Delta_h\pi_{lh}^* - S.$$

We therefore conclude that $\Delta_h = 1, S = (1-\Delta)\pi_h^{FB}$, and $\Delta_l = \Delta$ constitute an optimal mechanism. The first contract can therefore be implemented through a wholesale contract with full-credit returns, plus a fixed transfer payment, and the second contract can be implemented by a BR contract with parameters specified in the proposition. □

Note that, in the first contract intended for the type h manufacturer, the wholesale price is equal to the retail price minus the retailer's unit administrative cost so that the retailer is solely reimbursed for its administrative costs for selling the product; in the meantime, the manufacturer pays the retailer a fixed share of the supply chain's profit to reward the retailer for the distribution service. This resembles the "fee-for-service" (FFS) system that has recently become the dominant business model in the U.S. pharmaceutical industry (Zhao et al. 2011). Proposition 8 shows that, in the presence of asymmetric information, the optimal contract that coordinate the supply chain is a combination of FFS and BR contracts. When vaccine manufacturers' delivery performance is steady over time or known to the retailer, a BR contract is sufficient to coordinate the supply chain; when the manufacturers' delivery performance is heterogeneous, however, it is to the supply chain's benefit to introduce FFS contract to mitigate the inefficiency resulting from adverse selection.

7. Extensions

We consider several extensions to our model. §7.1 extends the outcome of early production mode to be continuous. §7.2 considers the case when the objective is to maximize the social welfare rather than the supply chain's expected profit. §7.3 incorporates delivery-improving capital investment, and characterizes the optimal allocation rules.

7.1 When Early Production Outcome is Continuous

So far we have captured the uncertainty in product design by assuming that all the units from early production will be discarded as long as the predicted design differs from the finalized design. In practice, the firm might be able to recover a proportion of the products. In the influenza vaccine industry, for example, the FDA determines the three strains used in the current year. If the manufacturer chooses a different combination of strains in the early production, then only those strains not chosen by the FDA will be discarded; the manufacturer can keep one or two, if not all, of the three strains from the early production.

We now assume that the early production will result in a continuous, uncertain outcome. Specifically, a proportion $\tilde{\theta}$—a random variable with a p.d.f. of $h(\cdot)$ and a c.d.f. of $H(\cdot)$—of all the units from the early production can be retained after the finalized design is announced. We further assume that the expected value of the proportion is $E[\tilde{\theta}] = \theta$. As such, our basic model is a special case with a two-point distribution: $\tilde{\theta} = 1$ with probability β, and $\tilde{\theta} = 0$ with probability $(1-\beta)$; the expected value is $E[\tilde{\theta}] = \beta$.

The sequence of events remains the same as shown in Figure 1, but the relationship between Q_e and Q_r is now generalized as:

$$Q_r = Q - \tilde{\theta} Q_e.$$

The manufacturer's production cost, captured by $c_r Q + (c_e - \beta c_r) Q_e$ (cf. (3)) in our main model, is now:

$$\begin{aligned} E[c_e Q_e + c_r Q_r] &= E[c_e Q_e + c_r (Q - \tilde{\theta} Q_e)] \\ &= c_e Q_e + c_r (Q - \theta Q_e) \\ &= c_r Q + (c_e - \theta c_r) Q_e. \end{aligned}$$

The expected quantity of sales, given $\tilde{\theta}$, can now be represented as

$$Z(Q, Q_e|\tilde{\theta}) = \begin{cases} \min\{Q, d\} & \text{with probability } \alpha \\ \min\{\tilde{\theta} Q_e, d\} + \min\{Q - \tilde{\theta} Q_e, (1-\gamma)(d - \tilde{\theta} Q_e)^+\} & \text{with probability } 1-\alpha. \end{cases}$$

The expected sales quantity can therefore be written as

$$\begin{aligned} E[Z(Q, Q_e)] &= E[E[Z(Q, Q_e|\tilde{\theta})]] \\ &= \alpha \left[\int_0^Q \xi dF(\xi) + \int_Q^\infty Q dF(\xi) \right] \\ &\quad + (1-\alpha) \int_0^1 \left[\int_0^{\tilde{\theta} Q_e} \xi dF(\xi) + \int_{\tilde{\theta} Q_e}^{\frac{Q - \gamma \tilde{\theta} Q_e}{1-\gamma}} [\gamma \tilde{\theta} Q_e + (1-\gamma)\xi] dF(\xi) + \int_{\frac{Q - \gamma \tilde{\theta} Q_e}{1-\gamma}}^\infty Q dF(\xi) \right] dH(\tilde{\theta}). \end{aligned}$$

Despite the more general representations of the production cost and the expected sales quantity, we can still replicate the procedures in designing the coordinating contract in §5.1 and §5.2. In other words, we can extend the BCS and BR contracts to the case in which early production yield a continuous outcome set. The following proportion is immediate from Propositions 4 and 5 by replacing β with θ:

PROPOSITION 9. *When the early production yields a continuous outcome set,*
(i) a BCS contract with

$$b^*_{BCS} = \frac{w - c_r}{p - c_r} \cdot p \text{ and } k^*_{BCS} = \frac{p - w}{p - c_r} \cdot (c_e - \theta c_r)$$

coordinates the supply chain when the retailer can observe the manufacturer's early production decision;
(ii) a BR contract with

$$b^*_{BR} = \frac{p(\theta w - c_e)}{\theta p - c_e}, \text{ and } \rho^*_{BR} = \frac{(p - w)(c_e - \theta c_r)}{(1 - \alpha)w(\theta p - c_e)}$$

coordinates the supply chain even when the manufacturer's early production decision is not observable to the retailer.

The two types of quantity-flexibility contracts, namely, QF and D-QF contracts, however, requires a new set of analysis. The results are in the same spirit as those in §4.3 and §5.3: the QF contract invariably never coordinates the supply chain, while the D-QF contract can conditionally coordinates the supply chain. Detailed proofs are available upon request.

7.2 Social Welfare

We have focused our attention on designing supply contracts that maximize the supply chain's operating income. Now we bring consumers' benefits into context and analyze the impact of supply contract design on the social welfare. Consistent with extant literature (e.g., Cho 2010), we define the social welfare as the sum of consumers' surplus and the supply chain's profit. Assume that consumers' aggregated net benefits from consuming the product is $V(z)$ when the selling quantity is $Z(Q, Q_e) = z$; $V(\cdot)$ is assumed to be a concave increasing function, reflecting the diminishing margin of social welfare improvement from a higher coverage. The expected social welfare $W(Q, Q_e)$ can be represented similar to (3) as follows:

$$W(Q, Q_e) = E[V(Z(Q, Q_e))] - [c_r Q + (c_e - \beta c_r)Q_e]. \tag{45}$$

In the influenza vaccine industry, when the initial fraction of infected population is small, Chick et al. (2009) show that the net benefits from vaccination can be represented in a piecewise-linear form. We define v as each consumer's surplus from consuming one unit of the product, and rewrite (45) as

$$W(Q, Q_e) = vE[Z(Q, Q_e)] - [c_r Q + (c_e - \beta c_r)Q_e]. \tag{46}$$

Comparing (46) with (3), we see that the expected social welfare differs from the expected supply chain profit mainly in the value associated with the expected selling quantity $Z(Q, Q_e)$. We focus on the case that $v \geq p$. Since the social planner values the units sold to end consumers more than the supply chain, we expect that the selling quantity in the social optimum is higher than in the market equilibrium; the three coordinating contract forms discussed in §5, therefore, lead to insufficient coverage because in equilibrium the manufacturer's chosen early production quantity is lower than in the social optimum.

One way to address the social welfare gap—due to the supply chain's relatively low incentive to increase the coverage of the product—is to have a third party that provides a subsidy s to the retailer—rather than the manufacturer—for each unit of products that are shipped and sold to consumers. At $s = v - p$, the retailer's total unit revenue from a consumer and the third party is exactly v, and the analysis of all the three coordinating contracts remains the same; we therefore expect that implementing one of three contracts with appropriate choice of parameters would lead to the social optimum. It is worth pointing out that the third party does not have to the government, as often suggested by the prior literature (cf. Chick et al 2008). Healthcare payers, for example, play a similar role in the case of the U.S. influenza market; the insurance coverage essentially reduces the price that the retailer charges each consumer for one unit of the product (i.e., the consumer's out-of-pocket expense) and ensure that the retailer has sufficient incentive to boost the sales amount.

7.3 Allocation of Delivery-Improving Capital Investment

As an extension to the dual-production-mode model, we consider the case in which the manufacturer and the retailer are engaged in a joint effort to improving the delivery performance for regular production (α) subject to capital investment. We denote the probability of on-time delivery for regular production without any delivery-improving investments as α_0. The amount of the required capital investment is defined by $\phi(\alpha)$, which has the following properties: (i) $\phi(\alpha_o) = 0, \phi(1) = \infty$, (ii) $d\phi(\alpha)/d\alpha > 0$, and (iii) $d^2\phi(\alpha)/d\alpha^2 > 0$. These assumptions guarantee that $\phi(\alpha)$ is a convex increasing function of α. The expected profit of supply chain can be rewritten from (3) as

$$\pi_S(Q, Q_e) = pE[Z(Q, Q_e)] - c_r Q - (c_e - \beta c_r)Q_e - \phi(\alpha). \tag{47}$$

Since its first term $pE[Z(Q, Q_e)]$ is a concave increasing function of α, (47) is a concave function of α. The optimal on-time delivery probability α^* can be determined by the first-order condition:

$$\frac{d\phi(\alpha)}{d\alpha} = p \cdot \frac{\partial E[Z(Q, Q_e)]}{\partial \alpha}, \text{ or}$$

$$\frac{d\phi(\alpha)}{d\alpha} = p\left[\int_0^Q \xi dF(\xi) + \int_Q^\infty Q dF(\xi)\right] \tag{48}$$

$$- p \cdot \beta \left[\int_0^{Q_e} \xi dF(\xi) + \int_{Q_e}^{\frac{Q-\gamma Q_e}{1-\gamma}} [\gamma Q_e + (1-\gamma)\xi] dF(\xi) + \int_{\frac{Q-\gamma Q_e}{1-\gamma}}^\infty Q dF(\xi)\right]$$

$$- p \cdot (1-\beta)\left[\int_0^{\frac{Q}{1-\gamma}} (1-\gamma)\xi dF(\xi) + \int_{\frac{Q}{1-\gamma}}^\infty Q dF(\xi)\right]. \tag{49}$$

The right-hand side of (49) consists of three parts with clear intuitions: the first term is the expected revenue when the regular production is on-time, the second term and the third term, combined together, represent the expected revenue when regular production is late. Hence the optimal on-time delivery level corresponds to the level at which the marginal capital investment is exactly equal to the expected benefits of on-time delivery.

Now the question is whether it is possible to design a contract to coordinate the supply chain. We break down this question into three goals: (i) The manufacturer and the retailer dedicate sufficient amount of resource to improve the on-time-delivery ratio to α^*, (ii) the retailer orders the right level (Q^*), and (iii) the manufacturer produces the right amount (Q_e^*) by the early production mode. In the new perspective, §5 essentially address the following problem: given α, design the right contract that motivates the retailer and the manufacturer to achieve goals (ii) and (iii) under a given α. With the added objective (i), the problem can be viewed as the optimal allocation of the capital investment between the manufacturer and the retailer.

We have the following proposition that summarizes the optimal capital investment allocation strategy. We focus our attention to the BR and BCS contracts and ignore the D-QF contract due to lack of analytical results.

PROPOSITION 10. *Let Q^{FB}, Q_e^{FB} and α^* be the first-best solutions, then*

(i) A variant of the BR contract can always coordinates the supply chain. The contract uses parameters b^* and ρ^* the same as in Proposition 5. In addition, the manufacturer and the retailer are responsible for θ and $(1-\theta)$ of the capital investment $\phi(\alpha^*)$, in which

$$\theta = \frac{\beta w - c_e}{\beta p - c_e}, \text{ and } 1-\theta = \frac{\beta(p-w)}{\beta p - c_e}.$$

(ii) If the manufacturer's early production quantity is observable to the retailer, then a variant of the BCS contract can also coordinate the supply chain. The contract uses parameters b^* and k^* the same as in Proposition 4. In addition, the manufacturer and the retailer are responsible for θ and $(1-\theta)$ of the capital investment $\phi(\alpha^*)$, in which

$$\theta = \frac{w - c_r}{p - c_r}, \text{ and } 1-\theta = \frac{p-w}{p-c_r}.$$

The proof of Proposition 10 has a clear intuition. We have shown in Propositions 4 and 5 that BCS and BR contracts coordinate the supply chain in different situations depending on whether the manufacturer's early production is observable to the retailer. Both of the two contracts effectively divide the supply chain profit between the manufacturer and the retailer, and the profit shares of the manufacturer and the retailer are b^*/p and $(1-b^*/p)$, respectively. As long as the introduction of the delivery-improving capital investment does not break the profit division between the manufacturer and the retailer, the supply chain should remain coordinated.

8. Concluding Remarks

In this paper, we study a contract design problem for a supply chain that faces uncertainties in the delivery, demand, and design of the product. The manufacturer has two production modes: the regular production mode that starts production after the design uncertainty is resolved, and the early production mode that starts production before the design is finalized; the main tradeoff is between the informational advantage of the regular production, and the delivery advantage of the early production. Our analysis reveals that without carefully designing the supply contract, a vicious cycle can be formed where the manufacturer lacks incentive to improve the on-time delivery, which leads to the retailer's shrunken order size in anticipated of lost demand, and in turns further discourages the manufacturers' delivery-performance-improving efforts.

We start by evaluating well-studied contacts in the supply chain coordination literature, and show that more complex contract forms are required to achieve the coordination of the manufacturer's production decision and the retailer's ordering decision. We also consider the scenario where the manufacturer is better informed about delivery performance than the retailer, and show that the effect of asymmetric information can effectively mitigated by a menu of screening contracts that resembles a combination of a fee-for-service contract for a manufacturer with high delivery performance, and a BR contract for a manufacturer with low delivery performance. Finally, we consider a few extensions and show that our key insights remain directionally unchanged.

Our study is primarily motivated by various complex contract forms commonly adopted in the influenza vaccine market. The value of our modeling and analysis is twofold. On the one hand, our work explains why a few complex delivery-time-dependent contracts—although never studied in the literature—are prevalent in the practice. On the other hand, the industry itself is also experimenting different contract terms aiming to improve the efficiency of the supply chain. These analytical models and insights can be used by practitioners to guide the design of supply contracts that induces improved on-time delivery performance of the supply chain.

Appendix

A. Analysis of the Scenario with a Single Production Mode

We consider a situation in which the manufacturer operates only in the regular production mode, which is essentially a newsvendor problem subject to uncertain delivery timing and demand loss due to late delivery.

Let Q denote the retailer's ordering quantity to the manufacturer. The retailer's problem is to choose Q that maximizes its expected profit

$$\pi_R(Q) = p \cdot E[\alpha \min\{Q, d\} + (1-\alpha)\min\{Q, (1-\gamma)d\}] - (w + c_o)Q.$$

Solving the above newsvender problem gives the retailer's ordering quantity Q^* characterized in the following lemma:

Lemma 4. *Under the wholesale contract and single production mode, the optimal ordering quantity Q^* satisfies*

$$\alpha \bar{F}(Q^*) + (1-\alpha)\bar{F}\left(\frac{Q^*}{1-\gamma}\right) = \frac{w + c_o}{p}. \tag{50}$$

The optimal ordering quantity Q^* increases in α but decreases in γ.

Proof. First, we show that $\pi_R(Q)$ is concave in Q by noticing that its second-order derivative in Q is negative:

$$\frac{d^2\pi_R(Q)}{dQ^2} = -p\left[\alpha f(Q) + (1-\alpha) f\left(\frac{Q}{1-\gamma}\right)\right] < 0.$$

Therefore, the optimal ordering quantity Q^* satisfy the first-order condition:

$$p\left[\alpha \bar{F}(Q^*) + (1-\alpha)\bar{F}\left(\frac{Q^*}{1-\gamma}\right)\right] = w + c_o,$$

which can be rewritten to (50).

We can show that Q^* increases in α and decreases in γ by observing that 1) the first-order derivative of $\pi_R(Q)$ in terms of Q increases in α and decreases in γ, 2) $\bar{F}(Q^*) > \bar{F}\left(\frac{Q^*}{1-\gamma}\right)$, and 3) both $\bar{F}(Q^*)$ and $\bar{F}\left(\frac{Q^*}{1-\gamma}\right)$ decrease in Q. □

The implications of Lemma 4 are two-fold. First, as the delivery becomes more unreliable (as α decreases), the retailer places a smaller order. The underlying explanation is that, if delay occurs, the demand shrinks due to loss of unfilled demand arriving during the ideal period. Given γ, a smaller α makes the retailer more likely to serve a smaller market. It is therefore reasonable for the retailer to order less from the manufacturer. Second, as consumers' time-sensitiveness increases (a larger γ), the retailer orders fewer units from the manufacturer due to a lower expected demand level. The above lemma shows that the time-sensitiveness of the demand, together with the uncertainty in the delivery timing, can create a *functional* shortage.

Lemma 5. *Under the single-production-mode setting, the first-best ordering quantity that maximizes the supply chain's expected profit satisfies*

$$\alpha \bar{F}\left(Q_s^{FB}\right) + (1-\alpha)\bar{F}\left(\frac{Q_s^{FB}}{1-\gamma}\right) = \frac{c_r + c_o}{p}. \tag{51}$$

Proof. We write the supply chain's profit as follows:

$$\pi_S(Q) = p \cdot E[Z(Q)] - (c_r + c_o)Q, \tag{52}$$

where

$$Z(Q) := \alpha \min\{Q, d\} + (1-\alpha)\min\{Q, (1-\gamma)d\}$$

is the actual sales. The lemma then follows immediately from the first-order condition. □

LEMMA 6. *Under the wholesale contract and single production mode, a buyback contract with buyback price* $b^* = \frac{w-c_r}{p-c_r-c_o} \cdot p$ *coordinates the supply chain if the manufacturer has no other production modes.*

Proof. Under a buyback contract with buyback price b, the retailer's profit function, denoted by $\pi_R^B(Q)$, is

$$\pi_R^B(Q) = p \cdot E[Z(Q)] - (w + c_o)Q + bE[Q - Z(Q)] \tag{53}$$
$$= (p - b)E[Z(Q)] - (w + c_o - b)Q. \tag{54}$$

Substituting $b = (w - c_r)/(p - c_r - c_o) \cdot p$ into (53) and comparing with (52), we have the following relationship between the retailer's and the supply chain's expected profit:

$$\pi_R^B(Q) = \frac{p - w - c_o}{p - c_r - c_o} \pi_S(Q).$$

Since the retailer's expected profit is always a proportion of the net expected profit of the supply chain, the retailer would choose Q that maximizes the supply chain profit. □

B. Data Source

According to our communication with production managers at Sanofi, the unit production costs for both early and regular production are roughly the same. We choose $c_r = c_e = \$3.0$ to be consistent with prior influenza vaccine literature, e.g., Deo and Corbett (2009) and Cho (2010). The wholesale price $w = \$12$ is chosen to be approximately the average unit price of influenza vaccine according to the CDC vaccine price list (CDC 2011c). The retail price, chosen as \$18, is the approximate average quoted price we collected from several local pharmacies including Rite Aid, CVS, Walgreen, and Giant Eagle, \$28, less the average administration cost, \$10, based on AAP (2007).

References

[1] American Academy of Pediatrics (AAP). 2007. Immunization financing: where is the breaking point? Available at `http://www.aap.org/immunization/pediatricians/pdf/TaskForceWhitePaper.pdf`.

[2] Cachon, G. 2003. Supply chain coordination with contracts. A. G. De Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science, Vol. 11. Supply Chain Management: Design, Coordination, and Operation.* Elsevier Science, Amsterdam, The Netherlands.

[3] Cachon, G., M. Lariviere. 2001. Contracting to assure supply: how to share demand forecasts in a supply chain. *Management Sci.* **47**(5) 629–646.

[4] Cachon, G. P., F. Zhang. 2006. Procuring fast delivery: sole-sourcing with information asymmetry. *Management Sci.* **52**(6) 881–896.

[5] Centers for Disease Control and Prevention (CDC). 2011a. Seasonal influenza vaccine supply and distribution in the United States. Available at: `http://www.cdc.gov/flu/about/qa/vaxdistribution.htm`. Accessed October 13, 2011.

[6] Centers for Disease Control and Prevention (CDC). 2011b. Selecting the viruses in the seasonal influenza (flu) vaccine. Available at: `http://www.cdc.gov/flu/professionals/vaccination/virusqa.htm`. Accessed October 13, 2011.

[7] Centers for Disease Control and Prevention (CDC). 2011c. CDC Vaccine Price List. Available at: `http://www.cdc.gov/vaccines/programs/vfc/cdc-vac-price-list.htm`. Accessed October 13, 2011.

[8] Chick, S. E., H. Mamani, D. Simchi-Levi. 2008. Supply chain coordination and influenza vaccination. *Oper. Res.* **56**(6) 1493–1506.

[9] Cho, S.-H. 2010. The optimal composition of influenza vaccines subject to random production yields. *Manufacturing Service Oper. Management* **12**(2) 256–277.

[10] Deo, S. C. J. Corbett. 2009. Cournot competition under yield uncertainty: The case of the u.s. influenza vaccine market. *Manufacturing Service Oper. Management* **11** 563–576.

[11] Donohue, K.L., 2000. Efficient supply contracts for fashion goods with forecast updating and two production modes. *Management Sci.* **46**(11) 1397–1411.

[12] Fukuda, K., D. O'Mara, J. A. Singleton. 2002. How the delayed distribution of influenza vaccine created shortages in 2000 and 2001. *Pharm. Ther.* **27** 235–242.

[13] Grout, J. R., D. P. Christy. 1993. An inventory model of incentives for on-time delivery in Just-In-Time purchasing contracts. *Naval Res. Logistics* **40** 863–877.

[14] Jackson, C. 2009. Integrated design-to-delivery for fast fashion: on time, on trend, and on profit. Aberdeen Group, Boston, MA.

[15] Katok, E., D. Thomas, A. Davis. 2008. Inventory service-level agreements as coordination mechanisms: the effect of review periods. *Manufacturing Service Oper. Management* **10**(4) 609–624.

[16] Laffont, J.-J., D. Martimort. 2002. *The Theory of Incentives.* Princeton University Press, Princeton, NJ.

[17] Matthews, J. T. 2006. Egg-based production of influenza vaccine: 30 years of commercial experience. *The Bridge* **36**(3) 17–24.

[18] Nowalk, M. P., R. K. Zimmerman, S. M. Cleary, R. D. Bruehlman. Missed opportunities to vaccinate older adults in primary care. *J. Amr. Board Fam. Pract.* **18**(1) 20–27.

[19] O'Mara, D., K. Fukuda, J. A. Singleton. 2003. Influenza vaccine: ensuring timely and adequate supply. *Infect. Med.* **20** 548–554.

[20] Plambeck, E., T. Taylor. 2005. Sell the plant? The impact of contract manufacturing on innovation, capacity and profitability. *Management Sci.* **51**(1) 133–150.

[21] Stecke, K. E., X. Zhao. 2007. Production and transportation integration for a make-to-order manufacturing company with a commit-to-delivery business mode. *Manufacturing Service Oper. Management* **9**(2) 206–224.

[22] Tsay, A. 1999. Quantity-flexibility contract and supplier-customer incentives. *Management Sci.* **45**(10) 1339–1358.

[23] Taylor, T. A. 2001. Channel coordination under price protection, midlife returns, and end-of-life returns in dynamic markets. *Management Sci.* **47**(9) 1220–1234.

[24] Taylor, T. A. 2002. Supply chain coordination under channel rebates with sales effort effects. *Management Sci.* **48**(8) 992–1007.

[25] Taylor, T. A., E. L. Plambeck. 2007. Simple relational contracts to motivate capacity investment. *Manufacturing Service Oper. Management* **9**(1) 94–113.

[26] Taylor T. A., W. Xiao. 2009. Incentives for retailer forecasting: rebates vs. returns *Management Sci.* **55**(10), 1654–1669.

[27] Vaccines and Related Biological Products Advisory Committee (VRBPAC). 2007. Transcripts and other meeting documents. `http://www.fda.gov/cber/advisory/vrbp/vrbpmain.htm`.

[28] Yano, C.A., H. L. Lee. 1995. Lot Sizing with Random Yields: a Review. *Oper. Res.* **43**(2) 311–334.

[29] Zhao, H., C. Xiong, S. Gavirneni, A. Fein. 2011. Fee-for-service contracts in pharmaceutical distribution supply chains: design, analysis, and management. *Manufacturing Service Oper. Management.* Forthcoming.